SYSTEM LIMITATION EXPERIENCES BY SWEDISH DRIVERS USING ACC AND LKA

Maria Klingegård,
Folksam
Sweden

Annika Larsson,
Veoneer
Sweden

Paper number: 23-0068

ABSTRACT

For many ADAS to reach its full safety potential they need to be activated and used by its drivers. There are thus several known (technical) limitations that could, as indicated by research, potentially affect the perception and use of the ADAS. This paper explores limitations as experienced by users for the lateral assistance systems Adaptive Cruise control (ACC) and Lane Keep assist (LKA). The paper partly reports on a larger online survey launched (n=1822) in 2021 aimed to explore self-reported use and non-use of six different ADAS among Swedish drivers using a 5-point Likert scale. Descriptive statistics including frequencies and a calculated summative level of agreement % is presented together with 95% confidence levels. Included in the analysis is those respondents reporting using ACC (n=1002), and/or LKA (n=461). Presented are limitations as experienced, frequency of use/non-use, and perceived driving experience. Results show that ACC is being activated (always/often) to a greater extent (84%) than LKA (57%), and for LKA it varies by frequency of driving. The majority of the participants had experienced more than one limitation (ACC:72%, LKA:68%), on a regular basis, which results in deactivation of the system. Only about 20 % (ACC:20%, LKA 18%) had never experienced that they could not use the ADAS. Those who do not experience any limitations, never experience the need to deactivate the ADAS to a greater extent- ACC: (38% vs 22%) and LKA (48% vs. 23%). Statistical significant tests relived a significant difference between LKA and ACC, in which LKA was affected to a greater extent for bad weather (48%), glare (48%), position in lane (27%), complex traffic (27%) while ACC was affected to a greater extent by dirty sensors (45%), complex traffic (43%), weather (31%). ACC also contribute (significantly) to a higher degree to a positive driving experience than LKA, likewise are more trusted and easier to use. This study highlights some of the reasons why ADAS are regularly turned off, diminishing their safety potential. Technological developments, together with standardization and infrastructure adaptation, may be required for ADAS to fully realize their safety potential.

INTRODUCTION

Today, vehicles often include systems that can help the driver steer, break, and keep distance to the vehicle ahead as well as warn when there is a crash risk. The introduction of advanced driver assistance systems, ADAS, (see table 1) is believed to have great potential for decreasing the number of fatalities in traffic [1], currently the 8^{th} most common cause of death in the world with 1.35 million lives lost every year [2]. Research shows that vehicle crash safety has increased steadily since the 1980s [3], but the number of crashes also need to go down. Growing evidence from simulation studies, field operational tests and crash data analysis demonstrates that ADAS, individually and together, increase the safety of the vehicle and decrease the risk of crashes (e.g.,[4], [5]).

ADAS such as adaptive cruise control (ACC) and lane keep assist (LKA) are systems which, when active continuously help the driver maintain distance to vehicle ahead and/or in their lane, see Table 1. These systems are often referred to as 'comfort' systems. Yet, LKA could potentially have safety benefits as they could reduce the risk of running off road, drifting into oncoming vehicles and side swipes. ACC could also potentially contribute to a higher could reduce the potential for rear-end crashes. Indeed, it has been shown that forward collision warning with break support (CWB) combined with ACC, reduced rear-end crashes with frontal impacts with 38% [6]. Studies have also shown that LKA-equipped vehicles were 9% less likely (HR=0.91) to run off the road [7]. LKA did not have a significant effect on risk of same-direction sideswipes or head-on crashes [7].

However, these lateral ADAS are not without (technical) limitations [8], potentially resulting in less usage. Known limitations typically communicated to drivers includes lane division unclear, lane markings damage, sensors obstructed/damages, very hot/cold temperature [8]. However, there are large difference between different vehicle manufacturers. Other challenges impacting the system's ability to function such as adverse weather, bright light is rarely communicated [8]. Other limitations could be narrow, winding, and sloping roads, or construction zones. In addition, there are the system constrains for which the ADAS are designed to operate within (e.g., vehicle speed, road type).

To be safely used drivers need to be aware of its limitations. Trust and correct understanding of system functionality are considered key variables for appropriate system use [9]. It is encouraged that limitations "should be clearly defined and effectively communicated to the driver, and that drivers should be unable to engage the systems outside of the ODD [Operational Design Domain]" [10].

The majority of people participating in surveys express a positive attitude toward ADAS [11, 12]. However, studies have shown that the frequency of ADAS usage vary ([13-17], and the knowledge of their presence [16] or technical limitations vary [18, 19]. As drivers learn to use ADAS, they become more aware of the limitations with time, with unwanted system actions such as harsh responses to cut-ins, limits in maximum brake force, and limits in lead vehicle detection [20]. There is also a potential connection between low use of assistance systems, though available, and a belief that systems will not provide much of a benefit [21]. Also, research indicate that when learning about the ADAS focus on its limitations results in negative bias towards ADAS [22], indeed it has been shown that the quantity and quality of device-specific feature systematically affected drivers perception [23].

There are thus several known (technical) limitations that could potentially affect the perception and use of the ADAS. This paper report on a survey exploring ADAS usage and the limitations as experienced for ACC and LKA by Swedish drivers, in winter, known conditions to impact the ability to use the systems.

Table 1 Schematic overview of ADAS system from an accident prevention perspective derived from a driver's point of view. Developed from [24, 25].

Traffic situation	***Inform*** Static/temporal information	***Warn*** Temporal information via sound, graphic, or haptics	***Act*** Brake, limit engine power, and/or steer	
Reverse and park	Camera feed Parking sensors	Object detection	Rear-AEB	Parking assistance
Distance and speed	Set speed Road speed	Speed warning	Intelligent speed assistance	Cruise Control **Adaptive Cruise Control (ACC)**
Crash avoidance and mitigation		Collision warning Distance warning	AEB car/pedestrian /cyclist Emergency steering	**Lane Keep Assistance (LKA)** Assistance driving*
Safe in lane		Lane Departure Warning Blind spot detection		
Driving safe		Driver Monitoring Systems	Alcohol lock	

*Lane centring with Adaptive Cruise Control, e.g., piloting functions

METHOD

Research aims and objectives

The aim of the paper is to provide an overview of limitations as experienced by Swedish drivers and explore the possible outcome in terms of none/usage and perceived driving experience. More specifically the objective is to: (1) present descriptive statistics on limitations as experienced, frequency of use, driving experience and perceived benefits, (2) identify potential differences between the experienced limitations between LKA and ACC. The goal is to determine if the limitations as experienced influence the usage of ADAS.

Digital survey design

The digital survey was engineered using Netigate software. Survey design was based on previous studies examining the usage of ADAS (e.g., [26]). A total of 6 ADAS was included: Adaptive Cruise Control (ACC), Lane Keeping Assistance (LKA), Lane Departure Warning (LDW), Blind Spot Detection (BLIS), Forward

Collision Warning with or without Automatic Emergency Break for vehicle (AEB car) or pedestrian/cyclist detection (FCW/AEB VRU). After a set of demographics questions, respondents were asked about system availability in their vehicle(s). After system availability, each ADAS was presented separately with corresponding questions in which the driver were asked to respond to a set of statement via a 5-point Likert scale. The respondents were asked questions only about the systems they expressed they had in their vehicle and were reporting using. The survey was expected to take up to 15 minutes to complete.

Distribution

The survey was digitally distributed during March 2021 via a social media advertisement (Facebook) by Folksam Insurance Company: "Help us in our research on driver support system". The Facebook campaign had a reach of 144 300 and about 3985 unique hits. The demographics of the distribution was Sweden and people above 18, with no further specification regarding interests or group memberships.

The starting page of the questionnaire explained and asked those only within the target group to continue: Swedish drivers with a vehicle no older than model year 2009, with at least one of the following ADAS; Adaptive Cruise Control (ACC), Lane Keeping Assistance (LKA), Lane Departure Warning (LDW), Blind Spot Detection (BLIS), Forward Collision Warning with or without Automatic Emergency Break for vehicle (AEB car) or pedestrian/cyclist detection (FCW/AEB VRU). If the respondents chose that they did not have experience of any of the system, the survey ended.

Respondents

A total of 2521 participants started the survey. If the respondents did not have any of the requested systems and/or if the respondent did not answer all questions in the survey, then they were excluded from the final dataset (Table 2). A total of 1153 respondents reported having the ADAS on their current car: ACC (n=1113) and LKA (n=636). In this paper only respondents who explicitly stated that they use the ADAS are included LKA (n=461) and ACC (n=1002)). A total of 37 (4%) of the ACC respondents reported only having experience of ACC. It should be noted that the respondents of the questionnaire do not represent the total population of drivers Sweden (Table 3).

Table 2 Overview of the respondents' experienced ADAS. Includes answer from question: "Do you have experience of the following ADAS".

ADAS	*Respondents who have the system*	*Respondents who have the system and use it*	*Respondents who have the system, but do not use it*	*Respondents who do not have the system*	*Respondents who do not know if they have it*	*Included in analysis*
ACC	1113 (61%)	1002 (55%)	111 (6%)	683 (38%)	26 (2%)	1002
LDW	997 (55%)	754 (41%)	243 (13%)	795 (44%)	30 (2%)	
LKA	636 (35%)	461 (25%)	175 (9%)	1117(61%)	69 (4%)	461
BLIS	718 (39%)	680 (37%)	38 (2%)	1051 (58%)	53 (3%)	
AEB car	1214 (67%)	1163 (64%)	51 (3%)	569 (31%)	39 (2%)	
AEB VRU	868 (48%)	841 (46%)	27 (2%)	763 (42%)	191 (11%)	
Total	1822	1822				1039

Table 3 Representativeness of population. Comparison of distribution between percent of respondents (Re), number of persons holding a driving license (Dr) and number of car owner (Ow) in Sweden. Presented statistics is based on Swedish official population data provided by Statistics Sweden (SCB), 2022.

Gender/Age	*18-29y*			*30-39y*			*40-49y*			*50-59y*			*60y or older*		
	Re	Dr	Ow	Re	Dr	Ow	Re	Dr	Ow	Re	Dr	Ow	Re	Dr	Ow
Female	11	15	7	17	16	15	31	16	19	20	17	24	21	37	36
Male	7	15	8	15	17	16	22	16	19	23	18	22	33	34	35
Total	8	15	7	15	16	16	23	16	19	23	17	23	31	35	35

Analysis
The dataset includes responses to the experience and usage of ACC and LKA. Reported are respondent's demographics, frequency of use, limitations as experienced, perceived benefits, and positive driving experience. Descriptive statistics (frequency count and proportion) were calculated per survey item. A summative level of agreement response was calculated by adding 4-5 (coded as 1) on the Likert scale. Statistical analyses include the proportions and difference of proportions with 95% confidence limits, CL. The CL for a proportion is calculated with assumption of simple normal approximation binomial intervals. The CL for a difference is calculated with the same assumption. The Z-statistics is only calculated if n1p1(1-p1) >= 9 och n2p2(1-p2) >= 9. No correction for finite populations has been done. Statistical tests include statistically significant differences between ACC and LKA. Excel Power Pivot (v. 2108) and SAS Enterprise Guide (v. 8.3.0.103) were used for statistical analysis.

RESULTS

Respondent demographics

Of the 1822 respondents who completed the questionnaire, a total of 1039 are included of which 1002 respondents had ACC (96%) and 461 respondents had LKA (44%). The demographics of the respondents are presented in Appendix, Table A1. The majority of the respondents were 50 years or older (ACC: 53%; 55%), and male (ACC: 85%, LKA: 87%). The majority of the respondents consider themselves less prone to take risks with a mean between 3,7-3,5 on a 10 point Likert scale; the majority indicated 3 or less on the scale (ACC: 54%, LKA: 55%). On the technology readiness scale the majority identified themselves as early adopter (ACC: 56%, LKA: 62%)) or early majority (ACC:33%, LKA 27%). Most of the respondents were positive towards using drivers support systems in general (ACC: 70%, LKA 75%). The majority lived in urban areas (55%), while the rest lived either in a large city (Stockholm/ Göteborg/ Malmö) (ACC: 24%, LKA:26%) or rural area (ACC: 20%, LKA: 19%). The majority of respondents lived in the south of Sweden which corresponds to the population concentration in Sweden. Most respondents had 5 ADAS or more (ACC: 49%, LKA,83%). A total of 37 (4%) of the ACC respondents reported only having experience of ACC.

Limitations as experienced by respondents

The majority of the respondents had experienced limitations and situations in which they could not activate the ADAS (ACC:72%, LKA:68%), table 4. Only about 20% (ACC:20%, LKA 18%) had never experienced any situation for which they could not use the ADAS. For ACC dirty sensors (45%) and complex traffic situation (43%) is commonly experienced. For LKA weather (48%) and glare (45%) is commonly experienced. Only about 25% had never experienced that they needed to deactivate the system due to negative driving experience (ACC: 25%, LKA:26%), table 6. Only a limited amount of people had often or always experienced that they had to turn off the system due to a negative experience (ACC: 3%, LKA: 6%), table 6. Statistical significant tests relived a significant difference between LKA and ACC, in which LKA was affected to a greater extent for bad weather, darkness, and glare, while ACC was affected to a greater extent by complex traffic situation, dirty sensors (Table, 4). A majority of the respondents had experience more than one limitation, see table 5. Those who did not experience any limitations (ACC: n=177, LKA n=52), never experience the need to deactivate the ADAS to a greater extent compared to those that experienced at least one limitation- ACC: (38% vs 22%) and LKA (48% vs. 23%). Subsequent statistical analysis shows a statistical significance between ACC and LKA, table 8 and 9. When excluding those who only had experience one of the systems, there are no significant difference in frequency (0-5) experienced limitations between ACC and LKA (cf., table 5).

Table 4 Specification of experienced situations the respondents had to turn off the system even though they wanted to use it. The respondents could answer one or more situations.

Reason to deactivate ADAS	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
*Total, turning off**	726 (72%)	312 (68%)	4.78	-0.31 – 9.86	0.0616
Weather	315 (31%)	221 (48%)	-16.20	-21.60 – -10.81	<.0001***
Darkness	68 (7%)	95 (21%)	-13.82	-17.83 – -9.81	<.0001***
Glare	62 (6%)	208(45%)	-38.93	-43.71 – -34.15	<.0001***
Positioning in lane	n/a	124 (27%)	n/a		

Complex traffic	432 (43%)	124 (27%)	16.22	11.14 – 21.29	<.0001***
Dirty sensors	451 (45%)	60 (13%)	31.99	27.64 – 36.34	<.0001***
No situations	203 (20%)	82 (18%)	2.47	-1.82 – 6.76	0.2674
Other	86 (9%)	33 (7%)	1.42	-1.50 – -1.50	0.3545

**due to negative driving experience / no time of turning off*

Table 5 overview of the number of limitations respondents indicated from a pre-set list (available alternatives: bad weather, darkness, glare, complex traffic situation, dirty sensors, no experienced situations, position in lane, other), respondents who indicated "no situation" is denoted as "0". Respondents could indicate one or more alternatives.

No. ADAS Limitations	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
0	177(18%)	34(7%)	10.29	6.93 – 13.65	<.0001***
1	421(42%)	219(48%)	-5.49	-10.98 – 0.00	0.0493*
2	272(27%)	125(27%)	0.03	-4.87 – 4.93	0.9902
3	92(9%)	62(13%)	-4.27	-7.86 – -0.68	0.0135*
4	27(3%)	19(4%)	-1.43	-3.50 – 0.65	0.1463
5	13(1%)	2(0%)	0.86	-0.06 – 1.79	1.0000
6	0(0%)	0(0%)	0.00	-0.00 – 0.00	1.0000

Table 6. Do you deactivate the system as it negatively contribute to your driving experience divided according to number of limitations as experienced.

Deactivate of ADAS	ACC 0 limitations experienced (n=177)	ACC 1-5 limitations experienced (n=825)	LKA 0 limitations experienced (n=52)	LKA 1-5 limitations experienced (n=409)
Always	2(1%)	11(1%)	0(0%)	7(23%)
Often	3(2%)	40 (5%)	2(4%)	34(8%)
Sometime	43(24%)	280 (34%)	10 (19%)	153 (37%)
Rarely	57(32%)	290 (35%)	10 (19%)	96(24%)
Never	68 (38%)	184 (22%)	25(48%)	95 (23%)

Table 7. Respondents experience the need to deactivate the system due to negative experience.

Deactivation due to negative driving experience	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
Never	252 (25%)	120 (26%)	-23.44	-28.74 – -18.15	<.0001***
Rarely	347 (34%)	106 (23%)	13.37	8.62 – 18.13	<.0001***
Sometime	323 (32%)	163 (35%)	11.85	7.17 – 16.53	<.0001***
Often	43(4%)	36 (8%)	0.60	-1.53 – 2.73	0.5885
Always	13 (1%)	7 (1%)	0.43	-0.67 – 1.53	1.0000
Cannot answer	24 (2%)	26 (6%)			

Table 8. Do you deactivate the ACC as it negatively contribute to your driving experience divided according to number of experienced limitations.

Deactivate of ADAS	ACC 0 experienced limitations (n=177)	ACC 1-5 experienced limitations (n=825)	Difference in proportion [%]	95% CI for difference	Significance level (p-value)
Always	2(1%)	11(1%)	-0.20	-1.95 – 1.54	1.0000
Often	3(2%)	40 (5%)	-3.15	-5.55 – -0.75	1.0000
Sometime	43(24%)	280 (34%)	-9.95	-16.74 – -2.55	0.0127*
Rarely	57(32%)	290 (35%)	-2.95	-10.56 – 4.67	0.4545
Never	68 (38%)	184 (22%)	16.12	8.41 – 23.82	<.0001***

Table 9. Do you deactivate the LKA as it negatively contribute to your driving experience divided according to number of experienced limitations.

Deactivate of ADAS	LKA 0 experienced limitations (n=52)	LKA 1-5 experienced limitations (n=409)	Difference in proportion [%]	95% CI for difference	Significance level (p-value)
Always	0(0%)	7(23%)	-1.71	-2.97 – -0.45	1.0000
Often	2(4%)	34(8%)	-4.47	-10.34 – 1.41	1.0000

Sometime	10 (19%)	153 (37%)	-18.18	-29.87 – -6.48	1.0000
Rarely	10 (19%)	96(24%)	-4.24	-15.71 – 7.23	1.0000
Never	25(48%)	95 (23%)	24.85	10.67 –39.03	0.0001**

Activation of ADAS

The respondents were also asked to indicate how often they used a particular ADAS when driving, Table 10. The result indicates that the participants use the system often (ACC: 37%, LKA: 25%) or always (ACC: 47%, LKA:32%) with ACC being used to a greater extent than LKA (p <.0001). The use of ADAS also varies by driving frequency for LKA (% agreement of activation of ADAS increase by the frequency of drive) but are more stable across the respondents for ACC (ranging between 83-90%), figure 1, table 12-13.

The respondents were asked in which specific traffic conditions they felt comfortable using the ADAS, Table 11. Most respondents were comfortable to use the system on highways (ACC 97%, LKA: 88%). Fewer respondents feel comfortable using ADAS near roadworks (ACC: 17%, LKA:8%) and on curvy roads (ACC: 32%, LKA 26%). There is also lower usage in high intensity traffic (ACC: 41%, LKA: 32%). ACC and LKA follow a similar pattern, but LKA consistently receives lower scores in each traffic condition. The difference is statistically significant for the different attributes (p<0.05).

Table 10. Activation of ADAS. Includes answer from question: How often do you use the ADAS in your current used vehicle?

Frequency of activation	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
% agreeness (always/often)	84%	57%			
Always	474 (47%)	148 (32%)	15.20	9.94 – 20.47	<.0001***
Often	371 (37%)	113 (25%)	12.51	7.58 –17.45	<.0001***
Sometime	128 (13%)	134 (29%)	-16.29	-20.92 – -11.66	<.0001***
Rarely	24 (2%)	43 (9%)	-6.93	-9.75 – -4.11	<.0001***
No knowledge	5 (1%)	23 (5%)	-4.49	-6.53 – -2.46	1.0000

Table 11 Respondents comfortable using ADAS in different traffic environments. Includes answer to question: on what roads/traffic conditions are you comfortable to use the system?

Type of road	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
Country roads	913 (91%)	375 (81%)	9.77	5.80 – 13.74	<.0001***
City streets	368 (37%)	113 (25%)	12.21	7.28 – 17.15	<.0001***
Highway	968 (97%)	405 (88%)	8.75	5.57 – 11.94	<.0001***
Roads with separated lanes (2+1 lanes)	833 (83%)	321 (70%)	13.50	8.71 – 18.30	<.0001***
curvy roads	316 (32%)	118 (26%)	5.94	1.03 – 10.85	0.0208*
Roads with traffic lights and/or round about	229 (23%)	65 (14%)	8.75	4.65 – 12.86	0.0001**
Low intensity traffic	567 (57%)	203 (44%)	12.55	7.08 – 18.03	<.0001***
High intensity traffic	410 (41%)	149 (32%)	8.60	3.35 – 13.84	0.0017*
Roadwork	167 (17%)	35 (8%)	9.07	5.73 – 12.42	<.0001***
Slow traffic	552 (55%)	159 (34%)	20.60	15.28 – 25.92	<.0001***

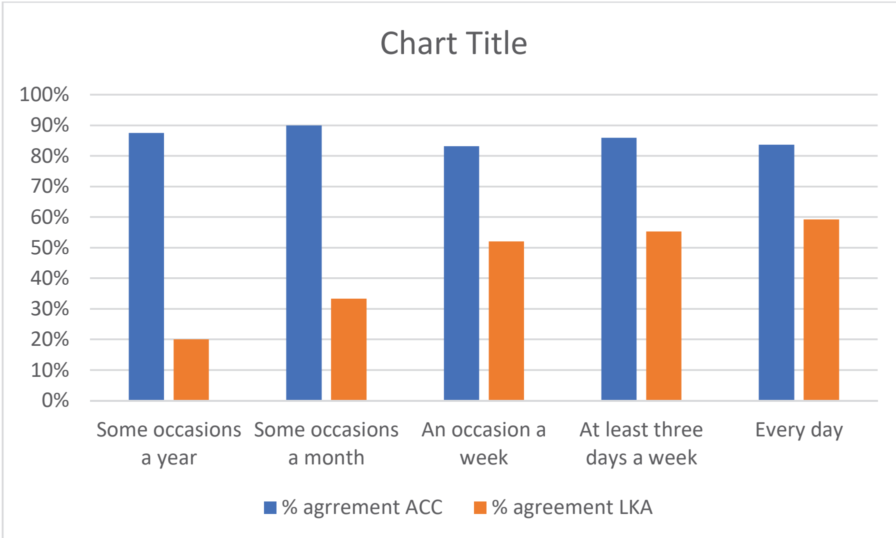


Figure 1 Overview of frequency of drive and activation of ACC and LKA (% agreement: (often/always) activated). See table 12, 13.

Table 12 Activation of ACC vs driving frequencies

Frequency of activation	*Some occasions a year*	95% CI	*Some occasions a month*	95% CI	*An occasion a week*	95% CI	*At least three days a week*	95% CI	*Every day*	95% CI	*Total*	95% CI
Always	2 (0%)	-0.08 – 0.48	8 (1%)	0.25 – 1.35	56 (6%)	4.17 – 7.01	120 (12%)	9.97 – 13.99	288 (29%)	25.94 – 31.54	474 (47%)	44.21 – 50.40
Often	5 (1%)	0.06 – 0.94	10 (1%)	0.25 – 1.35	43 (4%)	3.04 – 5.55	106 (11%)	8.67 – 12.48	207 (21%)	18.15 – 23.17	371 (37%)	34.04 – 40.02
Sometime	0 (0%)		2 (0%)	-0.08 – 0.48	18 (2%)	0.97 – 2.62	27 (3%)	1.69 – 3.70	81 (8%)	6.40 – 9.77	128 (13%)	10.71 – 14.84
Rarely	1 (0%)	-0.10 – 0.30	0 (0%)		2 (0%)		8 (1%)	0.25 – 1.35	13 (1%)	0.60 – 2.00	24 (2%)	1.45 – 3.34
No knowledge	0 (0%)		0 (0%)		0 (0%)		2 (0%)		3 (0%)		5 (1%)	0.06 – 0.94
No answer	0 (0%)		0 (0%)		0 (0%)		0 (0%)		0 (0%)		0 (0%)	
Total	8 (1%)	0.25 – 1.35	20 (2%)	1.13 – 2.86	119 (12%)	9.87 – 13.88	263 (26%)	23.52 – 28.97	592 (59%)	56.04 – 62.13	1002 (100%)	

Table 13. Activation of LKA vs driving frequencies

Frequency of activation	*Some occasions a year*	95% CI	*Some occasions a month*	95% CI	*An occasion a week*	95% CI	*At least three days a week*	95% CI	*Every day*	95% CI	*Total*	95% CI
Always	1 (0%)	-0.21 – 0.64	0 (0%)		15 (3%)	1.63 – 4.87	47 (10%)	7.43 – 12.90	85 (18%)	14.90 – 21.98	148 (32%)	27.84 – 36.37
Often	0 (0%)		2 (0%)	-0.17 – 1.03	10 (2%)	0.84 – 3.50	26 (6%)	5.36 – 10.26	75 (16%)	12.90 – 19.64	113 (25%)	20.59 – 28.44
Sometime	3 (1%)	-0.08 – 1.38	3 (1%)	-0.08 – 1.38	17 (4%)	1.97 – 5.41	35 (8%)	5.17 – 10.01	76 (16%)	13.10 – 19.87	134 (29%)	24.92 – 33.21
Rarely	1 (0%)	-0.21 – 0.64	0 (0%)		3 (1%)	-0.08 – 1.38	15 (3%)	1.63 – 4.87	24 (5%)	3.18 – 7.23	43 (9%)	6.67 – 11.98
No knowledge	0 (0%)		1 (0%)	-0.21 – 0.64	3 (1%)	-0.08 – 1.38	9 (2%)	0.69 – 3.22	10 (2%)	0.84 – 3.50	23 (5%)	3.00 – 6.98
No answer	0 (0%)		0 (0%)		0 (0%)		0 (0%)		0 (0%)		0 (0%)	
Total	5 (1%)	0.14 – 2.03	6 (1%)	0.27 – 2.34	48 (10%)	7.82 – 13.20	132 (29%)	24.51 – 32.76	270 (59%)	54.07 – 63.07	461 (100%)	

Perceived driving experience

The respondents were asked if the specific ADAS contributed to a positive driving experience. ACC contributed to a larger extent to a positive driving experience as compared to LKA (ACC 86% vs. LKA 64%, p<0.0001). Considering how the system contribute to the driving experience LKA is considered to a lager extent as a safety system by the respondents compared to ACC (LKA 72% and ACC 25%). Subsequent statistical analysis show statistically significant difference (p<0.001). LKA is also considered to a greater extent to increase mental ease (p< 0.001). ACC is mainly reportedly used to decrease fuel consumption (45%). Only a minority of the respondents uses the ADAS to enable the performance of other activities that are not related to driving (ACC: 6%, LKA: 10%).

There is a significant difference in trusting the ADAS to maintain distance to vehicle in front (ACC, 88%) or maintain position in lane (LKA, 56%) (p<0.001).

Table 14. ADAS Contributes to a positive driving experience. Includes respondents answer to the question: the ADAS contribute to a positive driving experience?

Positive driver experience	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
% agreement (4/5)	853 (85%)	294 (64%)	21.36	16.45 – 26.67	<.0001***
Strongly agree (5)	535(54%)	152 (33%)	20.42	15.13 – 25.71	<.0001***
Agree to large extent (4)	318 (32%)	142(31%)	0.93	-4.17 – 6.04	0.7208
Somewhat agree (3)	122(12%)	110 (24%)	-11.69	-16.07 – -7.30	<.0001***
Disagree (2)	15 (2%)	20 (4%)	-2.84	-4.85 – -0.84	0.0010*
Strongly disagree (1)	4 (0.4%)	11 (2%)	-1.99	-3.43 – -0.54	1.0000
Cannot answer	8 (1%)	26 (6%)	-4.84	-7.02 – -2.66	1.0000

Table 15 Respondents answer to the question: How does the ADAS use contributes to the driving experience?

Type of positive experience	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion*	*95% CI for difference*	*Significance level (p-value)*

			[%]		
Increased safety	252 (25%)	333 (72%)	-47.08	-51.98 – -42.19	<.0001***
Increased physical comfort	315 (31%)	215 (47%)	-15.20	-20.59 – -9.82	<.0001***
Increased mental ease	68 (7%)	221 (48%)	-41.15	-45.97 – -36.33	<.0001***
Decreased fuel consumption	451 (45%)	45 (10%)	35.25	31.15 – 39.35	<.0001***
Opportunities to do other things	62 (6%)	44 (10%)	-3.36	-6.43 – -0.29	0.0214*
Becoming a better driver	432 (43%)	125 (27%)	16.00	10.91 – 21.09	<.0001***
More enjoyable driving	203 (20%)	42 (9%)	11.15	7.53 – 14.77	<.0001***
Other	86 (9%)	36 (8%)	0.77	-2.23 – 3.77	0.6190

Table 16 Overall experience with ACC and LKA

Type of experience	*ACC (n=1002)*	*LKA (n=461)*	*Difference in proportion [%]*	*95% CI for difference*	*Significance level (p-value)*
Trust: Keeps speed and distance to vehicle in front / keep the car in the middle of lane	884 (88%)	256 (56%)	32.69	27.74 – 37.65	<.0001***
Accelerates and brakes smoothly	798 (80%)	n/a			
Been helpful in dangerous situations	448 (45%)	106 (23%)	21.72	16.79 – 26.64	<.0001***
Increases risk to be in dangerous situations	53 (5%)	23 (5%)	0.30	-2.12 – 2.72	0.8100
Have good understanding of function of system	956 (95%)	378 (82%)	13.41	9.67 – 17.15	<.0001***
Is easy to use	952 (95%)	375 (81%)	13.67	9.86 – 17.47	<.0001***
Good collaboration in driving the car	780 (78%)	n/a			
Fights against the system	n/a	32 (7%)			

DISCUSSION

Studies have shown that many ADAS, both individually and in combination, can increase the safety of a vehicle and reduce the risk of personal injury, e.g., [4, 5, 27, 28]. However, there are known limitations to ADAS that can impact their effectiveness [8]. A combination of technical limitations in sensor technology and variations in driving conditions reduce their ability to be used.

This study confirms previous research that ADAS are being used in daily driving, with Adaptive Cruise Control (ACC) being used (significantly) more frequently than Lane Keeping Assist (LKA). However, the study highlights that most respondents experienced limitations and the respondents regularly turn off the systems, limiting their potential effectiveness. Research has indicated that limitations can affect the use and perceived value of ADAS [13]. Even the naming of the systems influences the driver's perception of its capabilities [29].

Previous studies have listed possible technical limitations that could limit the use of ADAS [8]. This study revealed a significant difference between LKA and ACC in terms of situations for which the driver experienced that they could not use the system (i.e., limitations as experienced): LKA was more affected by bad weather (48%), glare (48%), position in lane (27%), and complex traffic (27%), while ACC was more affected by dirty sensors (45%), complex traffic (43%), and weather (31%).

Previous studies have shown that bad weather conditions had no substantial impact on driving behaviour (e.g., frequency of activation) [30]. This study shows that 31% of the ACC respondents and almost 50% of the LKA respondents were limited by bad weather. However, as the author note in [30], the absence of results may be due to the absence of extreme weather during the test period. ADAS performance in adverse weather and different light conditions will come into focus in the coming years, as EuroNCAP have released their Vision 2030 with an increased effort of testing systems in a multitude of conditions [31].

As the study of [30], and this study, complex traffic is experienced as a limited factor more frequently than bad weather, but, in this study it was only true for ACC (not LKA). Previous research has identify that traffic conditions to be the most critical part of the driving context [30]. The study shows that respondents, as previously indicated, are comfortable using ACC and LKA on highway (ACC: 97%, 88%) and country roads (ACC: 91%, LKA: 81%). Less on curvy roads and by road works. This study thus show that many respondents are affected by the driving context (e.g., road type, traffic intensity) and that it limits the use of ADAS. Interestingly, as much as 27% of the respondents using LKA have turned of the system due to position in lane. 7% of the LKA respondents consider that they have to argue with the system. Previous research has identified that: "Situations where drivers reported feeling uncomfortable with the automation during their drive were dominated by instances where lane centring struggled with common roadway features such as hills and intersections"[26]. This research confirms that the respondents feel less comfortable using the systems on curvy roads (ACC:32%, LKA: 26%); most participants feel comfortable using the systems on Personalisation, and a greater flexibility in the system may be required for decreasing the frequency of experience of this limitation.

Previous studies have shown a variety of frequency of use vary [13-17]. In this study the majority of the respondents uses the systems often or always (ACC:85%, LKA 64%). But we also see that a majority of the respondents had experienced the need to turn off the ADAS system due to a negative driving experience (only 25% for ACC and 26% for LKA had never experienced the need to turn it off). The frequency of turning off the system also significantly varies depending on the specific ADAS technology being used. Previous research has indiated a "strong relationship between system activation and the capability to prevent lane drifts and the timing of steering input" [32]. In this study we see that ACC positively contributes to the driving experience (85%) and is experienced to accelerate and break smoothly (80%), and is experienced to have a good collaboration (78%).

The benefits of ADAS to the driving experience have been previously identified as a factor influencing its usage [21]. This study highlights that of the choses given (table 13): ACC are being used to decrease fuel consumption (ACC: 45%) and to become a better driver (43%), while LKA is being used to increase safety (71%), physical comport (47%), mental ease (48%).

For these systems to be safely used, research has highlighted the need that drivers understand their limitations [18-19]. Previous research has demonstrated that there is a lack of awareness or understanding of key limitations in ADAS [12]. However, the respondents in this survey judge themselves having good understanding of the function of system (ACC: 95%, LKA 82%). The result from the presented survey indicates that even though respondents experience limitations in their use of the systems, they believe that the systems positively contribute to the driving experience; at least for ACC (ACC:85%, LKA 64%). Previous research has identified pleasantness of use and perceived benefits as most important factors determine the use of ADAS [23].

The results from this paper show that the trust towards the ACC (88%) is higher than LKA (56%) with a significant difference. Previous research has shown that "automation failures do not negatively affect trust and acceptance if they are known beforehand"[9]. This research indicates no significant difference in experienced limitations (comparing 0 limitations vs. 1-5 limitations) for those who turned off the system at least once due to negative driving experience. This study shows that 95% of those using ACC consider that they have a good understanding of system functionality, significant lower for LKA (82%). Trust and correct understanding of system functionality are considered key variables for appropriate system use [9]. Due to sensor limitations, not every situation can be handled by the system and, therefore, driver intervention is required.

Throughout this study it is shown that ACC significantly differ from LKA and, ACC consistently receives better scores. Previous research has identified a difference between LKA and ACC. For instance, the study by [26] indicated that "drivers reported significantly higher trust in adaptive cruise control than in lane centering".

There is thus more work needed, especially for LKA. Experienced limitations influence the frequency of use. Future studies include identifying the effect of respondent demographics, attitude towards the system and the frequency of experience of limitations. Future studies should also include a wider population to better represent the total of Swedish drivers.

Limitations

The study was distributed via social media and based on self-reported experiences, and though care has been made to describe systems clearly, there may still be some confusion. Also, one should take care with the results as the respondents cannot be considered to be representative of all Swedish drivers.

CONCLUSION

Studies have shown that many ADAS, on their own and together, increase the safety of the vehicle and lower the risk of personal injury. However, for the systems to fulfil their safety potential they need to be used. This study highlights that the majority of the respondents experience limitations, and they regularly experience the need to turn off the systems. The study reveals that the limitations as experienced significantly varies depending on the specific ADAS technology. For ADAS to fully realize their safety potential, technological advancements, standardization efforts, and infrastructure adaptations may be necessary. This study is based on a self-reported survey and may not represent the view of all drivers in Sweden.

REFERENCES

1. Lubbe, N., et al. *Predicted road traffic fatalities in Germany: The potential and limitations of vehicle safety technologies from passive safety to highly automated driving*. in *Proceedings of IRCOBI conference*. 2018. Athens, Greece.
2. WHO, *Global Status Report on Road Safety 2018*. 2018, World Health Organization: Geneva, Switzerland.
3. Kullgren, A., H. Stigson, and A. Axelsson, *Developments in car crash safety since the 1980s* in *IRCOBI*. 2020. p. 86-99.
4. Cicchino, J.B., *Effectiveness of forward collision warning and autonomous emergency braking systems in reducing front-to-rear crash rates.* Accident Analysis & Prevention, 2017. **99**: p. 142-152.
5. Sternlund, S., et al., *The effectiveness of lane departure warning systems—A reduction in real-world passenger car injury crashes.* Traffic Injury Prevention, 2017. **18**(2): p. 225-229.
6. Isaksson-Hellman, I. and M. Lindman, *Evaluation of Rear-End Collision Avoidance Technologies based on Real World Crash Data*, in *Proc. 3rd International Symposium on Future Active Safety Technology towards zero traffic accidents*. 2015: Gothenburg. p. 471-476.
7. Spicer, R., Vahabaghaie, A., Murakhovsky, D., Bahouth, G., Drayer, B., Lawrence, S, *Effectiveness of Advanced Driver Assistance Systems in Preventing System-Relevant Crashes.* SAE Int. J. Adv. & Curr. Prac. in Mobility 2021. **3**(4): p. 1697-1701.
8. Capallera, M., Meteier, Q., de Salis, E., Angelini, L., Carrino, S., Kahale, O A., Mugellini. *Owner Manuals Review and Taxonomy of ADAS Limitations in Partially Automated Vehicles*. in *AutomotiveUI '19: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2019. Netherlands.
9. Beggiato, M. and J.F. Krems, *The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information.* Transportation Research Part F: Traffic Psychology and Behaviour, 2013. **18**: p. 47-57.
10. Mueller, A.S., I.J. Reagan, and J.B. Cicchino, *Addressing Driver Disengagement and Proper System Use: Human Factors Recommendations for Level 2 Driving Automation Design.* Journal of Cognitive Engineering and Decision Making, 2021. **15**(1): p. 3-27.
11. Caber, N., P. Langdon, and P.J. Clarkson. *Designing Adaptation in Cars: An Exploratory Survey on Drivers' Usage of ADAS and Car Adaptations*. in *Advances in Human Factors of Transportation*. 2020. Cham: Springer International Publishing.
12. McDonald, A., C. Carney, and D.V. McGehee, *Vehicle Owners' Experiences with and Reactions to Advanced Driver Assistance Systems (technical report)*, in *AAA Foundation for Traffic Safety*. 2018: Washington, D.C.
13. Lijarcio, I., et al., *Availability, Demand, Perceived Constraints and Disuse of ADAS Technologies in Spain: Findings From a National Study.* IEEE Access, 2019. **7**: p. 129862-129873.
14. Boelhouwer, A., et al., *How are car buyers and car sellers currently informed about ADAS? An investigation among drivers and car sellers in the Netherlands.* Transportation Research Interdisciplinary Perspectives, 2020. **4**: p. 100103.
15. Reagan, I.J., et al., *Disengagement from driving when using automation during a 4-week field trial.* Transportation Research Part F: Traffic Psychology and Behaviour, 2021. **82**: p. 400-411.
16. Harms, I.M., L. Bingen, and J. Steffens, *Addressing the awareness gap: A combined survey and vehicle registration analysis to assess car owners' usage of ADAS in fleets.* Transportation Research Part A: Policy and Practice, 2020. **134**: p. 65-77.
17. Wozniak, D., et al., *Analysis of advanced driver assistance systems in police vehicles: A survey study.* Transportation Research Part F: Traffic Psychology and Behaviour, 2021. **83**: p. 1-11.
18. DeGuzman, C.A. and B. Donmez, *Drivers Still Have Limited Knowledge About Adaptive Cruise Control Even When They Own the System.* Transportation Research Record, 2021: p. 03611981211011482.
19. Hungund, A.P. and A.K. Pradhan, *A survey on knowledge and perceptions of advanced driver assistance systems in Massachusetts drivers.* Traffic Injury Prevention, 2022: p. 1-3.
20. Larsson, A.F., *Driver usage and understanding of adaptive cruise control.* Appl Ergon, 2012. **43**(3): p. 501-6.
21. Larsson, A. and K. Emanuelsson, *Examining Factors for Low Use of Driver Assistance Systems*, in *In FAST-zero*. 2021.
22. DeGuzman, C.A. and B. Donmez, *Drivers don't need to learn all ADAS limitations: A comparison of limitation-focused and responsibility-focused training approaches.* Accident Analysis & Prevention, 2022. **178**: p. 106871.
23. Biassoni, F., D. Ruscio, and R. Ciceri, *Limitations and automation. The role of information about device-specific features in ADAS acceptability.* Safety Science, 2016. **85**: p. 179-186.

24. NHTSA, *Vehicle Shopper's Guide: Driver Assistance Technologies*. 2018.
25. AAA, *Advanced driver assistance technology names: AAA's recommendation for common naming of advaced safety systems*. 2019.
26. Reagan, I.J., J.B. Cicchino, and D.G. Kidd, *Driver acceptance of partial automation after a brief exposure.* Transportation Research Part F: Traffic Psychology and Behaviour, 2020. **68**: p. 1-14.
27. Sternlund, S., *The safety potential of lane departure warning systems—A descriptive real-world study of fatal lane departure passenger car crashes in Sweden.* Traffic Injury Prevention, 2017. **18**(sup1): p. S18-S23.
28. Cicchino, J., *Effects of automatic emerency braking systems on pedestrian crash risk*. 2022.
29. Teoh, E.R., *What's in a name? Drivers' perceptions of the use of five SAE Level 2 driving automation systems.* Journal of Safety Research, 2020. **72**: p. 145-151.
30. Orlovska, J., et al., *Effects of the driving context on the usage of Automated Driver Assistance Systems (ADAS) -Naturalistic Driving Study for ADAS evaluation.* Transportation Research Interdisciplinary Perspectives, 2020. **4**: p. 100093.
31. NCAP, E., *Euro NCAP vision for 2030 A safer future for mobility* 2022.
32. Reagan, I.J., J.B. Cicchino, and C.J. Montalbano, *Exploring relationships between observed activation rates and functional attributes of lane departure prevention.* Traffic Inj Prev, 2019. **20**(4): p. 424-430.

APPENDIX

Table A 1. Overview of respondents' demographics.

Demographics	**ACC (n=1002)**	**LKA (n=461)**
Year		
18-29	73 (7%)	23 (5%)
30-39	151 (15%)	59 (13%)
40-49	243(24%)	128 (28%)
50-59	243 (24%)	127 (28%)
>60	292 (29%)	124 (27%)
Gender		
Female	143 (14%)	59 (13%)
male	853 (85%)	399 (87%)
Other /do not want to specify	6(0.6%)	3 (0.6%)
Living environment		
City (Stockholm, Göteborg, Malmö)	243(24%)	119(26%)
Urban	555(55%)	253(55%)
rural area	204(20%)	89(19%)
Risk taking in traffic (1-10)	3,7 Mean	3,5 Mean
Technology readiness		
Innovator	65 (6%)	40 (8%)
Early adopter	559 (56%)	284(62%)
Early majority	332 (33%)	126 (27%)
Late majority	30 (3%)	7(2%)
Laggards	9(1%)	2 (1%)
Top 5 represented vehicle brands	Volvo (343/34%), Volkswagen (183/18%), Kia (69/7%), Skoda (54/5%), Toyota (46/5%)	Volvo (150/32%), Kia (57/12%), Volkswagen (48/11%), Tesla (40/8%), Hyundai (20/4%)
Attitude towards ADAS systems		
Very negative	6 (1%)	1 (0.2%)
Fairly negative	15 (2%)	3 (1%)
Neutral	40 (4%)	13 (3%)
Little positive	246 (25%)	96 (21%)
Very positive	695 (70%)	348 (75%)
Frequency of driving		
Daily	592 (59%)	270 (59%)
Minimum 3 days a week	263 (26%)	32 (29%)
At least once a week	119 (12%)	48 (10%)
At least once a month	20 (2%)	6 (1%)
At least once a year	8 (1%)	5 (1%)
Experience of vehicle		
month	(262/26%)	28 (6%)
1-6 months	(185/18%)	113 (25%)
7-12 months	(165/16%)	83 (18%)

13-24 months	(262/26%)	126 (27%)
>24months	(350/35%)	110 (24%)
Number of ADAS systems		
1	37 (4%)	0 (0%)
2	101 (10%)	6 (1%)
3	145 (15%)	19 (4%)
4	226 (23%)	48 (10%)
5	238 (24%)	133 (28%)
6	255 (25%)	255 (55%)

Table 17 ACC contributes to positive driving experience versus driving frequencies

Type of positive experience	*Some occasions a year*	*Some occasions a month*	*An occasion a week*	*At least three days a week*	*Every day*	*Total*
Increased security	0 (0%)	3 (0%)	27 (3%)	63 (6%)	159 (16%)	252 (25%)
Increased fysical comfort	3 (0%)	5 (1%)	35 (3%)	79 (8%)	193 (19%)	315 (31%)
Increased mental relaxation	0 (0%)	1 (0%)	10 (1%)	20 (2%)	37 (4%)	68 (7%)
Opportunities to do other things	0 (0%)	0 (0%)	9 (1%)	18 (2%)	35 (3%)	62 (6%)
Becoming a better driver	5 (1%)	10 (1%)	63 (6%)	105 (10%)	249 (25%)	432 (43%)
Decreased fuelconsumption	6 (1%)	9 (1%)	44 (4%)	99 (10%)	293 (29%)	451 (45%)
More enjoyable cardriving	0 (0%)	7 (1%)	24 (2%)	52 (5%)	120 (12%)	203 (20%)
Other	0 (0%)	0 (0%)	9 (1%)	30 (3%)	47 (5%)	86 (9%)

Table 18 LKA contributes to positive driving experience versus driving frequencies

Type of positive experience	*Some occasions a year*	*Some occasions a month*	*An occasion a week*	*At least three days a week*	*Every day*	*Total*
Increased security	3 (1%)	4 (1%)	40 (9%)	99 (21%)	187 (41%)	333 (72%)
Increased fysical comfort	1 (0%)	4 (1%)	25 (5%)	56 (12%)	129 (28%)	215 (47%)
Increased mental relaxation	1 (0%)	5 (1%)	28 (6%)	60 (13%)	127 (28%)	221 (48%)
Opportunities to do other things	0 (0%)	0 (0%)	5 (1%)	13 (3%)	26 (6%)	44 (10%)
Becoming a better driver	0 (0%)	2 (0%)	22 (5%)	29 (6%)	72 (16%)	125 (27%)
Decreased fuelconsumption	1 (0%)	0 (0%)	5 (1%)	8 (2%)	31 (7%)	45 (10%)
More enjoyable cardriving	1 (0%)	0 (0%)	9 (2%)	11 (2%)	21 (5%)	42 (9%)
Other	1 (0%)	1 (0%)	3 (1%)	10 (2%)	21 (5%)	36 (8%)

Table 19 Overall experiences with ACC versus driving frequencies

Type of experience	*Some occasions a*	*Some*	*An*	*At least three days*	*Every day*	*Total*

	year	occasions a month	occasion a week	a week		
Contributes to a positive driving experience	5 (1%)	17 (2%)	103 (10%)	231 (23%)	497 (50%)	853 (85%)
Keeps speed and distances to vehicle in front	7 (1%)	18 (2%)	105 (10%)	238 (24%)	516 (52%)	884 (88%)
Accelerates and brakes smoothly	7 (1%)	14 (1%)	102 (10%)	211 (21%)	464 (64%)	798 (80%)
Been helpful in dangerous situations	2 (0%)	8 (1%)	56 (6%)	117 (12%)	265 (26%)	448 (45%)
Increases risk to be in dangerous situations	0 (0%)	0 (0%)	10 (1%)	15 (2%)	28 (3%)	53 (5%)
Have good understanding of function of system	7 (1%)	19 (2%)	112 (11%)	246 (25%)	572 (57%)	956 (95%)
Is easy to use	7 (1%)	18 (2%)	108 (11%)	253 (25%)	566 (56%)	952 (95%)
Good coworking in driving the car	5 (1%)	13 (1%)	92 (9%)	216 (22%)	454 (45%)	780 (78%)
Feels to disconnected from driving the car	1 (0%)	0 (0%)	2 (0%)	4 (0%)	22 (2%)	29 (3%)

Table 20 Overall experiences with LKA versus driving frequencies

Type of experience	*Some occasions a year*	*Some occasions a month*	*An occasion a week*	*At least three days a week*	*Every day*	*Total*
Contributes to a positive driving experience	2 (0%)	5 (1%)	35 (8%)	83 (18%)	169 (37 %)	294 (64%)
Keeps the car in middle of lane	3 (1%)	4 (1%)	31 (7%)	69 (15%)	149 (32%)	256 (56%)
Been helpful in dangerous situations	0 (0%)	0 (0%)	12 (3%)	28 (6%)	66 (14%)	106 (23%)
Increases risk to be in dangerous situations	1 (0%)	0 (0%)	1 (0%)	6 (1%)	15 (3%)	23 (5%)
Have good understanding of function of system	3 (1%)	4 (1%)	36 (8%)	107 (23%)	228 (49%)	378 (82%)
Is easy to use	3 (1%)	5 (1%)	38 (8%)	106 (23%)	223 (48%)	375 (81%)
Fights against the system	2 (0%)	0 (0%)	2 (0%)	9 (2%)	19 (4%)	32 (7%)
Feels to disconnected from driving the car	0 (0%)	0 (0%)	3 (1%)	6 (1%)	15 (3%)	24 (5%)

Table 21 Ability to use ACC on different types of roads and traffic enviroments versus driving frequencies

Type of road	*Some occasions a year*	*Some occasions a month*	*An occasion a week*	*At least three days a week*	*Every day*	*Total*
Country roads	7 (1%)	19 (2%)	110 (11%)	241 (24%)	536 (53%)	913 (91%)
Citystreets	2 (0%)	7 (1%)	47 (5%)	100 (10%)	212 (21%)	368 (37%)
Highway roads	8 (1%)	20 (2%)	116 (12%)	251 (25%)	573 (57%)	968 (97%)
Roads with separated lanes (2+1 lanes)	7 (1%)	17 (2%)	99 (10%)	219 (22%)	491 (49%)	833 (83%)
Winding roads	3 (0%)	2 (0%)	34 (3%)	78 (8%)	199 (20%)	316 (32%)
Roads with traffic lights and/or traffic circle	1 (0%)	4 (0%)	24 (2%)	56 (6%)	144 (14%)	229 (23%)

Table 22 Ability to use LKA on different types of roads and traffic environments versus driving frequencies

Type of Road	*Some occasions a year*	*Some occasions a month*	*An occasion a week*	*At least three days a week*	*Every day*	*Total*
Country roads	1 (0%)	4 (1%)	44 (10%)	108 (23%)	218 (47%)	375 (81%)
Citystreets	1 (0%)	1 (0%)	14 (3%)	29 (6%)	68 (15%)	113 (25%)

Highway roads	3 (1%)	4 (1%)	41 (9%)	117 (25%)	240 (52%)	405 (88%)
Roads with separated lanes (2+1 lanes)	2 (0%)	3 (1%)	33 (7%)	98 (21%)	185 (40%)	321 (70%)
Winding roads	1 (0%)	1 (0%)	14 (3%)	33 (7%)	69 (15%)	118 (26%)
Roads with traffic lights and/or traffic circle	0 (0%)	0 (0%)	6 (1%)	18 (4%)	41 (9%)	65 (14%)

PEER REVIEW PAPER

This paper has been peer-reviewed and published in a special edition of Traffic Injury Prevention 24(S1), by Taylor & Francis Group. The complete paper will be available on the Traffic Injury Prevention website soon. To access ESV Peer-reviewed papers click the link below
https://www.tandfonline.com/toc/gcpi20/24/sup1?nav=tocList

PEER REVIEW PAPER

This paper has been peer-reviewed and published in a special edition of Traffic Injury Prevention 24(S1), by Taylor & Francis Group. The complete paper will be available on the Traffic Injury Prevention website soon. To access ESV Peer-reviewed papers click the link below
https://www.tandfonline.com/toc/gcpi20/24/sup1?nav=tocList

Measuring the Transition Window in Conditional Automation

John G. Gaspar, Chris W. Schwarz, Omar Ahmad,

National Advanced Driving Simulator, University of Iowa

United States

Thomas Fincannon

National Highway Traffic Safety Administration

United States

Paper Number: 23-0135

ABSTRACT

Conditional (SAE driving automation level 3) automation is projected to see broader deployment by several manufacturers in multiple countries. Level 3 Automated Driving Systems (ADS) target performing the entire dynamic driving task (DDT) within an operational design domain (ODD), allowing the user to disengage while the system is active, but expect that the driver remains “takeover ready.” In practice, one expectation of conditional automation is that the system will issue a request to intervene (RTI) prior to exiting its ODD, with a “sufficient” time window for the user to reengage in manual driving. To date, research has yet to provide an understanding of what a sufficient transition window will be in the context of near-production systems, specifically in those designed to operate in low-speed traffic jams. This study used an adaptive approach to adjust the timing of the request to intervene during transitions of control in a high-fidelity driving simulator. The transition adjusted from transition event to event based on whether participants were able to successfully take control in the previous event. Success was defined in reference to a baseline group who drove with expectations of being fully engaged in driving (i.e., SAE driving automation level 2). The results show that most participants were able to successfully make transitions with transition windows in the range of 4.5 to 5 seconds. However, some participants took several seconds longer to make successful transitions and for the subset of participants looking away from the forward road at the onset of the RTI, transition windows in the range of 7-7.5 seconds led to more successful transitions. This study provides a starting point for developing an objective definition of a “sufficient” transition window in the context of low-speed conditional Automated Driving Systems.

INTRODUCTION

Several Automated Driving Systems (ADS) with conditional automation are projected to be more widely deployed around the world in the next five years. Conditional automation (also known as SAE International (SAE) driving automation level 3) refers to an ADS capable of performing simultaneous lateral and longitudinal vehicle control along with object and event detection and response (OEDR; [1]).

SAE defines expectations for the ADS and the human operator in conditional automation. The ADS is expected to be capable of performing the entire DDT within its ODD, with the exception of fallback to a takeover-ready driver upon system's request. The ADS is also expected to be able to preview an exit from its ODD, at which point it issues a RTI to the human. The ADS would need to provide a "sufficient" time window for the human to transition back into manual control. The expectation is that the human operator will not be forced into a time-critical response situation. There are corresponding expectations for the human operator. With the ADS active within its ODD, the operator can fully disengage from the DDT and perform non-driving tasks (e.g., emailing on their phone). The operator must remain "receptive" to RTIs. The operator also functions as the "takeover-ready user." This means that the human operator is responsible for bringing the vehicle to a minimal risk state. For example, if a vehicle component malfunctions and impacts ADS operation, the system would issue an RTI and the takeover-ready human operator is responsible for resuming control to take necessary manual actions to assure safety of the vehicle.

A large body of research exists on Transition of Control (TOC) from SAE levels 2 or 3 to level 0. McDonald and colleagues [2] and Zhang and colleagues [3] wrote useful reviews of this literature. These reviews focus on the impact of different parameters (RTI timing and modality, secondary task engagement, etc.) on TOC performance. Importantly, this research has largely focused the crash-imminent transitions of control in unrestricted ODDs.

However, this situation does not meet the expectation laid out for level 3 ADS. By categorizing their ADS as level 3 according to the SAE J3016 definitions [1], the expectation is that the ADS will provide a sufficient transition window that is compatible with human capabilities for the driver to resume manual control. A recent report from the German Ethics Commission on Automated and Connected Driving states, "The software and technology in highly automated vehicles must be designed such that the need for an abrupt handover of control to the driver ('emergency') is virtually obviated" [4]. This includes so-called "silent" automation failures, where the ADS fails without notifying the user. This implies that the most common control transitions initiated by level 3 features will involve the ADS notifying the user of an upcoming exit of the ODD. The ADS is expected to provide a "sufficient" transition window for the user to reengage in driving before reaching the ODD limits [1].

To date, there is no consensus definition of what constitutes a sufficient transition window. Furthermore, most research on TOC in conditional automation has focused on relatively broad ODDs, despite the fact that the first deployments of these ADSs will be in well-defined ODDs, such as low-speed traffic jams.

The objective of this study was to examine what may constitute a minimum sufficient transition window (MSTW) in the context of low-speed traffic jam conditional automation. To achieve this objective, the study used an adaptive methodology, where the duration of the transition window changed from event-to-event based on success or failure in the TOC.

METHOD

Participants

The sample consisted of 50 licensed US adult drivers. The sample consisted of drivers between ages 21 and 65, with each group being comprised of 50% male and 50% female drivers. Participants were recruited from the National Advanced Driving Simulator's subject registry and through campus-wide emails. All study procedures were approved by the University of Iowa Institutional Review Board and all study participants provided written informed consent.

Simulator and Virtual Database

The NADS-1 simulator consisted of a 24-foot diameter dome enclosing a full-size 2015 Toyota Camry sedan with active steering and pedal feedback (Figure 1). A 13-degree of freedom motion system provided participants realistic acceleration, braking, and steering cues. Sixteen high-definition (1920x1200) LED (light emitting diode) projectors display seamless imagery on the interior walls of the dome with a 360-degree horizontal field of view. The data sampling rate was 240 Hz.

Figure 1. Interior Views of the NADS-1 Simulator

The study drive utilized a virtual database consisting of a divided interstate highway with three lanes of traffic in each direction. Within this 40-minute route, the drive alternated between periods of higher-speed (55mph) dense traffic and low-speed (22mph) traffic jams. Transition points, where Traffic Jam Auto Drive (TJAD) requested a manual takeover, occurred when TJAD identified an approaching ODD exit.

Traffic Jam Auto Drive (TJAD)

A level 3 ADS designed specifically for operation in congested traffic was to be used in this project. This system was called TJAD and mimicked the functionality and interfaces of near-production level 3 conditional automation (based on an earlier informal technology review). TJAD had the following ODD and functionality:

- Available only on divided multi-lane highways.
- Available only in dense traffic, with vehicles detected ahead and in the adjacent lane(s), traveling 35 mph or less.
- When ODD criteria are met, the user received notification that TJAD was available. It could be activated via a steering wheel button.
- When active, TJAD controlled the longitudinal and lateral control position of the vehicle. The user was told they could disengage from driving.
- When TJAD sensed an upcoming ODD exit, it issued an RTI to the user.

Non-Driving Task

To encourage disengagement from driving when TJAD was active, participants engaged in a demanding non-driving task during periods where TJAD was engaged. Participants were instructed to perform an email task using a provided smart phone. The inbox in the email application contained fictional messages that were either work-related, personal, or promotional. The participant was given rules for which actions to take with each type of message and told that their final compensation for their participation depended on how well they performed this task and how many emails they processed during the study drive.

Transition of Control and RTI

At the end of each traffic jam window, TJAD encountered one of two situations that fell outside its ODD, prompting the system to issue an RTI to the participant. There were 12 total TOC events. These events were a revealed stopped vehicle (6) and a revealed work zone (6). In both cases, the system identified a stationary object (either a vehicle or work zone) in the travel lane and issued an RTI. The RTI process for the two types of events was identical. The participant needed to take control and execute an avoidance maneuver around the object. Figure 2 shows timeseries diagrams of the two types of events. For all events, the object (i.e., stopped vehicle or work zone) was revealed at 15s time-to-collision based on the participant's speed. Because speed was controlled by TJAD, the reveal distance was nearly identical across all events. The direction of the reveal (left vs. right) was balanced across events. The order of events was fixed across participants.

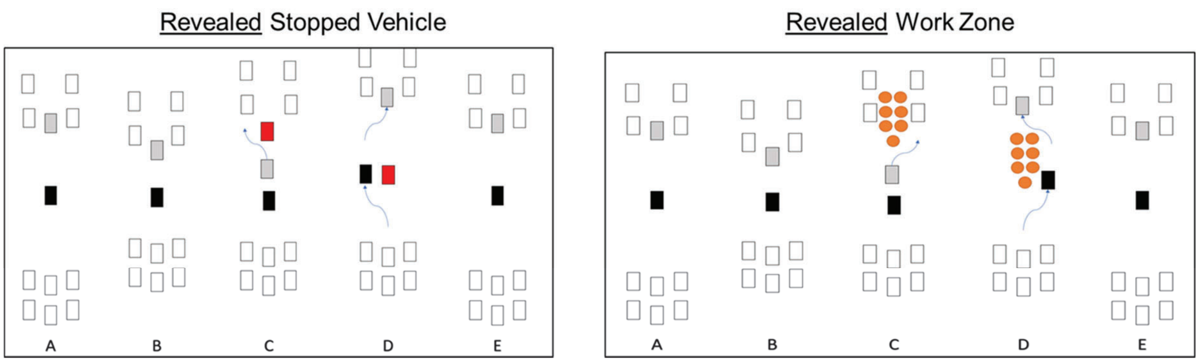


Figure 2. Study Drive Events

The RTI was designed to be representative of near-production systems. The RTI consisted of two auditory-visual stages, as shown in Figure 3. Stage 1 was intended to be a notification to the participant of an upcoming ODD exit. Stage 2 was intended to be a more severe alert prompting the participant to take control immediately. If participants did not take over by the end of the transition window, the system entered a failure mitigation strategy (FMS) where the vehicle began braking automatically. The transition window constituted the time from the start of RTI stage 1 until the start of the FMS.

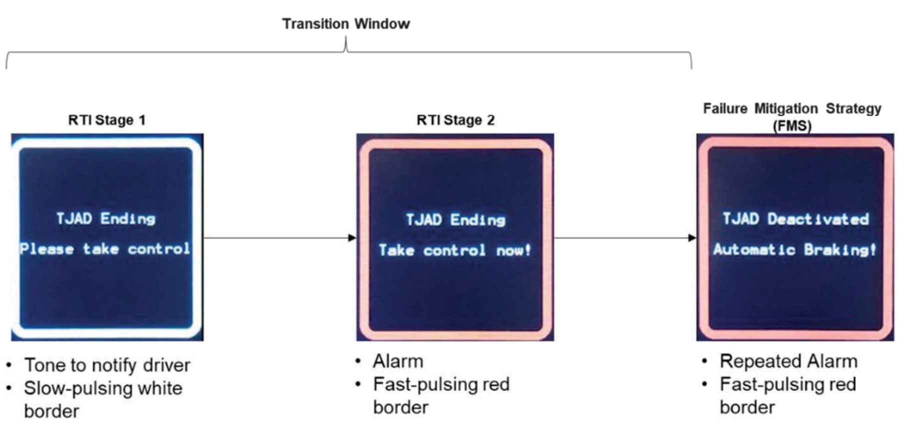


Figure 3. Request to Intervene

Adaptive Transition Window

The study was designed to estimate the MSTW using an adaptive approach, whereby the transition window changed based on whether the previous transition was successful or not. The duration of the transition window was adjusted from event to event by changing the timing of the RTI Stage 1 in relation to the object reveal. Figure 4 shows the sequence of the transition of control events. Note that RTI timing was adjusted based on time to collision (TTC).

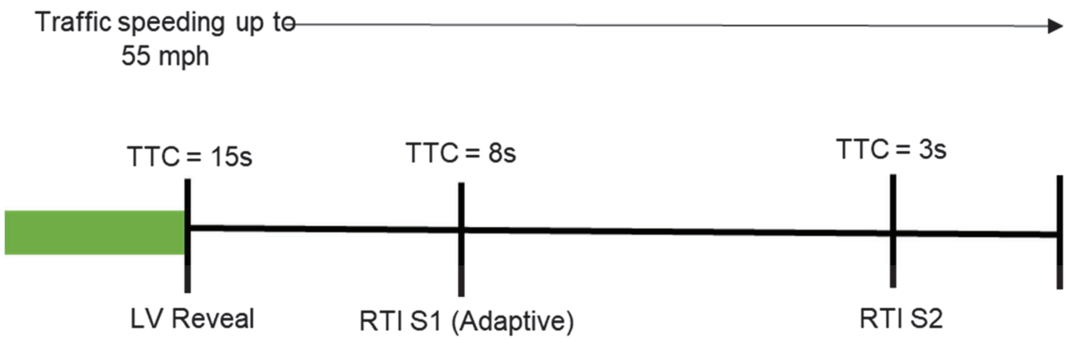


Figure 4. Transition of Control Event Sequence

To examine participant's MSTW, the step size (i.e., change to transition window) decreased across the events, as shown in Table 1. Note that this adaptive sequence was independent for the two types of events. For each event type, the initial transition window was 7.5s, which was also considered the maximum transition window (i.e., transition windows could not be longer than 7.5s). If participants responded successfully in all 6 events, they could reach a final transition window near 4s (for RTI stage 1).

Table 1. Adaptive Step Size

Trial	Step Size (+/- s)
1	-
2	1.6
3	0.8
4	0.4
5	0.2
6	0.1

Baseline Group

In order to determine success or failure during the transitions, data from a baseline group who remained engaged in the driving task (i.e., did not perform the email task) were collected before the main study. This group did not receive an RTI but were told they needed to take back control if they saw any stationary objects in the roadway, as these fell outside the system's ODD. For this baseline group, we calculated the TTC with the stationary object at the point where the participant began responding. Using the 85th percentile of this distribution of TTC values from the baseline group, we computed a success/failure threshold that was then used for the adaptive approach described above.

Procedure

Potential participants completed an online screening to verify they met the inclusion criteria for the study. Participants provided written informed consent and completed a driving questionnaire. During the study visit, participants received training on the simulator, TJAD, and the email task (other than the baseline group). Participants were instructed about TJAD's capabilities and its ODD, as well as the RTI sequence. Training provided information and practice for the email task and informed participants about the incentive for engaging in the non-driving task.

Participants then entered the simulator and completed a practice drive, where they practiced engaging and disengaging TJAD during traffic jams. They also practiced the non-driving email task while TJAD was active. Participants then completed a wellness questionnaire to screen for simulator sickness. Eligible participants then completed the study drive, which lasted approximately 35 minutes. This was followed by a second wellness questionnaire, a post-drive questionnaire, and debriefing.

RESULTS AND DISCUSSION

Data were reduced using custom MATLAB scripts. R statistical software was used to analyze and visualize the reduced data. Video data were coded using Boris open-source video coding software and synchronized with simulator data based on frame numbers.

Baseline Transition of Control

TTC was calculated at the point at which participants began their first response following the lead vehicle reveal. Responses could constitute either the first braking or steering response following the reveal. Braking was defined as a brake pressure of 5 pounds or greater. Steering response was defined as a change in steering wheel angle of 10 degrees or more. TTC represents the time before the participant would collide with the stationary object if no response was executed and speed remained constant. These thresholds were selected based on previous studies using the NADS-1 simulator platform. Larger TTC indicate that participants began responding with longer time gaps, indicating they would have more time to execute a safe response.

Figure 5 shows the distribution of TTC values for the baseline group. The vertical bar represents the 85th percentile of the distribution, which was 4.6s. This threshold was used as the cutoff to delineate success vs. failure for the main study. If responses occurred at a TTC of 4.6s or more, they were classified as successful. If responses occurred at less than 4.6s TTC, they were classified as unsuccessful.

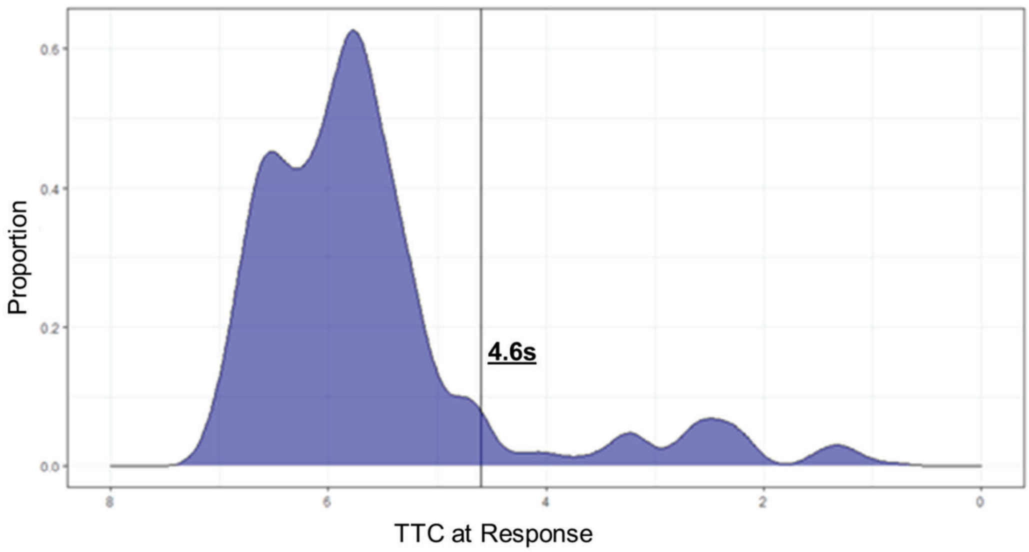


Figure 5. Time to Collision Distribution of Baseline Group

Adaptive Paths

Figure 6 shows the adaptive sequence for all participants (N = 40) in the main study. This figure traces the different paths through the adaptive sequence for each of the two types of events. The figure is color-coded for successful/unsuccessful responses.

Most participants showed a decreasing transition window for both event types over the course of the study. Each event type began with an initial transition window of 8s. By the end of the sequence for each event type, participants were able to successfully transition into control in a little more than half that initial time.

Minimum Window by First Failure

One way to think about the minimum sufficient transition window is the window where participants made their first unsuccessful transition (i.e., failure). Figure 6 shows that several participants (approximately 15% across the aggregated event types) failed with ~6s transition windows, resulting in the adaptive logic increasing the transition window. Figure 6 shows there was a subset of participants who struggled to transition into control successfully (compared to baseline) with transition windows ranging from 6-8 seconds.

One conclusion from these adaptive sequences is the variability between individuals with respect to transition of control success as a function of the transition window. Most participants were successful across the majority of events for each event type, consistently reducing the duration of the transition window. Importantly, however, the distribution of transition windows across each event (for each event type) shows long tails of the distribution, indicating the likelihood of outlier participants (or individual events).

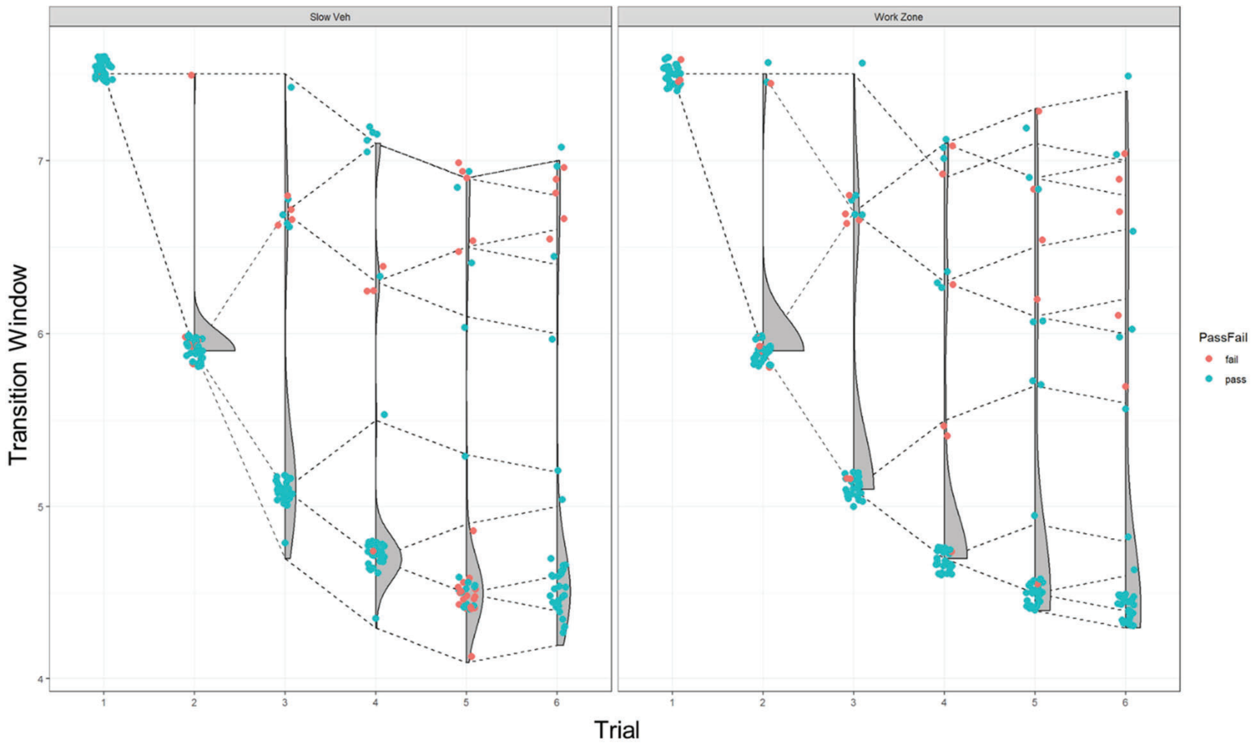


Figure 6. Adaptive Paths

Minimum Window by Final Event

Another way to estimate the minimum sufficient window is to examine the distribution of transition windows for the final event, which represent the conclusion of the adaptive sequence for each event type. These distributions are shown in Figure 7. Again, most participants achieved a final transition window between 4 and 5 seconds. However, it is also important to consider the tails of the distributions, as they may be the most important contributors to crashes and therefore key considerations in selection of system thresholds. The 85th percentile of the distributions was at approximately 6.5s.

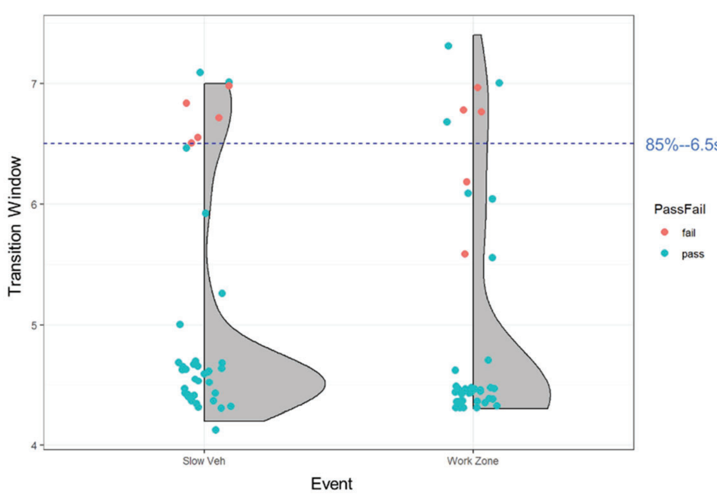


Figure 7. Distribution of Responses in Final Events. Blue line represents 85th percentile

Minimum Window by Glance Location

One consideration in the identification of thresholds for transition of control is the impact of the behaviors that precede the transition on the likelihood of transition success. Design needs to consider a range of potential user states, particularly in the context of conditional) automation where it is permissible for the user to fully disengage from the DDT while the system is active.

Review of driver-facing video suggested two distinct patterns of behavior, specifically visual attention in the form of glances to the forward road, preceding the transitions. One behavior pattern involved occasional or frequent glances forward, presumably to check on the driving task and potentially to anticipate upcoming transitions of control. The other pattern of behavior involved few if any checks forward.

To determine whether these different glance patterns contributed to transition of control success in the adaptive experiment, we coded glance location at the time of the reveal event. Using the frame where the reveal occurred, an independent rater coded glance location as either forward (i.e., at the road) or down (i.e., at the email task).

To examine the potential impact on transition success across the different transition windows captured in this study, we binned data based on the available preview time, which reflects the duration of the transition window. For example, a preview time of 7 seconds corresponds to a 7 second transition window where the RTI gives 7 seconds for the participant to transition into control (assuming the participant looks up at the RTI).

Figure 8 shows the success rates across different preview time bins. These success rates are divided based on the glance location at the RTI, looking up or looking down. For events classified as looking up, participants had high success rates across the range of transition window bins. In three cases, success was 100% when participants were looking up at the RTI. Importantly, even at the short transition window bins, success rates were high when participants were looking up.

Conversely, success rates were lower when participants were looking down at the RTI. Success rates were approximately 50% across all time bins for the subset of events where participants were looking down. Figure 8 also shows that, within the subset of looking down events, success rates were higher for transition windows in the 7-7.5s range (82%) compared to success rates in all the other shorter transition window bins. No other transition window range had success rates greater than 65%.

These results provide two important conclusions. First, glance location immediately preceding a transition of control has a key impact on the likelihood of transition success. When participants were looking forward at the time of the RTI, the duration of the transition window did not matter, and participants were able to successfully respond within even the shortest transition windows. Conversely, participants who were looking down were much less likely to show successful responses across all transition windows. This reiterates the importance of considering driver state in evaluation of transition of control and RTI design.

Second, for those participants who were looking down at the RTI, the longest transition window duration resulted in higher success rates than all shorter transition windows. This suggests that a transition window in the range of 7-7.5s may be necessary to enable successful transition of control in situations where drivers are disengaged from the DDT at the time the RTI is issued.

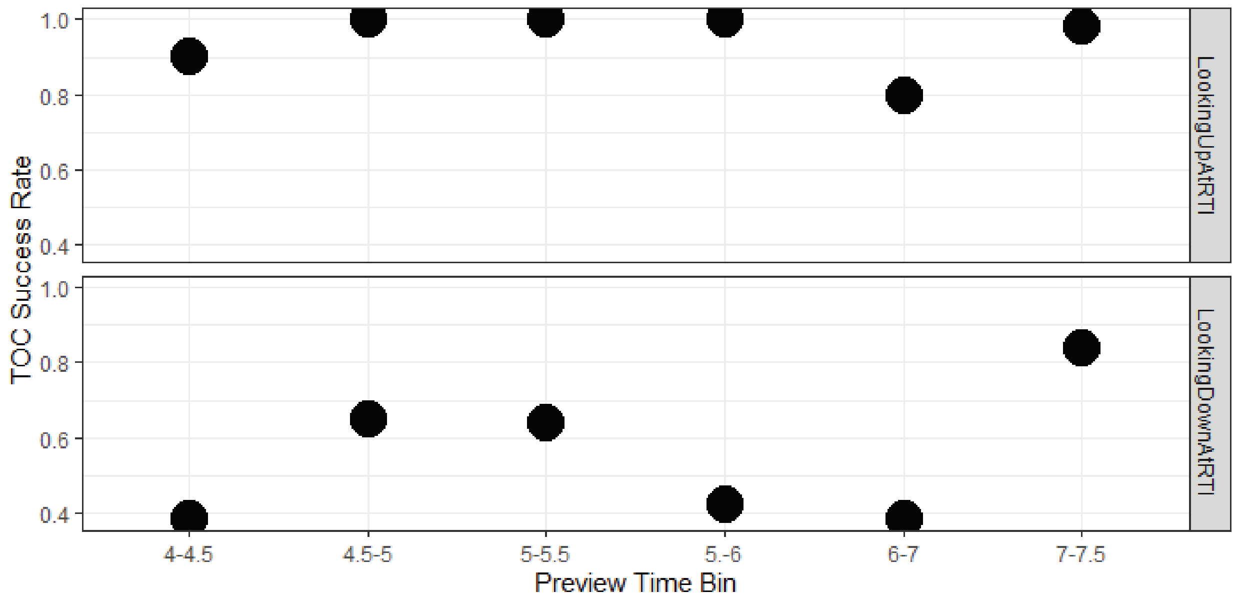


Figure 8. Transition of Control Success by Binned Transition Window (Preview Time) Duration

Limitations and Future Directions

Several limitations from the present study are worth noting. One potential limitation to consider in this study is the repeated nature of the transition of control events. It is possible that over time participants altered their visual attention strategy to better detect the onset of the lead vehicle reveal events. In real-world interactions with Automated Driving Systems, time-critical transitions of control will be rare. The majority of transitions are likely to be uneventful, allowing drivers larger time cushions to transition back into manual control. This study focused on quantifying the minimum sufficient window that may be necessary in rarer transition of control situations. While rare, such "edge case" events are likely to contribute disproportionately to crash risk during human-automation interactions.

This study used a high-fidelity driving simulator to assess transition of control performance. While simulation may lack some of the complexity of on-road driving, the motion cues provided by the simulator in this study mimicked those experienced by drivers in real vehicles and were likely important cues, particularly for disengaged drivers, in determining when TJAD might issue an RTI. There is a need for future research, particularly naturalistic driving data, to understand the myriad of ways in which drivers will behave when conditional automation is available. This will help inform the design of future experimental work, which can help inform the creation of objective best practices for some of the automated vehicle terminology.

The results of this study are best considered in the context of low-speed traffic jam conditional automation. The study focused on traffic jam systems because they were identified as one of the first likely deployment situations. Future research should focus on expanded and more diverse ODDs to understand whether the minimum sufficient transition windows identified in this study generalize to a wider range of systems and situations. Along these same lines, it will be important to further understand the impact of different RTI design characteristics (e.g., alert modality) on transition of control in conditional automation.

Finally, the results of this study highlight the importance of understanding individual differences with respect to driver interactions with vehicle automation. It is worth noting that some participants were consistent in their pattern of behavior (e.g., always executing check glances) while other participants changed their behavior from event to event. In addition to changing overall glance patterns, participants may have changed the frequency of their glances, which again may have impacted the likelihood of early detection of the reveal event. While beyond the scope of this study, analysis of glance behavior during periods of automated driving will be important for understanding the impact of individual differences on transition of control. Understanding the impact of individual differences and demographic characteristics on transition of control is another important avenue for future ADS research.

CONCLUSIONS

The objective of this study was to examine the minimum sufficient transition window in the context of traffic jam conditional automation. By using a unique adaptive design, the study procedure used a high-fidelity simulator to systematically tune the transition window necessary for participants to successfully regain manual control following periods of automated driving. The results indicate that many participants were able to make successful transitions of control with relatively short (4.5-5s) transition windows. However, the analysis also suggests that some participants, in particular those drivers who were looking away from the forward roadway at the onset of the RTI, required longer transition windows (i.e., 7-7.5 seconds). It should be noted that these absolute transition window durations may depend on factors such as travel speed, traffic density, and other operational design domain characteristics. Though more research is needed to understand what other factors may impact a human driver to successfully resume control in conditional automation after being disengaged from the driving task, this study provides empirical data of observed human driver needs related to the minimum transition window for the traffic jam conditional automation use-case. Other factors can only increase these observed transition windows.

REFERENCES

[1] S. A. E. International (2021). Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles-. Publication J3016_202104.

[2] McDonald, A. D., Alambeigi, H., Engström, J., Markkula, G., Vogelpohl, T., Dunne, J., & Yuma, N. (2019). Toward computational simulations of behavior during automated driving takeovers: a review of the empirical and modeling literatures. Human factors, 61(4), 642-688.

[3] Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving: A meta-analysis of 129 studies. Transportation research part F: traffic psychology and behaviour, 64, 285-307.

[4] Luetge, C. (2017). The German ethics code for automated and connected driving. Philosophy & Technology, 30(4), 547-558.

DRIVERS' RESPONSE TO AUTOMATION INITIATED DISENGAGEMENT IN REAL-WORLD HANDS-FREE DRIVING

Pnina Gershon
Bruce Mehler
Bryan Reimer
Massachusetts Institute of Technology
USA

Paper number 23-0137

ABSTRACT

Driving automation features in the form of advanced driver assistance systems (ADAS) that can control the longitudinal and lateral vehicle kinematics on sustained bases (SAE Level2) are becoming increasingly available in consumer vehicles, making the study of drivers' behavioral adaptation and the impacts of automation central to driving safety. This study used real-world data to assess drivers' responses to automation-initiated disengagements by quantifying changes in drivers' moment-to-moment visual attention and vehicle control behaviors.

Fourteen drivers (36% female) drove for one-month each a Cadillac CT6 equipped with a data acquisition system that recorded driving kinematics, miles driven, automation use, GPS, and video of the driver and driving environment. Cadillac's Super-Cruise (SC) is one of the most advanced, commercially available partial automation systems that, when engaged, enables hands-free driving while directly monitoring the driver's head orientation.

A total of 265 SC initiated disengagement events were identified (mean=18.9; SD=16.5 per driver) across 5,514 miles driven with SC. In general, SC initiated disengagements were associated with substantial changes in glance distribution. Immediately after disengagement, the proportion of glances to the Road decreased from 83% to 68% and at the same time the proportion of glances to the Instrument Cluster increased substantially, from 8% to 27%. The period following SC initiated disengagement was also characterized by a 44% increase in the overall number of transitions between glance areas (from 845 transitions before to 1218 after the disengagement across events). The most dominant visual attention patterns after SC disengagements were Road to Instrument Cluster (57% increase) and Instrument Cluster to Road (222% increase). Linear quantile mixed-effects models were used to estimate glance duration before and after disengagements. Findings indicate that on-road glance duration following SC disengagement decreased significantly and was 4.86sec shorter in the 85^{th} quantile ($Q15_{Before}$=0.5, CI=[0-2.24], $Q15_{After}$=0.43, CI=[0-2.11], p=.04; $Q50_{Before}$=2.02, CI=[0.8-3.24], $Q50_{After}$=1.45, CI=[1.02-1.88], p=ns; $Q85_{Before}$=6.63, CI=[2.06-11.2] to $Q85_{After}$=1.77, CI=[0-3.67], p<.001). Analysis of driver hands on-wheel behavior indicate that drivers adopted SC's hands-free feature to a substantial degree, taking both hands off the steering wheel more than 75% of the time SC was engaged. Takeover duration when driving hands-free was significantly longer (2.4sec) compared to driving with at least one hand on the steering wheel (1.8sec).

In conclusion, concerns over the phenomenon of driver out-of-the-loop, coupled with known limitations of partial automation systems, have led research to focus on driver response to automation-initiated disengagement and the ability to regain manual control. We find that real-world automation-initiated disengagements trigger substantial changes in driver glance behavior including shorter on-road glances and frequent transitions between the Road and the Instrument Cluster glance areas. This behavior pattern likely represents drivers' searching for information related to the disengagement or the automation state and may be shaped by the automation design. Higher levels of automation may introduce more substantial changes in visual and vehicle control behaviors during automation initiated-disengagements. This data provides useful information to designers charged with developing assistive and automated systems and empowers regulators and safety advocates with insights needed to better guide appropriate utilization of ADAS technologies.

INTRODUCTION

Driving automation features in the form of advanced driver assistance systems (ADAS) are becoming increasingly available in consumer vehicles, making the study of drivers' behavioral adaptation and the impacts of automation central to driving safety. The Society of Automotive Engineers (SAE) defines a taxonomy of six automation levels, ranging from manual driving (i.e., SAE Level 0 - no driving automation) to fully self-driving vehicle under all conditions (i.e., SAE Level 5 - full automation) [1]. Partial automation systems (SAE Level 2) currently available in consumer vehicles can simultaneously control the longitudinal (e.g., adaptive cruise control) and lateral (e.g., lane centering) elements of the dynamic driving task on a sustained basis. When using partial automation, the driver delegates elements of operational control of the dynamic driving task over to the automation, while remaining responsible for monitoring, preforming object and event detection, and responding when needed. Essentially, with partial automation systems, driving has become a collaboration between driver and automation, where drivers are expected to maintain attention to events on the roadway and be ready to regain control of the vehicle at all times [2-4]. One such commercially available partial automation system is Cadillac CT6's Super Cruise (hereafter SC) which is a geofenced system that can be used only on mapped limited-access highways. When engaged, SC enables legs-free and hands-free driving while continuously monitoring the driver's visual attention using a camera-based Driver Monitoring System (DMS). Leveraging driver head-pose data derived from the DMS, SC supervises driver attention to the road through a series of escalating alerts designed to work in cohesion with features of the ADAS. SC and other partial automation systems have been associated with reports of increased driver comfort and reduced driving demand; however, the safety benefits of such systems are still debated as there is a potential for over-reliance [3, 5].

As driving demands are lowered by automation and the driver's role pivots toward monitoring, a task in which humans inherently underperform, drivers are more likely to experience challenges to maintain sufficient attention to the driving task and increased propensity for phenomena like driver out-of-the-loop (OOTL), cognitive underload, mode confusion, and distracted driving [6-8]. These concerns over driver inattention, coupled with known limitations of partial automation have led research to focus on time-critical, system-initiated disengagements where the automation issues an immediate takeover request and drivers are required to takeover either the lateral (steering) or both lateral and longitudinal (steering and speed) control [7-8]. The growing body of research that focuses on driver-automation interactions and system-initiated disengagement indicates that extended use of automation may result in slower hazard detection and longer reaction time to obstacles compared to manual driving [7, 9]. Driver attention monitoring and support systems are one mechanism intended to mitigate lapses in driver engagement by providing feedback to the driver or adapting the automation functionality in real-time [10-13]. Currently available driver monitoring systems use steering wheel torque-based sensors and/or driver facing cameras to track gaze and/or head position to infer driver state and intervene when a threshold for apparent inattention is exceeded. SC, for example, has a camera-based driver monitoring system that employs multimodal cues (visual, auditory, and haptic) to support driver attention on moment-to-moment basis.

While partial automation systems are increasingly available, the literature to date is limited by the lack of objective, real-world data on the extent to which drivers use partial automation, the context and frequencies in which the automation initiates disengagements, and how drivers respond to such events. Furthermore, in situations of automation-initiated disengagement events, it is still largely unknown to what extent the use of systems like SC, that allows hands-free driving, will impact driver ability to regain control in a timely manner. Naturalistic studies that directly and continuously record real-world driver behavior, capturing the use of automation and automation disengagements and driver takeovers, along with comprehensive, moment-to-moment driving data including vehicle speed, g-forces, engagement in non-driving related activities, driver visual attention patterns, traffic density, and other environmental conditions [4, 14-16] allow researchers to systematically address these gaps.

The current study used real-world driving data to characterize the occurrences of SC-initiated disengagement requests and to assess how drivers respond to these events. By quantifying changes in visual attention prior and immediately after SC disengagements, and analyzing the moment-to-moment distributional properties of driver glance and hand-on-wheel behaviors, this study provides insights on how drivers dynamically allocate their visual attention and the time it takes them to regain basic control following automation-initiated disengagement. The uniqueness of this paper comes from the large naturalistic driving dataset of consumers using Cadillac SC, a commercially available partial automation system that, when engaged, allows hands free driving. The findings from this paper can inform improvements in the design of assistive and automated vehicle technology as a whole, by enhancing knowledge on how drivers leverage automation under real-world operating conditions.

METHOD

Participants

A total of fourteen drivers (36% female) with an average age of 42 years old (SD=13.3 years old) participated in the study. Drivers from the greater Boston area of Massachusetts were recruited through flyers and online advertisements. Potential participants were screened according to inclusion criteria that required them to pass background and driving record checks, and to have highway driving as part of their regular commute. Drivers were excluded if they had been involved in a police-reported crash or received two or more traffic violation convictions in the past year, or had other risk markers (e.g., selected criminal records, or previous license suspension). Participants were provided with an MIT owned vehicle for one month along with paid tolls and a monetary incentive of $50 to complete a post drive interview. Participants received training on the available automation features including Adaptive Cruise Control (ACC, SAE Level 1) and the SC (SAE Level 2) systems. The training session started with a 30-minute static in-vehicle instruction period followed by an hour of on-road training. During the training drive, participants were familiarized with and were asked to interact with the different automation systems.

Data Source and Data Reduction

Data were drawn from the ongoing MIT Advanced Vehicle Technology (MIT-AVT) naturalistic data collection effort. As part of the study, participants drove MIT's instrumented 2018 Cadillac CT6 vehicles for a period of one month each (between April 2018 and May 2019). The study vehicles were instrumented with RIDER (Real-time Intelligent Driving Environment Recording) data acquisition system (Fridman et al. 2019) that continuously collected data from: (i) Controller Area Network (CAN) bus to determine vehicle kinematics, driver interaction with the vehicle controllers, and the state of in-vehicle automation systems, (ii) Global Positioning System (GPS) to record location; and (iii) four 720p video cameras that continuously captured (30 fps) the driver's face, vehicle cabin, instrument cluster, and the view of the forward roadway (See Figure 1). Together, these multiple data sources and data types provided rich and comprehensive data related to the vehicle state, driving environment, driver behavior, and the use of automation.

Figure 1. **Example of hands-free, legs-free driving in the Cadillac CT6 with SC engaged as captured by the four RIDER cameras, including: the driver's face, in cabin view, instrument cluster, and the view of the forward roadway.**

Automation initiated disengagement, operationally defined as an event where the automation (SC) issued an immediate takeover request triggered by conditions such as: (i) a failure of the automation system (e.g., sensing, computation, and planning), (ii) changes in the operational design domain (ODD), like entering a construction zone, and (iii) driver behaviors like accelerating beyond the ACC braking authority. To identify these events, a filter was run over the continuous CAN data and each flagged event was evaluated by experienced coders who viewed a video segment of 10sec before and 10sec after SC initiated the disengagement. Coding employed a systematic protocol to validate the occurrence and annotate the context of the disengagement events.

Measures

<u>Glance Location</u>

Each SC disengagement event was manually annotated capturing the driver's glance location at a frame-by-frame level at 30 fps. Glances were classified according to the following categories: (i) Road (any glance directed outside the windscreen); (ii) Instrument Cluster (any glance to the instrument cluster or steering wheel region); (iii) Down & Center Stack (any glance to the center stack, the in-car multimedia touch-screen, or down when looking at a smartphone or other object in the lap region); (iv) Mirrors (any glance to the left, right windows or mirrors, and rearview mirror). When a glance did not fall under the categories listed, it was annotated as Other. The Other category contains non-specific glances away from the road. For example, over the shoulder glances, glances towards objects positioned at head level or higher, but also those rare moments when the eyes were closed for longer than a typical blink duration (0.3sec). If the driver's eyes were temporarily not visible due to lighting conditions, etc., glances were coded as Not available. The Other and Not available glance categories were excluded from the glance analysis. Lastly, following ISO 15007-1:2014, a single glance consisted of the transition time toward an area of interest and the subsequent dwell time on that area.

Steering Wheel Control Level

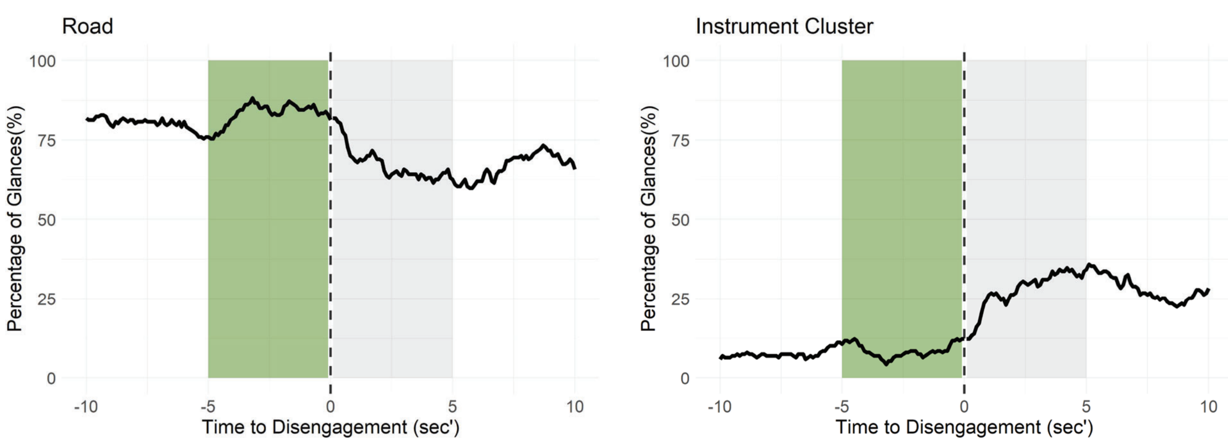
A driver's hand position relative to the steering wheel was manually annotated frame-by-frame at 30 fps and included the following labels: (i) No hands (driving hands free with no hand on the steering wheel); (ii) One hand (driving with one hand on the steering wheel); and (iii) Two hands (driving with two hands on the steering wheel); and (iv) Not visible (used when the driver's hands were momentarily not visible or in poor video quality).

RESULTS

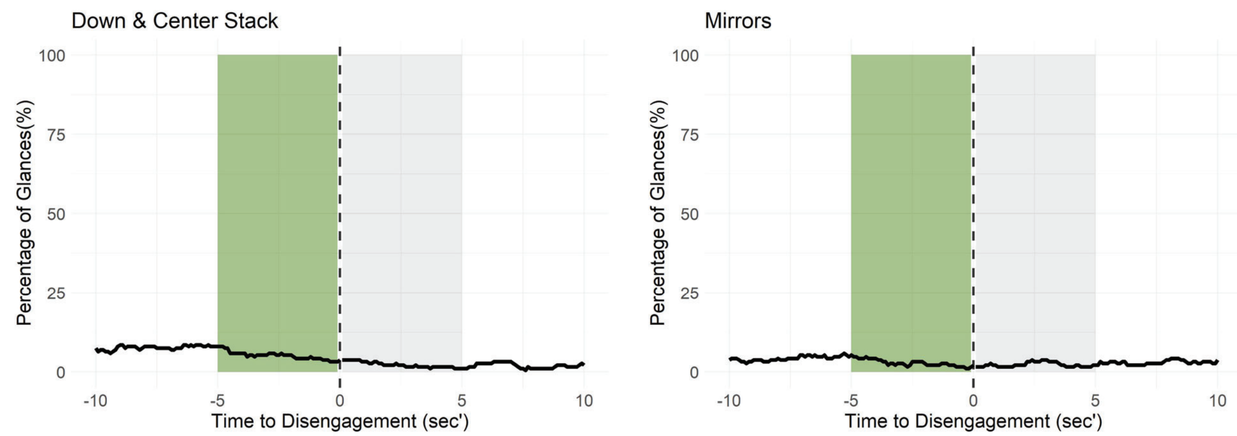
During the data collection period and within the study sample, 265 SC initiated disengagement events were identified (mean=18.9; SD=16.5; range= 0 to 52 events per driver) across 5,514 miles driven with SC engaged. The analysis below included only disengagement events for which drivers' eyes and hands were visible and SC was engaged for at least 10sec before the disengagement (n=187). Following disengagement, the transition in the automation state was primarily to ACC (89%); only in a limited number of the disengagement events was the transition directly to manual driving (11%). During the evaluated 10sec after SC disengagement, drivers may or may not have reengaged SC.

Glance Behavior

Figure 2 shows the time course of glance behavior from 10sec before to 10sec after SC initiated disengagement, across the evaluated glance areas. In general, SC initiated disengagements were associated with substantial changes in glance distribution. During the transition phase (i.e., in the 5sec around the disengagement), the distribution of glances prior to the disengagement (highlighted in green) were 83% (SD=3%) directed to the Road and only 8% (SD=2%) to the Instrument Cluster. Immediately following SC disengagement (highlighted in gray), glance proportions to the Road decreased substantially (68%, SD=6%) and the proportion of glances to the Instrument Cluster increased by 19% (27%, SD=6%) (see Figures 2a and 2b). This shift in glances towards the Instrument Cluster may capture drivers' search for information related to the reason for disengagement and/or to SC state. At the start of the transition phase, the average proportion of glances to the Down & Center Stack area was 5% (SD=1%), and this value decreased to 2% (SD=1%) in the 5sec after the disengagement (Figure 2c). The proportion of glances to the Mirrors showed a slightly decreasing trend as well.



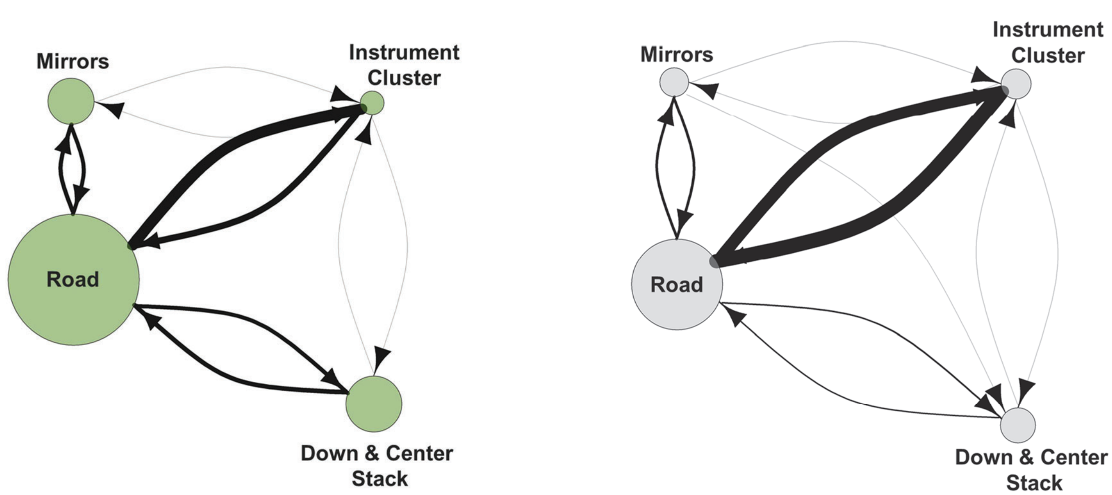
(a) **(b)**



(c) **(d)**

Figure 2. **Proportions of glances to the different areas at any given time during SC initiated disengagement events. SC initiated disengagement is at 0sec, marked with black dashed line. The shaded bands indicate the intervals of the transition phase starting 5sec before (highlighted in green) and 5sec after (highlighted in gray) SC initiated disengagement.**

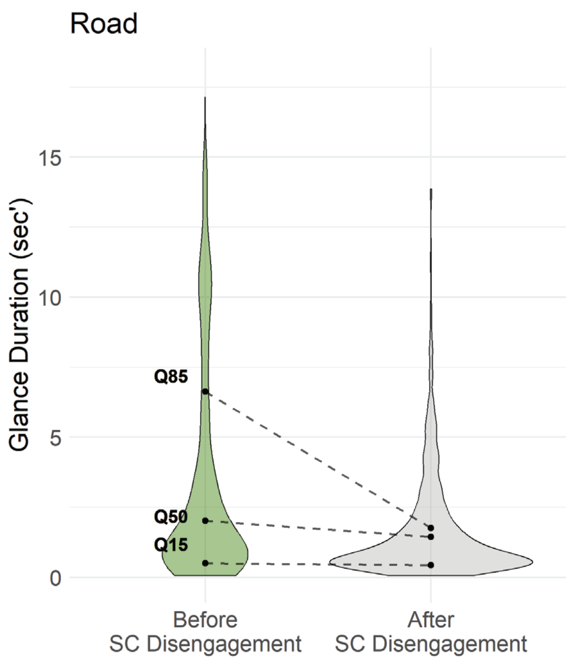
Figures 3a and 3b summarize driver visual attention patterns, capturing both the number of transitions (normalized proportion of transitions indicated by the edge width) and the mean duration spent in each glance area (indicated by the relative size of the vertex) in the 10sec before and 10sec after SC initiated disengagement. Overall, the number of transitions between glance areas following SC disengagement increased by 44%, from 845 transitions before the disengagement to 1218 after the disengagement across all events. The most prevalent visual attention patterns following SC initiated disengagement were Road → Instrument Cluster and Instrument Cluster → Road, which increased by 57% (from 262 to 412 transitions) and by 222% (from 156 to 503 transitions), respectively. In-parallel, the number of transitions from the Road → Down & Center Stack decreased by 45% (from 99 to 54) and the number of transitions from Road → Mirrors decreased by 10% (from 103 to 93) (see Figures 3a and 3b). All glances to all off-road areas, before and after SC initiated disengagements, were most likely to return to the Road before transitioning elsewhere. Figures 3a and 3b also illustrate that, on average, the time drivers spent in each glance area changed before and after SC initiated disengagements. The mean duration of glances to the Road decreased and was 2sec shorter following the disengagement. The mean durations of glances to the Mirrors and the Down & Center Stack areas decreased from 0.79sec to 0.68sec and from 1.34sec to 0.89sec respectively, and the glance duration to the Instrument Cluster increased slightly following the disengagements (from 0.71sec to 0.84sec).



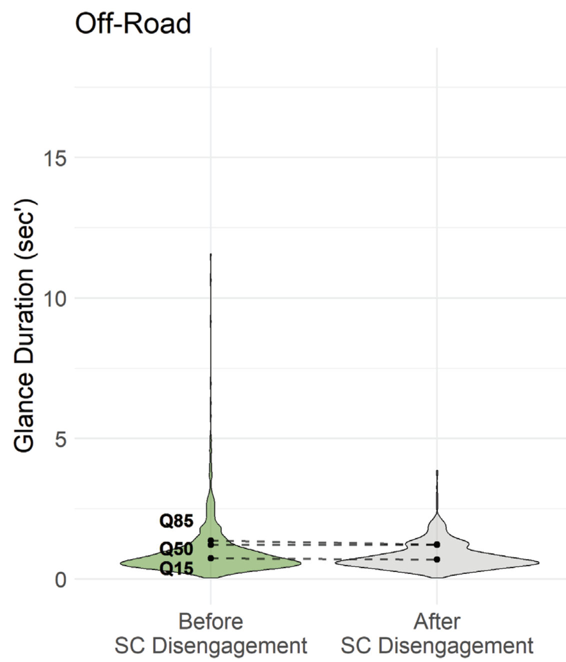
(a) **(b)**

Figure 3. **Transition plots showing the number of transitions (edge width - normalized proportion of transitions) and the mean duration spent in each glance area (vertex size) in the (a) 10sec before and (b) 10sec after SC initiated disengagement.**

Linear quantile mixed-effects models with subject specific random intercept were used to estimate variations in the duration of glances to the Road and to the Off-Road glance areas (i.e., aggregated across Mirrors, Instrument Cluster, and Down & Center Stack) before and after SC disengagement and across the 15th, 50th and 85th quantiles [17]. The duration of long glances to the Road decreased significantly and was 4.86sec shorter following the disengagement ($Q15_{Before}$=0.5, CI=[0-2.24], $Q15_{After}$=0.43, CI=[0-2.11], p=.04; $Q50_{Before}$=2.02, CI=[0.8-3.24], $Q50_{After}$=1.45, CI=[1.02-1.88], p=ns; $Q85_{Before}$=6.63, CI=[2.06-11.2] to $Q85_{After}$=1.77, CI=[0-3.67], p<.001) (Figure 4a). Off-Road glance duration did not change significantly following SC initiated disengagements in any of the evaluated quantiles (see Figure 4b). Analysis of the glance duration for the individual off road glance areas did not yield any significate difference between the period before and after SC initiated disengagements.



(a)

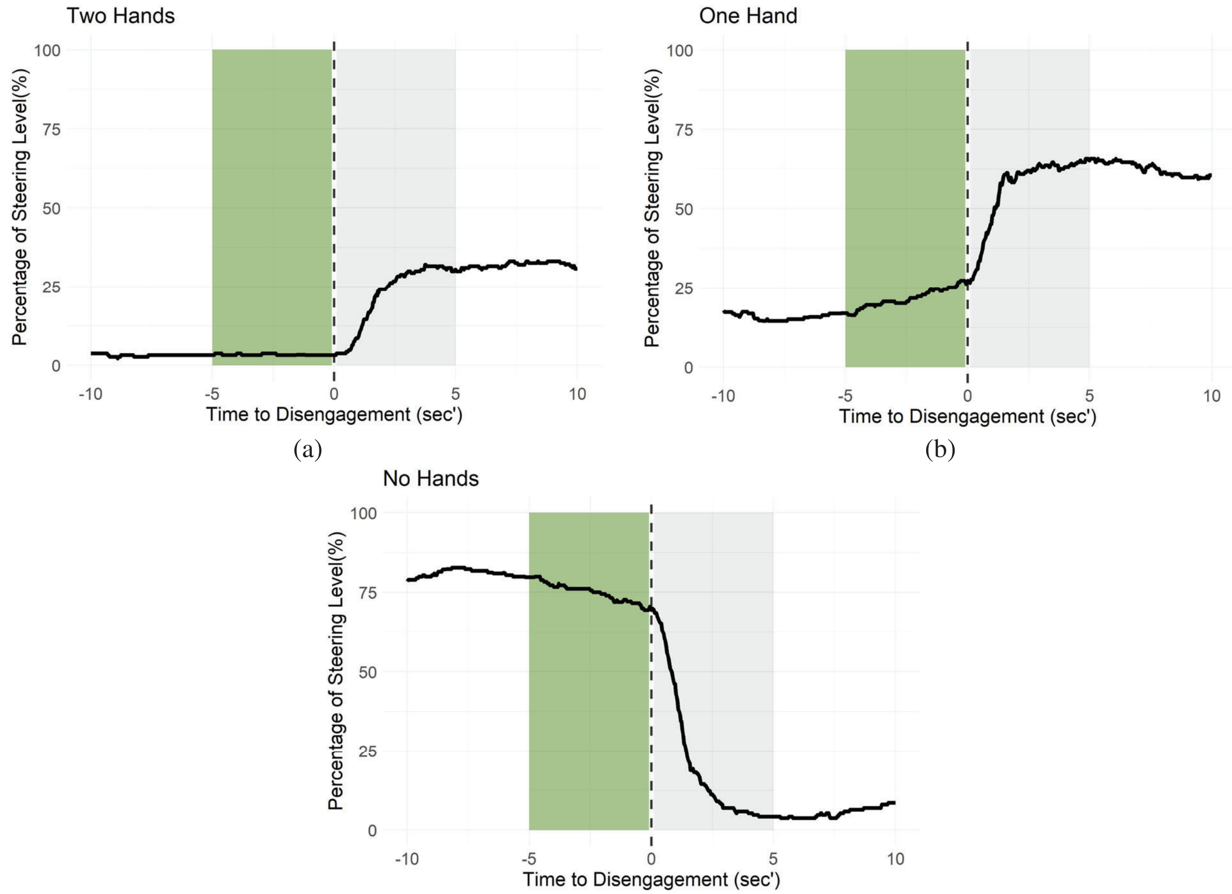


(b)

Figure 4. **Estimated mean glance duration 10sec before and 10sec after SC initiated disengagement across the 15th, 50th, and 85th quantiles, (a) to the Road, and (b) to the Off-Road glance areas.**

Steering Wheel Control Level

Figure 5 shows the percentage of steering wheel control levels expressed by hand on wheel position, across all events during the 10sec before and 10sec after SC initiated disengagements. We find that prior to SC initiated disengagement, on average, drivers drove hands-free 78% of the time and rarely had both hands on the steering wheel (3%). During the transition phase (i.e., in the 5sec around the disengagement), drivers' percentage of steering control level changed dramatically. While at the beginning of the transition phase the proportion of No Hands on the steering wheel was still about 75%, it dropped shortly after SC disengagement to 21% on average. Meaning, shortly after the onset of the disengagement, drivers increased their level of steering wheel control by 54%, grabbing the wheel with at least one hand. A linear mixed-effects model with a driver-specific random intercept was used to assess the relationship between level of steering wheel control and the time it took drivers to takeover in response to SC initiated disengagement. The analysis compared steering wheel control levels (hands-free vs. driving with at least one hand on the steering wheel) at the onset of the disengagement request and the succeeding takeover duration. The disengagement duration from its onset is dictated by the time it takes the driver to resume steering control (i.e., active steering that is reflected by a change in the automation state). On average, the takeover duration when driving hands-free was significantly longer (2.4sec) compared to driving with at least one hand on the steering wheel (1.8sec) [$\chi^2(1)$=4.97, p=.026].



(c)

Figure 5. **Percentage of steering wheel control levels across time and hand position. SC disengagement is at 0sec and marked with a black dashed line. The shaded bands indicate the interval of the transition phase starting 5sec before (highlighted in green) and 5sec after (highlighted in gray) SC initiated disengagement.**

DISCUSSION

This study provides new objective insights on how drivers behave with partial automation system that allows drivers to drive hands free. The study findings capture naturalistic utilization of Cadillac's SC through shifts in attention and hand position during automation-initiated disengagements and transitions of control. Cadillac SC was one of the first systems to use a driver facing camera to infer driver attentional state and issue multimodal alerts when a threshold for apparent inattention is exceeded. Driving safety research has long been concerned with the driver's ability to maintain adequate attention to the driving task when using automation and to regain vehicle control in response to a potential hazard or following automation-initiated disengagement requests [18-19]. Similar to previous studies that focused on driver visual attention when driving with partial automation, we also find that the distributions of glances prior to SC initiated disengagement were relatively stable with glances primarily directed to the Road and to the Down & Center Stack area, and only a limited proportion of glances were directed towards the Instrument Cluster [3, 20]. Our results extend the understanding of drivers' glance behavior around automation-initiated disengagements, indicating that shortly after SC disengagement, there was a sharp and substantial decrease in the proportion of glances that drivers directed to the Road along with an increase in the proportion of glances towered the Instrument Cluster, which lowered attention to all other glance areas. Glances to Down & Center Stack areas are often associated with non-driving related activities, as drivers may glance down when using a smartphone or glance to the center stack while interacting with the infotainment touchscreen.

As part of a driver's response to SC initiated disengagement, the proportion of glances to the Down & Center Stack areas diminished almost to zero and remained low for the rest of the evaluated period. Changes in driver glance behavior following the disengagement were also evident in more frequent transitions between a limited number of glance areas and in that long on-road glances were shorten by 73%. Taken together, these changes in glance behavior resulted in an overall lower and more fragmented visual attention to the Road during a sensitive and possibly time-critical driving situation. This shift in glance behavior is consistent with information seeking behavior that may capture drivers' search for information related to the reason for the disengagement and/or the automation state.

Adequate control of the steering wheel is critical both to correct lateral position in normal driving and to evade emergency situations. There is limited quantitative information on how drivers maintain control of the steering wheel during manual or automated driving, especially when using partial automation systems that allow hands free driving. The current analysis found that drivers adopted the SC's hands-free driving feature to a substantial degree, taking both hands off the steering wheel more than 75% of the time when SC was engaged. When both hands were off the steering wheel, the response to SC initiated disengagement requests took on average an additional 0.6sec for drivers to regain control and actively resume steering which is an increase of 33% compared to drivers who had at least one hand on the wheel. The significance of this time interval should be evaluated in the context of readiness to respond and considering the design of SC which continues its lane-centering support until the driver actively steers.

Limitations and Future Directions

Limitations to the current study include a volunteer sample of drivers that were enrolled for one-month. Driving context was not investigated, although driver visual response has been shown to depend on, for example, the presence of a lead vehicle. Future research could study the interplay between driver behavior and the complexity of the traffic situation. Future work may also consider different types of, and reasons for, system-initiated disengagement. Furthermore, as increases in inattention and distraction may become more pronounced after prolonged periods of automated driving, future work may benefit from longer periods of study to better estimate the longer-term use of automation and document more instances of automation-initiated disengagements. This study provides a reference point for looking at the response behaviors when interacting with other system designs and implementations (e.g., Tesla's Autopilot or Ford's Blue Cruise). Future research should evaluate how the use of partial automation and the characteristics of driver attention monitoring and support systems may impact the way glance behavior and steering wheel control intertwin.

CONCLUSIONS

As driver behavioral adaptation to automation is complex and hard to predict, developing human-centered automated systems will benefit from an evolutionary process that builds upon insights on how drivers use currently available ADAS in real-world settings. The use of partial automation is changing and reshaping drivers' visual and vehicle control behaviors, including the response to automation-initiated disengagements. Based on the observed trends, it is likely that the use of higher levels of automation will introduce even more substantial changes in visual and vehicle control behaviors during automation disengagements. As the performance of automation depends on the interaction between the human and the system, we hope that the data provides useful information to designers charged with developing assistive and automated systems and empowers regulators and safety advocates with data needed to objectively assess how to guide appropriate utilization of such technology.

ACKNOWLEDGMENTS

The data are from work supported by the Advanced Vehicle Technology (AVT) Consortium. The views and conclusions expressed are those of the authors and have not been sponsored, approved, or endorsed by the consortium.

REFERENCES

[1] SAE International, Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (SAE Standard J3016, Report No. J3016-201806) (2018).

[2] Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013). "take over!" how long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1938–1942. https://doi.org/10.1177/1541931213571433

[3] Morando, A., Gershon, P., Mehler, B., & Reimer, B. (2020). Driver-initiated Tesla Autopilot Disengagements in Naturalistic Driving. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. https://doi.org/10.1145/3409120.3410644

[4] Gershon, P., Seaman, S., Mehler, B., Reimer, B., & Coughlin, J. (2021). Driver behavior and the use of automation in real-world driving. *Accident Analysis &Amp; Prevention*, *158*, 106217. https://doi.org/10.1016/j.aap.2021.106217

[5] Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, *68*, 138–145. https://doi.org/10.1016/j.apergo.2017.11.010

[6] Norman, D. A. (1990). The 'problem ' with automation: inappropriate feedback and interaction, not 'over-automation.' *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *327*(1241), 585–593. https://doi.org/10.1098/rstb.1990.0101

[7] Louw, T. (2017). Are you in the loop? Using gaze dispersion to understand driver visual attention during vehicle automation. *CERN European Organization for Nuclear Research - Zenodo*. https://doi.org/10.5281/zenodo.235571

[8] Gaspar, J., & Carney, C. (2019). The Effect of Partial Automation on Driver Attention: A Naturalistic Driving Study. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(8), 1261–1276. https://doi.org/10.1177/0018720819836310

[9] Eriksson, A., & Stanton, N. A. (2017). Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *59*(4), 689–705. https://doi.org/10.1177/0018720816685832

[10] Donmez, B., Boyle, L., Lee, J.D. (2008). Designing Feedback to Mitigate Distraction. Driver Distraction: Theory, Effects, and Mitigation. *CRC Press*, pp. 519–531.

[11] Lee, J.D., Moeckli, J., Brown, T.L., Roberts, S.C., Schwarz, C., Yekhshatyan, L., Nadler, E.,

Liang, Y., Victor, T.W., Marshall, D. (2013). Distraction detection and mitigation through driver feedback *(Report DOT HS 811 547A). National Highway Traffic Safety Administration*.

[12] Reimer, B. (2020). Should Tesla take the initiative to better monitor and manage driver behavior with Autopilot? *Forbes*.

[13] Reimer, B., Angell, L., Mehler, B., Fitch, G. M., Noble, A. M., Feit, S., & Skrypchuk, L. (2022). Managing distraction holistically through driver attention support. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 66, No. 1, pp. 310-314). Sage CA: Los Angeles, CA: SAGE Publications.

[14] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J. M., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., & Knipling, R. R. (2006). The 100-Car Naturalistic Driving Study Phase II – Results of the 100-Car Field Experiment. *Chart*.

[15] Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., ... & Dingus, T. A. (2013). The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk *(No. DOT HS 811 757)*.

[16] Gershon, P., Ehsani, J. P., Zhu, C., Sita, K. R., Klauer, S., Dingus, T., & Simons-Morton, B. (2018). Crash risk and risky driving behavior among adolescents during learner and independent driving periods. *Journal of Adolescent Health*, 63(5), 568-574.

[17] Geraci, M., & Bottai, M. (2014). Linear quantile mixed models. *Statistics and computing*, 24(3), 461-479.

[18] Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation Expectation Mismatch: Incorrect Prediction Despite Eyes on Threat and Hands on Wheel. *Human Factors, 60*(8), 1095-1116. doi:10.1177/0018720818788164.

[19] Yang, S., Kuo, J., Lenn´e, M.G. (2020). Effects of distraction in on-road level 2 automated driving: Impacts on glance behavior and takeover performance. *Human Factors* 0018720820936793. https://doi.org/10.1177/0018720820936793.

[20] Pipkorn, L., Dozza, M., & Tivesten, E. (2022). Driver visual attention before and after take-over requests during automated driving on public roads. *Human factors*, 00187208221093863.

DEVELOPMENT OF A PERFORMANCE-BASED PROCEDURE FOR SAE LEVEL 2 DRIVER ENGAGEMENT ASSESSMENTS

André Wiggerich
Federal Highway Research Institute (BASt),
Germany

Paper Number 23-0162

ABSTRACT

Adapting the performance and design of Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems (ADS) to human capabilities and safety needs is an important requirement for a safe market introduction of new technologies in this field. A specific challenge for SAE Level 2 systems is the engagement of the driver in the driving task. A high level of driver engagement is necessary to ensure that drivers are able to fulfill their role and responsibility to monitor the system performance and to intervene in system limit situations. However, effectiveness of current driver monitoring technologies to ensure driver engagement are limited with broadly diversified performance parameters. Therefore, the aim of the current research was to develop and validate a standardized procedure for performance-based assessments of driver engagement of Level 2 systems with a direct link to safety by focusing on controllability in accident-prone system limit situations. In total, 39 drivers without prior experience in continuously assisting systems participated in the evaluation study on a test track. To assess the validity of the procedure and the standardized test scenario, half of the participants experienced a Level 2 system (Tesla Autopilot) while the other half drove the same vehicle (Tesla Model 3) conventionally (fully manual). The participants task was to constantly follow a lead vehicle, driven by a second experimenter, on a round course for approx. 30 minutes. At the end of the test drive an accident-prone system limit situation without a prior system-initiated warning was triggered: The lead vehicle performed a cut-out maneuver revealing a stationary crash target in front of the participant. Without driver intervention, the Level 2 system was not able to avoid a collision. Therefore, the participant was required to react by braking and/or steering. Results indicate that the test scenario is controllable by conventional drivers. No driver of this group caused a collision or had a Time-To-Collision minimum below 1 second. However, drivers of the Level 2 system specifically used in the study had difficulties in controlling the system limit situation and intervening adequately to avoid a collision with the stationary target. 15% of these drivers collided with the target and approx. 50% had a Time-To-Collision minimum below 1 second. Furthermore, the median Time-To-Collision minimum of the Level 2 drivers was approx. 1.5 seconds lower compared to conventional drivers. Concluding, it can be stated that the test scenario is in general controllable by conventional drivers but, due to a lack of driver engagement of the Level 2 system tested in the study, participants of this group had problems in intervening adequately to the system limit. In summary, the developed procedure is a pragmatic, reliable and valid way to assess driver engagement independently and in a design-neutral way by focusing on safety critical interaction behavior in system limit situations.

INTRODUCTION

Safety of SAE Level 2 Systems is strongly linked to a high level of driver engagement. This is especially important as drivers are required to intervene immediately to system limits, even without any prior warning or information given by the system [1]. Hence, driver engagement includes a continuous monitoring of the system performance as well as of the driving environment by the driver, being able to identify situations requiring an immediate intervention. Taking this into account, driver engagement is at least a visual as well as a cognitive task [2]. However, effectiveness of current driver monitoring technologies to ensure driver engagement is rather limited. These technologies are able to asses motoric (hands on the steering wheel) as well as visual (eyes on the road) driver behavior only. There is currently no technical solution available in series production vehicles to assess if a driver has understood the system performance as well as the driving situation by means of driver monitoring technologies. Furthermore, although all of the systems available today are labeled as Level 2 systems, systems of different manufacturers show a different operational behavior and system performance. Hence, these differences in system design can influence the quality and safety of human-machine-interaction (HMI) and driver engagement in several ways making it an important topic for consumer protection organizations and type approval authorities to focus on [3]. Therefore, the aim of the current research is to develop and validate a standardized procedure for performance-based assessments of driver engagement of Level 2 systems with a direct link to safety by focusing on controllability in accident-prone system limit situations.

Safety assessments of series production vehicles in Europe are performed by the European New Car Assessment Program (EuroNCAP). In 2018 and 2020, EuroNCAP already performed safety assessments of Level 2 systems by a separate grading in addition to the regular safety rating [4, 5]. For the first time, the 2020 grading used a balance principle for the final scoring and grading, combining the assessment of vehicle assistance competence on the one hand as well as driver engagement on the other hand. The balance principle was introduced to emphasize the importance of a good driver engagement in relation to an increasing vehicle assistance. Finally, considering also safety backup functions in addition, a 4-point grading system, from 'entry' over 'moderate' to 'good' and 'very good', was created. To assess the driver engagement component of the grading, consumer information, system status, driver monitoring and driving collaboration aspects are analyzed by EuroNCAP experts. The assessment basically is focused on concrete system design aspects, for example display icons and colored indicators for certain system status information. However, the safety impact of these design aspects in relation to real driver behavior was neither parameterized nor analyzed so far. The balance principle here is lacking a real comparison of objective driver behavior to objective system performance. Assessing driver engagement directly by analyzing real driver behavior, could create a more naturalistic and valid assessment approach with a direct link to measurable safety outcomes and can therefore close this gap. However, the feasibility and reliability of direct driver-based assessment procedures in consumer protection is still an open research question.

Controllability and Safety Assessment of Human-Machine-Interaction

Safety assessments of HMI for assisted and automated driving are often performed by human factors experts based on a set of defined design requirements and design checklists (for example see [6]). The requirements or checklist items are usually derived from international standards or guidelines. However, the safety outcome is limited to the individual interpretation of the requirements by the experts. Furthermore, requirements consider only individual aspects of a safe system design (for example certain display icons for system status indications), but usually do not testify about the user-system-interaction as a whole (for example do users understand their role and responsibility in interaction with the system?). This can only be done by analyzing user understanding and behavior in concrete interaction with a holistic system. The Take-over controllability (TOC) rating for Level 3 systems, developed by [7], goes one step further. Video material of control transitions from automated to manual driving is assessed by trained raters in a standardized procedure. The TOC rating is used to analyze concrete and observable driver behavior in a safety relevant scenario, the takeover situation. However, since the TOC rating is limited to control transitions from automated to manual driving only and safety of HMI with regard to driver engagement for assisted driving requires different aspects, at least an adaptation of the procedure and rating categories seems necessary. There is a strong relation between research on controllability and the safety assessment of HMI. However, relevant safety indicators are often operationalized and assessed heterogeneously. [8] and [9] applied subjective ratings as well as objective criteria in different driving simulator studies, to assess controllability of Level 2 systems. On a test track, [10] tested the controllability in accident prone-system limit situations in different experiments. The authors describe, that participants often showed severe problems in controlling the vehicle in front of a target, although having their hands on the steering wheel as well as their eyes on the road. [11] found similar effects. Results suggest the importance for consumer protection and type approval to take a closer look at controllability and safety of HMI and driver engagement. However, appropriate procedures are still missing here.

Research Questions

To address the feasibility, reliability and validity of driver-based assessment procedures for safety assessments of HMI and driver engagement of SAE Level 2 systems, the study focused on three different research questions: 1) Is it possible to measure and assess safety of HMI of SAE Level 2 systems in a standardized procedure with performance-based metrics? 2) Do drivers of a Level 2 System show a different interaction behavior in a safety critical test scenario compared to conventional drivers driving completely in manual control without any continuous system support? 3) Is the defined test scenario in general controllable by conventional drivers driving completely in manual control?

METHODS

Test vehicle

The study was conducted with a Tesla Model 3 (model year 2020), using the Level 2 system Tesla Autopilot (ACC combined with Tesla Auto-Steer) in the Level 2 system condition (software version v10.2, 2021.4.18.2).

While using the system, drivers are required to permanently leave their hands on the steering wheel. Taking hands away from the steering wheel would result in hands-off warnings. ACC speed should be set to 60 kph by the participants. The initial distance to the lead vehicle was put to the closest setting. However, it was up to the participant to change the distance to another setting, but no participant made use of this option. In the control condition, participants where driving the same vehicle conventionally (fully manual) without any continuous system support. They were asked to keep a close distance to the lead vehicle, as they would usually do when driving at a similar speed on a two-lane road.

Test track

The study was conducted on a test track in Bad Sobernheim, Germany. The 1.6 km-long one-way round course with two fully marked lanes (lane width: 3.30 m) provided different radii of left and right curves as well as straight sections of different length to allow participants to experience the system behavior in different driving situations (see Figure 1). Due to safety reasons, apart of a second experimenter driving a vehicle in front of the participant during the whole experiment, there was no surrounding traffic or other road users on the track.

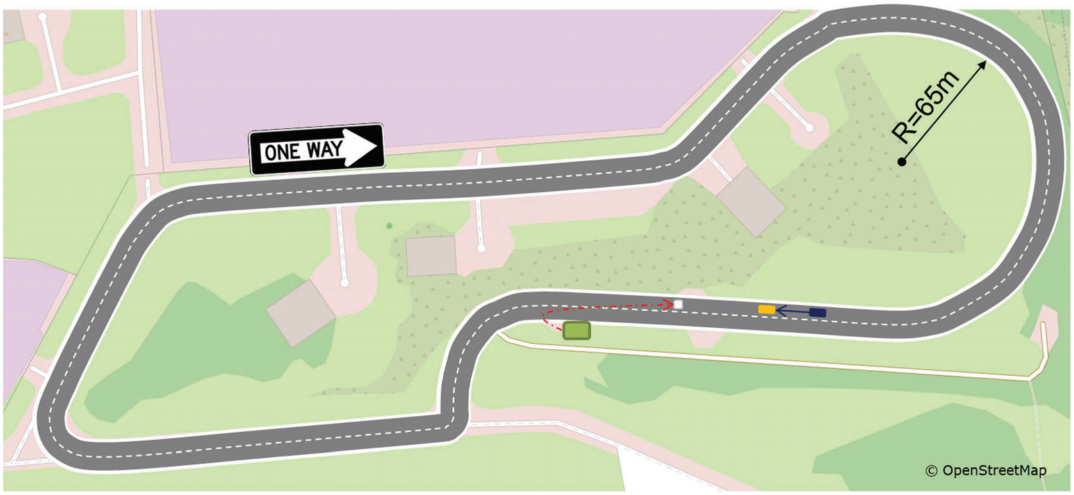


Figure 1. Test track in Bad Sobernheim, Germany.

Test Scenario

To be a suitable test scenario for the safety-related results of adequate driver engagement in Level 2, we consider the situation needs to be a system limit requiring the human driver to intervene. To emphasize the intervention by the driver, the situation should be highly salient and intuitively understood by manual drivers as well as drivers using a Level 2 system. Therefore, a standardized vehicle cut-out scenario was used in the study. A crash target was positioned manually in the center of the driving lane on a straight section of the test track. To reduce the possibility for the participant to anticipate the situation beforehand, the target was hidden in a tent next to the track prior to the scenario. When the lead vehicle, driven by a professional driver, approached the stationary target at a speed of 60 kph, the driver performed an evasive steering maneuver with a full lane change to the left lane. The driver was trained to start the maneuver at a TTC of approximately 1 s (distance to target approx. 15 m) with a maximum lateral acceleration of approximately 7 m/s^2. In case the participant did not react early enough, a Forward Collision Warning (FCW) was issued and an Autonomous Emergency Braking (AEB) maneuver started. However, without driver intervention, a collision with the crash target was caused. The participant could avoid a collision by braking or steering to the left lane. Due to the curvature, the target was visible for the participant for at least 10 seconds prior to the collision.

Study Procedure

Arriving at the test track, participants were welcomed and told a cover story that in the following they would experience a new vehicle and should assess its use and comfort during a test drive. After signing an informed consent form, participants received a safety briefing about safe behavior at the test facility. Afterwards, they were asked to read an excerpt of the user manual. In the Level 2 system condition, a part of the user manual explaining ACC and Auto-Steer was provided to participants while participants in the manual driving condition were provided with an excerpt of general vehicle functions. The user manual of the Level 2 system instructed drivers to always keep their hands on the steering wheel and pay attention, since use of the system can be limited at any time and the driver is still fully responsible for driving safely. In case of a system limit, drivers were told to intervene immediately. After reading the user manual, participants were escorted to the test track and seated in the driver seat. One experimenter took a seat on the passenger seat right next to the participant. The experimenter explained briefly the general vehicle control as well as basic system functionality. The test drive was started, if participants indicated that they felt comfortable with the vehicle control. The experimenter on the passenger seat gave the instructions to drive in the right lane, follow the lead vehicle driving in front and use the Level 2 system as often as possible during the drive. Overtaking was not permitted. The speed of the Level 2 system should be set by the participants to 60 kph and distance to the vehicle ahead was pre-set to the closest possible option. The distance to the lead vehicle could be changed by the participant, but no participant attempted to do so. Manual drivers were asked to follow the lead vehicle as close as they would usually do when driving at a similar speed. Effective driving speed during the experiment was adjusted to the respective track section by the second experimenter (professional driver) driving the vehicle in front. After 14 laps (approximately 30 minutes driving time), a standardized system limit test scenario was initiated by a third experimenter who was waiting out of sight of the participant around the track. The third experimenter placed a crash target in the center of the driving lane (see Figure 2). Beforehand, the target was hidden in a tent right next to the test track, so that the scenario and situation could not be anticipated by the participant. When the second experimenter driving the lead vehicle performed a sudden lane change in front of the target to the left lane, the target came into direct sight. It was now the participant's task to avoid a collision with the stationary target by immediately intervening by braking and/or steering.

Figure 2. Left: Schematic overview of the cut-out maneuver performed by the lead vehicle (green) in front of the stationary crash target (white). Center, Right: The target is hidden in a tent and not visible for participants prior to the test scenario, where it is placed in the center of the driving lane by a third experimenter.

During the whole test drive, participants were observed by the first experimenter on the passenger seat and their interactional behavior was assessed using a standardized, tablet-based assessment and rating tool developed by [12]. Altogether, each test drive lasted about 35 min. Participants received a financial compensation of 30 Euros for their participation in the study.

Test Criteria

The assessment was performed according to recommendations of the RESPONSE III Code of Practice (CoP) [15]. The CoP defines that 85% of a driver population is able to control a defined system limit situation, if all drivers (100%) from a representative sample of n=20 participants are able to control the vehicle in the respective situation. The CoP recommends that the definition of adequate pass-fail criteria should be situation and system specific. Accordingly, for this study, in total three different criteria based on the dependent measures were defined:

- All participants (100%) are able to avoid a collision with the stationary target in the cut-out situation [15].
- The mean TTCmin value of each test vehicle group to the stationary target is >1.0 s ([16], [8]).
- Each participant´s reaction in the cut-out situation is assessed uncritical (< 7) on the standardized scale (see Figure 3) for a general safety rating of driver-vehicle-interaction by the first experimenter [8].

The scenario is failed, if one or more of the three criteria are failed.

Figure 3. General safety rating of driver-vehicle-interaction, standardized on a 11-point scale based on [13] and [14].

Experimental Design
The study was carried out in a between-subjects design with one experimental factor ´Level of Assistance`. Participants were either experiencing a Tesla Autopilot SAE Level 2 system ('Level 2') during the test drive or driving fully manually without any continuous system support ('Conventional Driving'). Participants were assigned randomly to the experimental condition. All participants independent of experimental condition experienced the standardized system limit at the end of the driving trial. There was no variance in order or length of the driving trial.

Participants
In total, *n*=39 drivers (12 females) participated in the study. Participants were recruited by newspaper and social media announcements in the local area of the test track. Mean age of the recruited sample was 41.31 years (*SD*=15.74, *MIN*=19, *MAX*=68). All participants held a valid driver's license, had normal or corrected to normal vision and no previous practical experience with advanced driver assistant systems (e.g. ACC, active lane centering etc.). With regard to age, gender or reported annual driving experience, the two experimental groups did not differ significantly. In total, *n*=20 participants were experiencing the Tesla Autopilot system ('Level 2') while *n*=19 participants were driving fully manually ('Conventional Driving').

RESULTS

In the following, the impact of the experimental factor ´Level of Assistance` on the different objective and subjective dependent measures of the study is described and analyzed. For the cut-out scenario, the mean TTCmin-values of the test vehicle to the stationary target as well as the frequency of collisions with the stationary target are used as dependent objective variables and the general safety rating of driver-vehicle-interaction by the first experimenter as a dependent subjective measure.

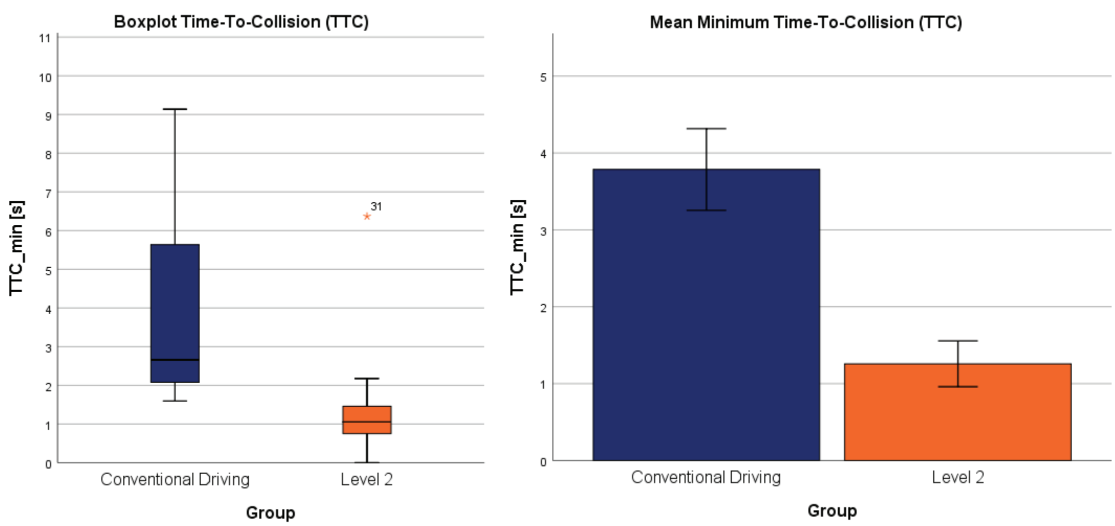


Figure 4. Left: Boxplots for the TTCmin-values by experimental condition (group). Whiskers represent minimum/maximum values within ± 1.5*IQR. 1st Quartile defines, that 25% of the TTCmin-values are lower. 3rd Quartile defines, that 75% of the TTCmin-values are lower. Right: Mean TTCmin-values by experimental condition.

The TTCmin-value distribution as well as the mean TTCmin-values of two different groups for the scenario are shown in Figure 4. The factor ´Level of Assistance` has a significant influence on the reaction of the drivers in front of the target, leading to significantly lower TTCmin-values for drivers driving with the Tesla Autopilot system (*M*=1.26s, *SD*=1.30s) compared to drivers driving fully manually without any continuous system support (*M*=3.79s, *SD*=2.25s, $t(28.45)=4.15$, $p<.001$). The mean TTCmin of manual drivers is more than 2.5s higher compared to the Tesla Autopilot drivers. Although the mean TTCmin-value for both groups is higher than the defined critical value of 1s, results show that 50% of Tesla Autopilot drivers have a TTCmin value <1s compared to 0% of drivers driving fully manually. However, the second controllability criterium is passed for both groups.

Taking the frequency of collisions with the stationary target vehicle into account, on a descriptive basis, there is a tendency towards higher collision rates for drivers driving with the Tesla Autopilot system (*n*=3) compared to drivers driving fully manually (*n*=0), but the difference is statistically not significant (X^2 =3.07, *df*=1, *p*=.079). However, since it was defined that all participants (100%) should be able to avoid a collision with the stationary target in the cut-out situation, the first controllability criterium was failed for the Tesla Autopilot system.

The driver's response towards the stationary target is further analyzed by means of the standardized safety ratings of the first experimenter. As can be seen in Figure 5, the response is rated dangerous and critical (standardized scale value >7) for 20% of the drivers driving with the Tesla Autopilot system compared to 0% of drivers driving fully manually. Therefore, the third controllability criterium is failed for the Tesla Autopilot system. Furthermore, the response for drivers driving with the Tesla Autopilot system is also rated significantly more critical on average (*M*=4.00, *SD*=3.23) compared to drivers driving fully manually (*M*=0.47, *SD*=0.96, *t*(22.53)=-4.67, *p*<.001).

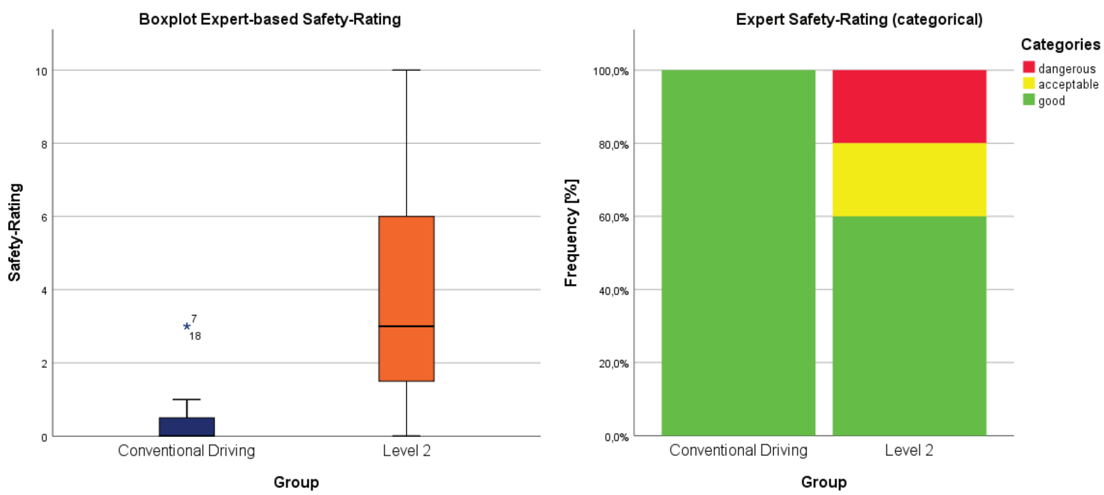


Figure 5. Left: Boxplots for the standardized safety ratings of the first experimenter by experimental condition (group). Right: Frequency of safety ratings per category by experimental condition.

Considering the objective TTCmin-values as an external validity criterium for the measurement of the subjective safety rating by the first experimenter, there is a strong correlation effect of *r*=-.574 (*p*<.01) for the cut-out scenario. The standardized safety ratings can therefore be seen as a valid indicator for HMI safety assessments. Furthermore, considering reliability of the method, there is also a strong inter-rater agreement between the first experimenter performing the rating within the vehicle and a second rater analyzing recorded video data (*ρ*=.706, *p*<.01).

DISCUSSION

The aim of the current research was to address the feasibility, reliability and validity of a driver-based assessment procedure for safety assessments of HMI and driver engagement for SAE Level 2 systems. In total, 39 drivers without prior experience in continuously assisting systems participated in the study on a closed test track. To assess the validity of the procedure and the standardized test scenario, one half of the participants were experiencing a Level 2 system (Tesla Autopilot) while the other half were driving the same vehicle conventionally (fully manual). The participants task was to constantly follow a lead vehicle, driven by a second experimenter, on a round course for approx. 30 minutes. At the end of the test drive an accident-prone system limit situation without a prior system-initiated warning was triggered: The lead vehicle performed a cut-out maneuver revealing a stationary crash target in front of the participant. Without driver intervention, the Level 2 system was not able to avoid a collision neither was the AEB-System given the initial following distance to the vehicle in front and speed. Therefore, the participant was required to react by braking and/or steering. The study aimed at assessing safety of HMI in a standardized procedure with performance-based metrics. The first research questions focused on the possibility to use these metrics to define interaction-related safety outcomes. In accordance to the RESPONSE III CoP three different pass-fail criteria were defined. Results show that all three criteria could be applied for the safety assessment in a practical way. Approx. 15% of the Level 2 drivers were not able to avoid a collision with the stationary target and 20% of these drivers were rated critical or dangerous on the standardized safety rating scale by the first experimenter. Therefore, the first and the third safety criteria was failed by the respective Level 2 system (Tesla Autopilot) used in the study. The second criteria did not focus on individual participants behavior but the mean TTCmin on a group level. As defined by [16], a TTCmin below 1s is related to a safety critical situation. Due to one outlier with a TTCmin of 6.3s in the group of the Level 2 drivers, the mean TTCmin of the group was >1s. However, 50% of drivers in this group still had values below 1s compared to no driver in the group of conventional drivers. It is therefore questionable, if the criterium is useful and valid on a group level by calculating the mean value. The assessment could probably be improved by defining specific quantities or frequencies (e.g. median or 1st quartil >1s).

The second research question focused on the behavior differences of Level 2 drivers and conventional drivers. Results indicate a difference between the two groups. No manual driver caused a collision in the test scenario and the reaction of all manual drivers was rated 'good' by the trained rater. Furthermore, the TTCmin of all conventional drivers was above 1 s. Therefore, the procedure seems to be a suitable approach to detect and assess difference in safety relevant driver behavior in the chosen scenario. Since no one of the conventional drivers showed any difficulty in avoiding a collision with the stationary target, the scenario itself seems to be in general controllable. Therefore, the third research question can clearly be answered yes. However, results also indicate that there is a large variance in the behavior of manual drivers in the face of the stationary target. Some drivers changed the lane even before the lead vehicle started the cut-out maneuver. This shows that human drivers are highly capable of anticipating risky situations, even before they actually happen. However, in most cases Level 2 drivers in the study lacked this anticipatory behavior. Most of these drivers waited for the system to solve the situation or warn the driver to intervene. Although all drivers independently of experimental condition had their hands on the steering wheel and their eyes on the road in the test scenario, the reaction of Level 2 drivers was delayed most of the time. Manual drivers often reduced their speed and increased their distance to the lead vehicle when the stationary target became visible for the first time, approximately 5-10 s prior to the calculated target impact. By doing so, these drivers gained time to observe what would happen in the ambivalent situation, long before the actually critical cut-out maneuver would take place. Level 2 drivers in the study usually missed their opportunity to increase distance to the lead vehicle when the target became visible. One reason for this could be that increasing the distance to the lead vehicle while using ACC would either take some time and require several pushes on a button or mean to directly lose system support completely by applying the brakes. Assuming that drivers are motivated to maximize their individual use of the system, this might demotivate them from using the brakes early to increase the distance to the lead vehicle when the situation is not yet a clear system limit.

CONCLUSIONS

Based on the experience of the study, the standardized assessment procedure for safety of human-machine-interaction was perceived pragmatic and feasible for testing purposes of consumer protection and type approval. There was a large difference in safety indicators between conventional driving and the Level 2 system used in the study. This is an indicator, that the test scenario is in general controllable by human drivers and due to specific interaction related problems of Level 2 systems, drivers show difficulties in intervening at the system limit. It became clear, that manual and Level 2 drivers showed different behavior patterns facing the target situation. However, the concrete variance between different Level 2 systems is still unclear. More experience and data are necessary for a comprehensive consideration and further development of the method.

REFERENCES

[1] SAE International Standard J3016. (2021). "Taxonomy and Definitions for Terms Related to On-Road-Motor Vehicle Automated Driving Systems"

[2] Pradhan, A. K., Crossman, J., & Sypniewski, A. (2019, June). Improving driver engagement during L2 automation: a pilot study. In *Driving Assesment Conference* (Vol. 10, No. 2019). University of Iowa.

[3] Carsten, O., & Martens, M. H. (2019). How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work*, 21(1), 3-20.

[4] EuroNCAP (2018). 2018 Automated Driving Tests: https://www.euroncap.com/en/vehicle-safety/safety-campaigns/2018-automated-driving-tests/

[5] EuroNCAP (2020). 2020 Assisted Driving Tests: https://www.euroncap.com/en/vehicle-safety/safety-campaigns/2020-assisted-driving-tests/

[6] Naujoks, F., Wiedemann, K., Schömig, N., Hergeth, S., & Keinath, A. (2019). Towards guidelines and verification methods for automated vehicle HMIs. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60, 121-136.

[7] Naujoks, F., Wiedemann, K., Schömig, N., Jarosch, O., & Gold, C. (2018). Expert-based controllability assessment of control transitions from automated to manual driving. *MethodsX*, 5, 579-592.

[8] Naujoks, F., Purucker, C., Neukum, A., Wolter, S., & Steiger, R. (2015). Controllability of partially automated driving functions–does it matter whether drivers are allowed to take their hands off the steering wheel?. *Transportation Research Part F: Traffic Psychology and Behaviour*, 35, 185-198.

[9] Naujoks, F., Purucker, C., Wiedemann, K., Neukum, A., Wolter, S., & Steiger, R. (2017). Driving performance at lateral system limits during partially automated driving. *Accident Analysis & Prevention*, 108, 147-162.

[10] Victor, T. W., Tivesten, E., Gustavsson, P., Johansson, J., Sangberg, F., & Ljung Aust, M. (2018). Automation expectation mismatch: Incorrect prediction despite eyes on threat and hands on wheel. *Human Factors*, 60(8), 1095-1116

[11] Pipkorn, L., Victor, T. W., Dozza, M., & Tivesten, E. (2021). Driver conflict response during supervised automation: Do hands on wheel matter?. *Transportation Research Part F: Traffic Psychology and Behaviour*, 76, 14-25.

[12] Schömig, N., Wiedemann, K., Julier, R., Neukum, A., Wiggerich, A., & Hoffmann, H. (2021). Methoden für die Bewertung der Mensch-Maschine-Interaktion beim teilautomatisierten Fahren. Berichte der Bundesanstalt für Straßenwesen, Reihe F: Fahrzeugtechnik (F 141). Bremen: Fachverlag NW in der Carl Ed. Schünemann KG.

[13] Kaussner, Y. (2013). Assessment of driver fitness: An alcohol calibration study in a high-fidelity simulation. Paper presented at the Fit to Drive 7th International Traffic Expert Congress, Berlin, 25.04.-26.04.2013.

[14] Kaussner, Y., Hoffmann, S., Fischer, F. & Kenntner-Mabiala, R. (2013). SPDE_DFA: S.A.F.E: Anleitung zur Durchführung von Fahrverhaltensbeobachtungen mit der Standardized Application for Fitness to Drive Evaluations (S.A.F.E.) am Tablet-PC. Würzburg, 2013.

[15] RESPONSE Consortium (2006). Code of Practice for the Design and Evaluation of ADAS. *RESPONSE 3: A PReVENT Project.*

[16] Hayward, J. C. (1972). Near-miss determination through use of a scale of danger. *Highway Research Record* (384).

RESEARCH ON ATTENTION KEEPING TECHNOLOGY TO REDUCE CARELESS DRIVING ACCIDENTS

Ryo Kimura, Tsuyoshi Nojiri, Hiroko Adachi,

Honda R&D Co., Ltd.

Japan

Paper Number 23-0195

ABSTRACT

We explore how a driver's attention changes when exposed to auditory and haptic stimuli. In many cases, accidents caused by internal factors, typified by drowsy driving and careless driving, result in serious accidents Therefore, it is important to keep the driver from losing attention. Thus far, manufacturers have built systems to promote arousal by engaging visual and auditory alarms when the driver loses attention. However, in such cases the driver has already lost attention when the alarm is engaged. Depending on driving circumstances, it is impossible to immediately stop the vehicle or take a break, and in some cases the driver has no choice but to continue driving despite their decreased level of attention. This paper focuses on auditory and vibratory stimuli as realistic methods of stimulus that can actually be supplied to the driver by products, and establishes stimulus methods with indications of a relationship to autonomic nerve activity. The paper also evaluates the effectiveness of these methods in preventing loss of attention, by supplying the established stimulus under conditions in which drivers have begun to lose attention in the past. The stimulus methods are as follows. (1) Music: Comparatively high-tempo music at around 100–130 beats per minute (bpm). (2) Music with amplified bass: The same music described in (1), but with its bass range amplified. (3) Music with vibration: Music with superimposed vibrations in sync with the bpm of the music described in (1), from a sound source in the seatback peaking at 60 Hz. Test subjects were put into a driving simulator that employed these stimuli and asked to follow the vehicle in front of them for 30 minutes. Each of the 11 test subjects repeated four trials, including trials with no stimuli. Several indicators were collected during these trials. For driving behavior, the indicator was Time-to-Collision (TTC), for subjective

sleepiness it was time-dependent change in the Karolinska Sleepiness Scale (KSS), and for a parasympathetic indicator it was Standard Deviation of NN intervals (SDNN), which is the standard deviation of the R-R Interval. The eight test subjects for which data was properly collected tended to exhibit higher minimum TTC during the trails with stimulus than those without. This increased greatly ($p < 0.01$) when bass amplification and superimposed vibrations were used. It was found that subjective sleepiness, which was the time for which test subjects were aware of being sleepy, decreased by 58% with the music stimulus (1), 86% with the music with amplified bass stimulus (2), and 77% for the music with vibration stimulus (3), compared to with no stimulus. Moreover, the SDNN trend revealed parasympathetic acceleration when there was no stimulus, but this was suppressed for both (1) music only and (3) music with vibration. For (2) music with amplified bass in particular, it remained in the same state from the beginning of the test.

Applying the knowledge above to inhibit loss of concentration before it occurs can be expected to help reduce traffic accidents associated with internal factors such as drowsy driving and careless driving.

INTRODUCTION

Honda aims to achieve zero fatalities in traffic accidents involving Honda motorcycles and four-wheeled vehicles by 2050. In Japan, internal factors typified by drowsy driving and careless driving account for 26.2% of the human factors in traffic fatalities (505 incidents) [1]. In the United States, it is estimated that drowsy driving leads to 90,000 collisions, 50,000 traffic accidents resulting in injury, and 800 fatalities per year [2]. In this way, drowsy driving and careless driving often lead to serious accidents. That is why it is important to help prevent the loss of driver attention brought about by drowsy driving and careless driving. Attention detection systems that detect loss of attention, through wandering of the driver’s vehicle and drive duration, are already in practical use. These systems output visual and auditory alerts when a loss of attention is detected. Also, in recent years, cars have started coming to market with functions that assess sleepiness based on facial features obtained from onboard driver monitoring cameras. When the system determines that the driver is sleepy, these functions trigger visual and auditory alarms. All detection/alarm systems are installed with the goal of encouraging the driver to take a break. But in the interest of acceptability, most of them do not activate an alarm until the driver’s level of sleepiness has reached a certain point. However, in such cases the driver has already lost attention when the alarm is engaged. Depending on driving conditions, it is not possible to immediately stop the vehicle or take a break, and in some cases the driver has no choice but to continue driving despite their decreased level of attention. In addition, it is challenging to wake a driver from a state of deep sleepiness of which they themselves are aware [3]. While an alarm stimulus may temporarily restore arousal, as the stimulus is intended to encourage the driver to take a break, it cannot be expected to be continually effective. Because of this, it is hoped that a method for continuously maintaining concentration and inhibiting sleepiness will be found.

It is known that the parasympathetic nerves of the autonomic nervous system are accelerated when people have lost attention during drowsy and careless driving. While it is challenging to control the autonomic nervous system through the force of one's own will, it is known to change in response to various external stimuli. Therefore, it stands to reason that there needs to be a stimulus that can inhibit the acceleration of the parasympathetic nerves in order to maintain concentration while continuing to drive. One example of a phenomenon in which a person's state can change in response to an external stimulus is that known as entrainment. Entrainment describes the phenomenon in which two different rhythms that are close to each other naturally synchronize. It is said that a similar thing happens in humans [4]. There are many events in daily life in which a person's state is changed or guided by entrainment. For example, it has been reported that entrainment, such as humming along when listening to music, becoming one with the music at a live concert and getting excited, or listening to sounds that are a different tempo from one's own heart rate, all influence the autonomic nervous system [5][6]. Entrainment of the autonomic nervous system often focuses on the tempo of music, and it has been reported that the tempo of a song itself can influence how the subjective groove of the music is felt, and influence the autonomic nervous system [7][8]. Also, when attention is given to what qualities of music make it easy to synchronize with, it turns out that such music has many low sounds such as drums [9]. It has been reported that emphasizing beat influences the autonomic nervous system and the desire to move one's body [10][11]. As can be seen, there are numerous examples of studies that examine musical tempo and beat as individual elements, but there are few looking into their combined effects. In addition, while there have been reports on the effects of sound and music stimuli on driving behavior, few studies have mentioned their relationship to subjective sleepiness and autonomic nerve activity.

As such, this study focused on loss of attention due to sleepiness, which is considered to be the most influential factor in loss of attention while driving. Using music focused on entrainment, it evaluated the effect of external stimuli on driving behavior, subjective sleepiness, and autonomic nerve activity, with the goal of developing a human–machine interface (HMI) that can inhibit loss of attention due to sleepiness. In considering stimuli, the study sought to use the vehicle's onboard audio as a stimulus to induce entrainment and intervene in a natural, unobjectionable way. Also, as vibrations have been added to some vehicle seats in recent years, the effect of superimposing vibrations on the music using sound from the seatback was also examined. In the experiment, test subjects were placed in a driving simulator (DS) and told to follow the car ahead of them. This monotonous driving situation can easily induce loss of attention. To help test subjects maintain attention, the music and music with superimposed vibrations mentioned earlier were supplied as stimuli, and the effectiveness of these stimuli in inhibiting the loss of attention was evaluated.

This experiment has been reviewed by the Bioethics Committee meetings for Honda's R&D activities. (Bioethics Committee No. 98HM-056 H)

TESTING METHOD

DS Tasks

In this study, researchers constructed an environment to allow the test subjects to perform the task of following the vehicle ahead of them. The constructed course was a straight road with three lanes, and test subjects were instructed to drive in the middle lane. Figure 1(a) shows the experimental environment and Fig. 1(b) shows an example of what participants saw as they drove during the experiment.

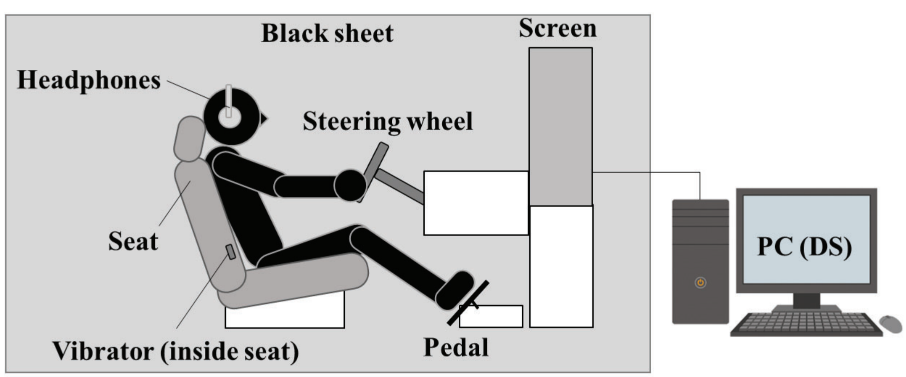


Figure 1(a). Experimental environment

Figure 1(b). Example test run

At the start of the test, the initial vehicle speed was 70 km/h and the initial distance from the vehicle ahead was 30 meters. Test subjects were not told how much distance to leave between their car and the vehicle ahead, and were asked to drive at a distance that they did not feel was dangerous. Test subjects drove for 30 minutes under each condition. Four tests with different conditions were conducted over 4 days, for a total of 120 minutes. To operate their own vehicle, test subjects needed to operate the accelerator and brake to maintain a suitable distance from the car ahead, and steer to maintain a straight line. The car being followed in the simulation was configured to perform two types of random deceleration, as shown in Fig. 2, slowing down to either 50 km/h or 30 km/h. Test subjects were instructed to perform the necessary operations to maintain distance from the car ahead when they noticed it slow down. In order to encourage loss of attention, the car in front was programmed not to decelerate until 15 minutes after the start of the test. Also, to help prevent test subjects from simply reacting to the brake lights of the car ahead, its brake lights were set to stay off. The experiment was also conducted in an environment in which light was blocked by dark curtains to induce sleepiness, as shown in Fig. 1(a).

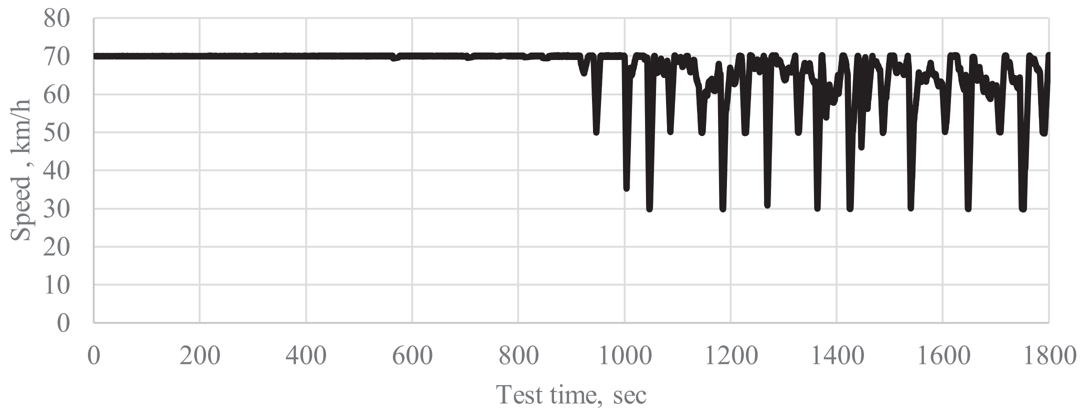


Figure 2. The behavior of the car ahead

Moreover, test subjects wore headphones when performing the driving task, to play back the driving/engine sounds from the simulator, and to present the type of musical stimuli described in the following section.

Stimulation Method

Test subjects performed operations while listening to sounds generated by their own vehicle and the surrounding vehicles that were output from the driving simulator, and music that was output from a PC. Tests were conducted with and without HMI stimuli. Scenario A used no HMI stimulus, and scenario B used HMI stimulus. Before the test, test subjects selected at least five songs that they would like to listen to during the experiment. These were chosen from a set of songs prepared by the experimenters (J-pop and contemporary Western music: 100–130 bpm). The selected music was arranged in a musical playlist set to start playing when the test began, and loop until the test ended. The sound volume during the test was set to a level of 65 dB, and then decreased from that level to a level that the test subject found acceptable. In scenario B1, the selected songs were played without processing as an original sound source. In scenario B2, the selected songs were first processed using the music creation software Cubase to amplify low sounds of 200 Hz or below, and then played back as a sound source. In scenario B3, vibrations corresponding to the bpm of each song were supplied from an vibrations installed in the lumbar section of the seatback (Fig. 1(a)) to correspond with the playback of the original sound source. The oscillator used a Vibro-Transducer VP to convert the electronic signal of the sound into mechanical vibrations, and employed a sound source with a peak of 60 Hz. Furthermore, the music used in scenario B was the same for each test subject.

Table 1 below shows the stimulus levels employed this time.

Table 1. Test stimulus levels

<table>
<tr><th colspan="2">Scenario No.</th><th>Test condition (HMI)</th></tr>
<tr><td colspan="2">A</td><td>No music</td></tr>
<tr><td rowspan="3">B</td><td>B1</td><td>Music(100-130 bpm)</td></tr>
<tr><td>B2</td><td>Music with amplified bass</td></tr>
<tr><td>B3</td><td>Music with vibration</td></tr>
</table>

Test Subjects

Eleven healthy adults from whom informed consent had been obtained (six men, five women, average age 37) participated in the experiment. Each held an ordinary motor vehicle license and drove on a daily basis. Physiological indicators could not be properly obtained from some participants, and others were already quite sleepy at the start of the test. In total, accurate data could not be obtained for three people. Therefore, the following data was collected and processed from eight participants.

Measurement Items

Subjective sleepiness index The Karolinska sleepiness scale (KSS) was used as an indicator of subjective sleepiness [12]. As shown in Table 2, the KSS is a questionnaire that test subjects can use to evaluate their own level of sleepiness. The odd numbers in the scale include sensitivity words indicating sleepiness. To allow test subjects to record their own changes in sleepiness during the test in real time, the researchers focused on the expressions used in the odd numbers of the scale. This made the scale easier for test subjects to remember and make judgments about during the driving tests. The evaluation scale for this experiment was defined to consist of levels 1 through 5. Test subjects were instructed to press buttons installed on the steering wheel, which corresponded to the Table 2, according to changes in subjective sleepiness while driving.

Table 2. KSS scale and subjective drowsiness index used in the experiment

KSS scale		Experiment KSS Scale	
1	extremely alert	1	extremely alert
2	very alert		
3	alert	2	alert
4	rather alert		
5	neither alert nor sleepy	3	neither alert nor sleepy
6	some signs of sleepiness		
7	sleepy, but no effort to keep alert	4	sleepy, but no effort to keep alert
8	sleepy, some effort to keep alert		
9	fighting sleep	5	fighting sleep

Physiological index The R-R interval (RRI) obtainable from electrocardiogram data was acquired as a physiological index for evaluating the effect of external stimuli on the autonomic nervous system. A Silmee wearable heart rate sensor made by TDK was used so as not to interfere with driving. The sampling rate was set to 1 kHz.

Driving behavior index Drivers who possess adequate attention to follow a car traveling ahead of them in the same lane can drive safely by using visual information and other factors to appropriately evaluate the risk of collision with the vehicle ahead. Time-to-Collision (TTC) is calculated as the distance from the driver's vehicle to the vehicle ahead divided by the relative velocity of the two vehicles. This was used as an indicator of driving behavior to determine whether the driver took action to decelerate in response to deceleration of the vehicle ahead, when attention has or has not been lost. TTC is an indicator that predicts when the driver's vehicle will collide with the vehicle ahead if the current relative velocity is maintained.

Procedure

Test subjects were asked to avoid excessive work on the day of the experiment, and to refrain from consuming caffeine.

In addition, living organisms have a biological rhythm called the circadian rhythm, and it is thought that sleepiness and physiological index can easily change depending on the time the study is performed. For this reason, the tests were only conducted after lunch during the period from 1 p.m. to 4 p.m., when sleepiness tends to increase. Figure 3 shows the stimulus level course for Test 1.

Trial Run ～10 min	Rest 1 min	Test Run 30 min	Questionnaire 5 min

Figure 3. Test 1 Stimulus Level Course

After the details of the experiment and precautions to be taken were explained to each participant, they were asked to put on a heart rate monitor and get into the driving simulator. Participants practiced driving for a few minutes, rested for 1 minute after practice, then performed the test run. Each test subject performed the experiment under a total of four test stimulus levels. To avoid order effects, the order of test stimulus levels was randomized for each test subject, and they performed one test per day for a total of 4 days.

ANALYSIS METHOD

Minimum TTC

To ascertain the difference in TTC for loss of attention in each test subject at each stimulus level, researchers calculated the minimum TTC value (>0 sec) in each test run for the period from when the preceding vehicle slowed down until the test subject began to decelerate. In comparing the average values of each stimulus level, significance was first confirmed using one-way ANOVA, then Tukey was used for multiple comparison.

Period of Low Attention

It is thought that attention has been lost at Level 4 or higher of the subjective sleepiness index shown in Table 2 when a person is aware of their own sleepiness. Therefore, the cumulative time driving during the experiment at KSS Level 4 or higher was documented for each stimulus level. In comparing the average values of each stimulus level, significance was first confirmed using one-way ANOVA, then Tukey was used for multiple comparison.

Average Rate of Change of Standard Deviation of NN intervals

Sleepiness extrapolated from heart rate information is based on findings related to presumed autonomic nerve activity based on electrocardiogram frequency analysis. However, when a person is fighting sleepiness in order to perform a task, such as in this report, a conflict arises between the body's desire to sleep and the arousal effort required to stay awake and drive. The level of arousal rises and falls when arousal effort is applied, effecting each physiological index. This is said to complicate the change in level of arousal, making it challenging to estimate sleepiness [13]. The Standard Deviation of NN intervals (SDNN), which is the standard deviation of the RRI and the analytical index in the time domain, was thus used to lessen the influence of arousal effort. This value tends to fall during arousal and rise during sleepiness. The average SDNN was compared at 120-second intervals (240 sec, 360 sec, ...), using the average SDNN level 120 seconds directly after the test as a base.

EXPERIMENTAL RESULTS

The change in TTC and SDNN accompanying sleepiness was confirmed. Figure 4a shows the change in TTC and sleepiness for one test subject under Stimulus Level A. The horizontal axis shows time passed and the vertical axis shows the TTC and KSS level at that time. The KSS level rose in the latter half of the test, and the TTC more frequently dropped greatly when the car in front began slowing down. Figure 4b shows the relationship to physiological index at that time. The horizontal axis shows time elapsed, and the vertical axis shows the change in average SDNN at 120-second intervals during the test run with the SDNN at 120 seconds after the start of the test set as 100%. The rate of change was observed to increase from the start of the test run as sleepiness rose, and reached 118% by the end of the test.

Figure 5a shows the change in TTC and sleepiness for the same test subject under Stimulus Level B2. There was little change in TTC despite the elapsed time, and KSS remained at KSS Level 1 for the majority of the experiment. In addition, there was little rise in the rate of change in SDNN shown in Figure 5b as the test progressed, and the value at the end of the test was 81%, which was even lower than at the start of the test.

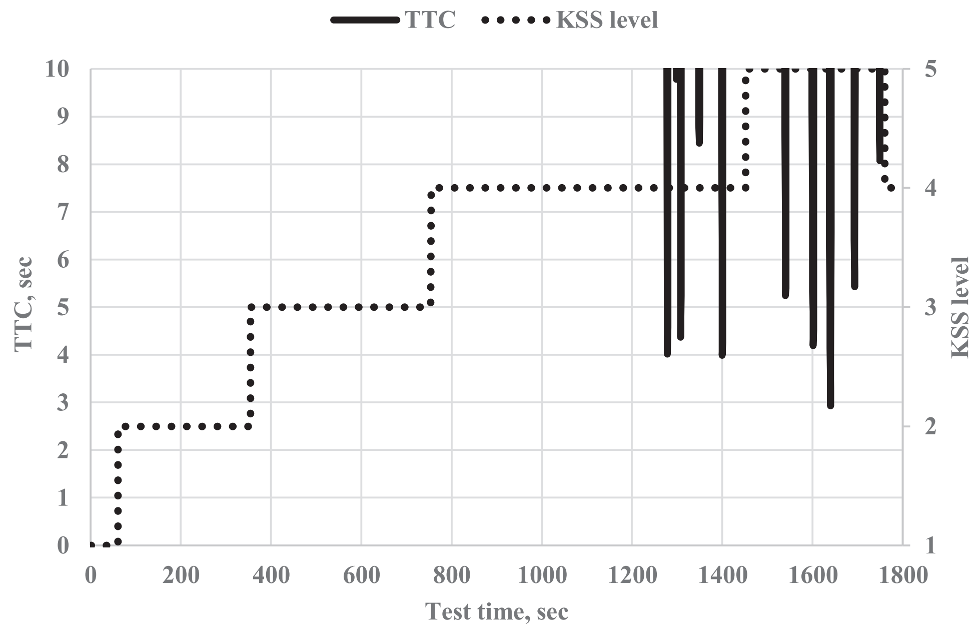


Figure 4a. Change in TTC and sleepiness (Stimulus Level A)

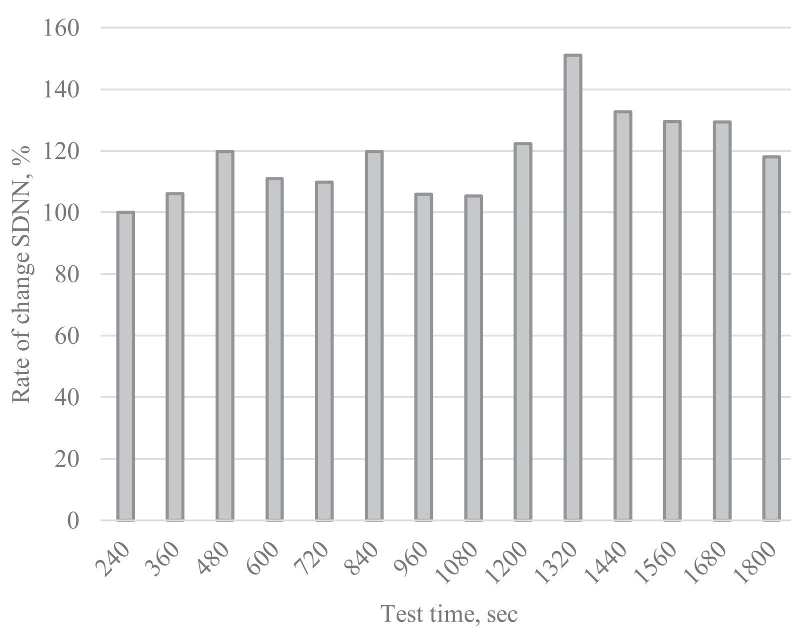


Figure 4b. Average rate of change of SDNN (Stimulus Level A)

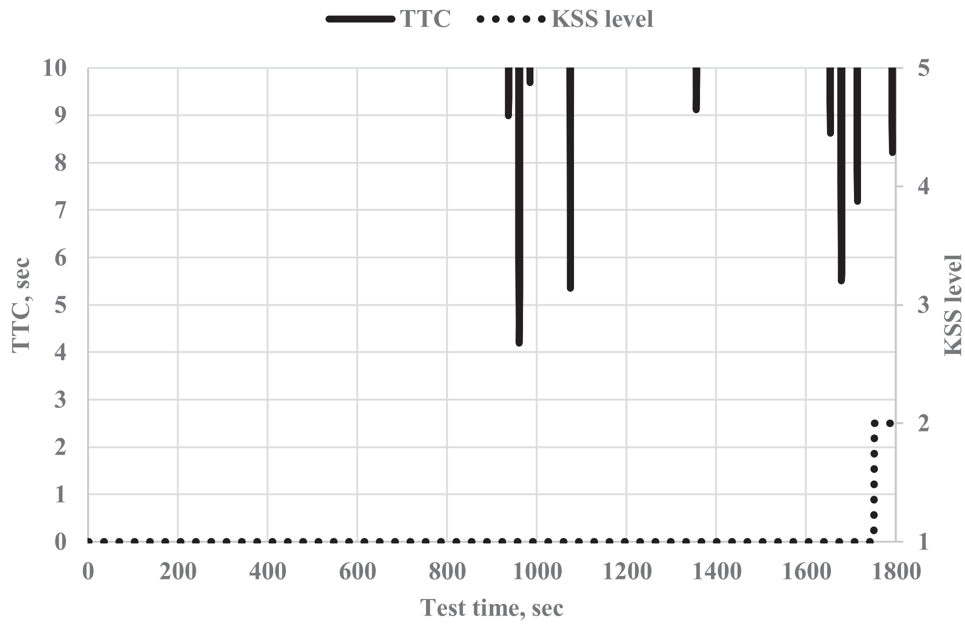


Figure 5a. Change in TTC and sleepiness (Stimulus Level B2)

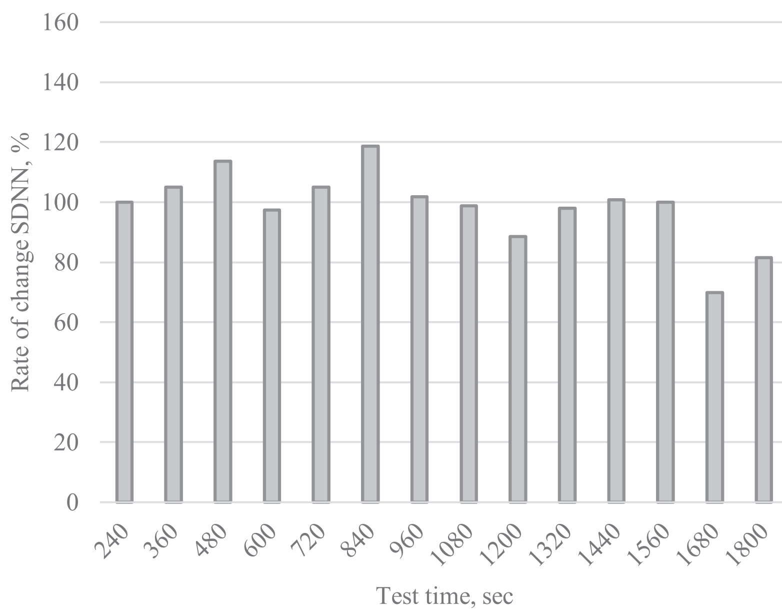


Figure 5b. Average rate of change of SDNN (Stimulus Level B2)

The influence of each stimulus condition on driving behavior was confirmed. Figure 6 shows the distribution of minimum TTC for each of the eight test subjects over all test intervals at each level. The forward collision warning (FCW) operation standard shown in the Forward collision warning system confirmation test [14] stipulates a TTC of >2.4 sec. In contrast to the TTC of <2.4 sec expected with FCW operating, around half of the test subjects drove with a TTC of 2.4 sec or less under Stimulus Level A. Under B1, one of the test subjects drove with a TTC of 2.4 sec or less. It was found that under stimulus levels B2 and B3, all of the test subjects drove with a TTC of 2.4 sec or more. When one-way ANOVA was performed with the experimental conditions as factors, the main effect between experimental conditions was found to be ANOVA F (3,32) = 8.20, p <(0.01). When Tukey multiple comparison was performed, stimulus levels B2 and B3 were found to be greatly higher than Stimulus Level A, at a level of 1%.

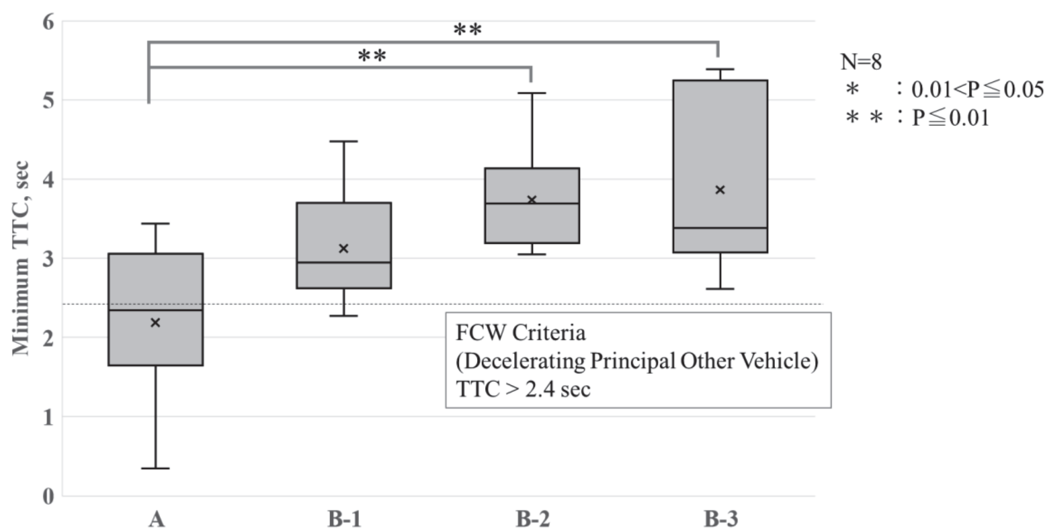


Figure 6. Minimum TTC at each stimulus level

The influence of each stimulus condition on sleepiness was confirmed. Figure 7 shows the total time at KSS Level 4 or higher for each stimulus level. While participants were subjectively aware of their own sleepiness for approximately 1,000 seconds under Stimulus Level A, it was found that applying a stimulus reduced the time that participants were aware of being sleepy by 58% for Stimulus Level B1, 86% for B2, and 77% for B3. When one-way ANOVA was performed with the experimental conditions as factors, the main effect between experimental conditions was found to be ANOVA F (3,32) = 9.65, p <(0.01). When Tukey multiple comparison was performed, Stimulus Level B1 was found to be greatly higher than Stimulus Level A at a level of 5%, while levels B2 and B3 were found to be greatly higher at a level of 1%.

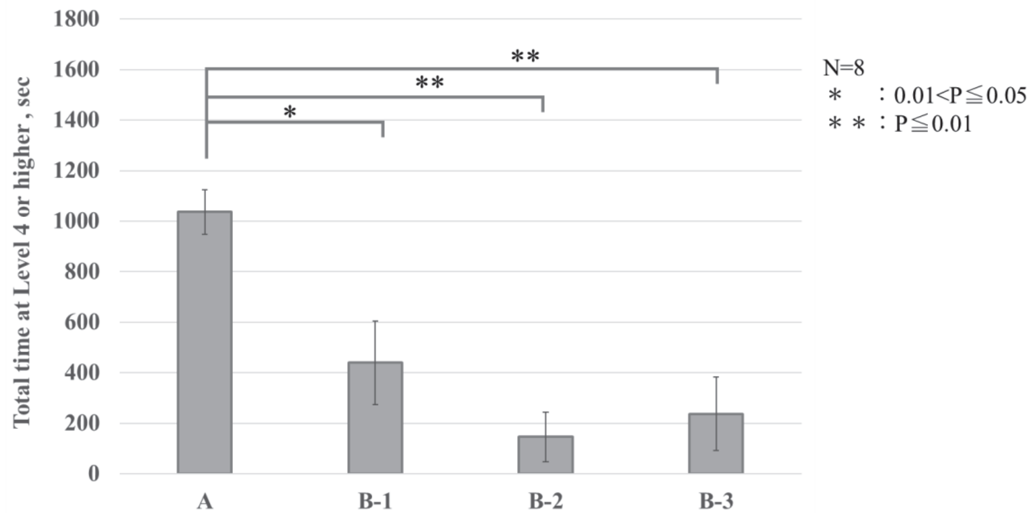


Figure 7. Time aware of sleepiness (KSS ≥4) at each stimulus level

The influence of each stimulus condition on the autonomic nervous system was confirmed. Figure 8a shows the average rate of change in SDNN from Stimulus Level A to Stimulus Level B1, and Fig. 8b shows the comparison of change in KSS. For A, the KSS level was found to increase in the latter half of the test, and the rate of change in SDNN also tended to increase. For B1 as well, a similar trend was observed in the latter half of the test, but both KSS level and the rate of change in SDNN tended to be lower than those of A. For B2, shown in Fig. 9a, it was found that the increase in KSS level was suppressed even during the latter half of the test. Also, the rate of change in SDNN shown in Fig. 9b remained almost unchanged from the start. For B3, shown in Fig. 10a, it was observed that the rise in KSS level was suppressed just as for B2, but the rate of change in SDNN rose in the middle of the test (960 to 1,440 sec), as shown in Fig. 10b.

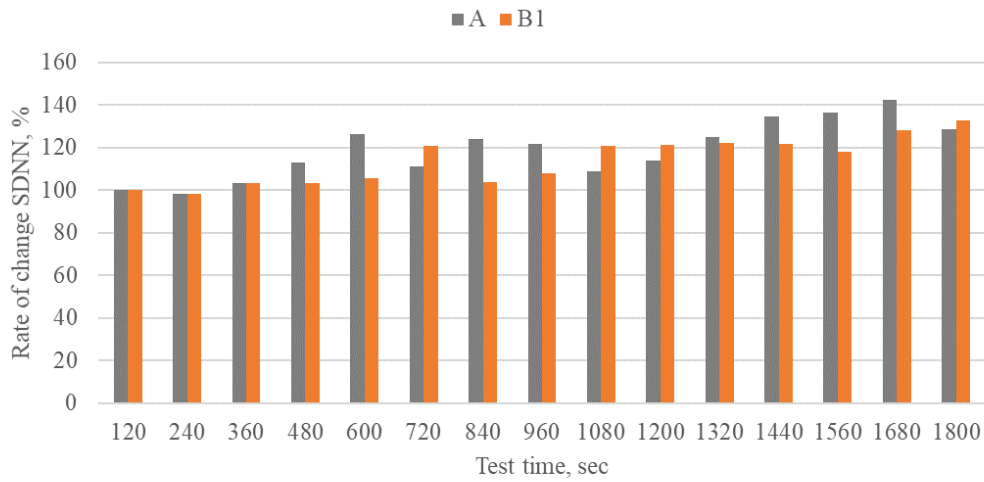


Figure 8a. Comparison of rates of change in SDNN (A vs. B1)

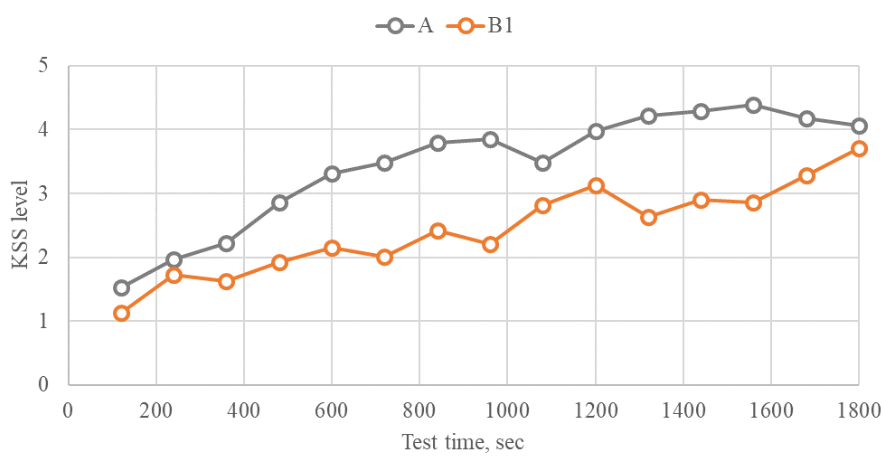


Figure 8b. Comparison of change in average KSS level at 120-sec intervals (A vs. B1)

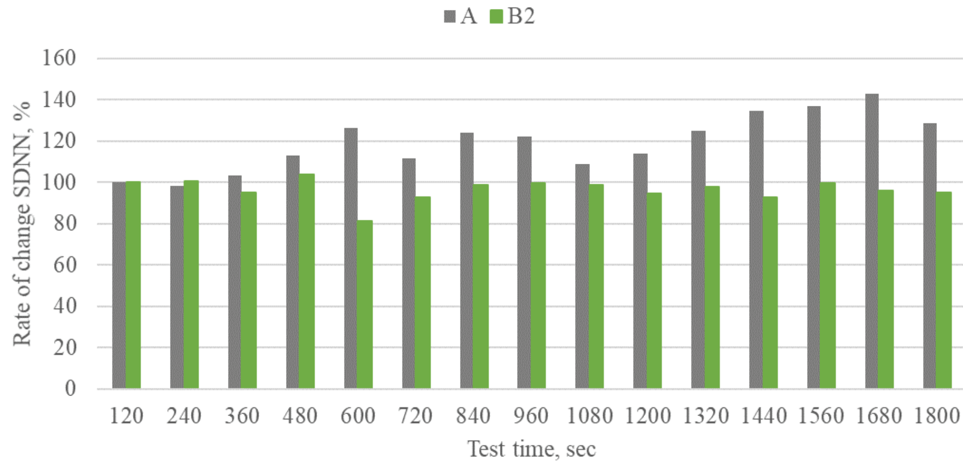


Figure 9a. Comparison of rate of change in SDNN (A vs. B2)

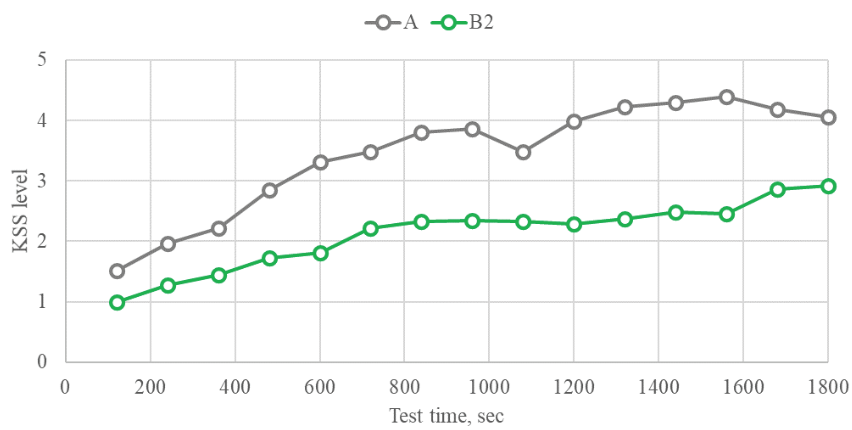


Figure 9b. Comparison of change in average KSS level at 120-sec intervals (A vs. B2)

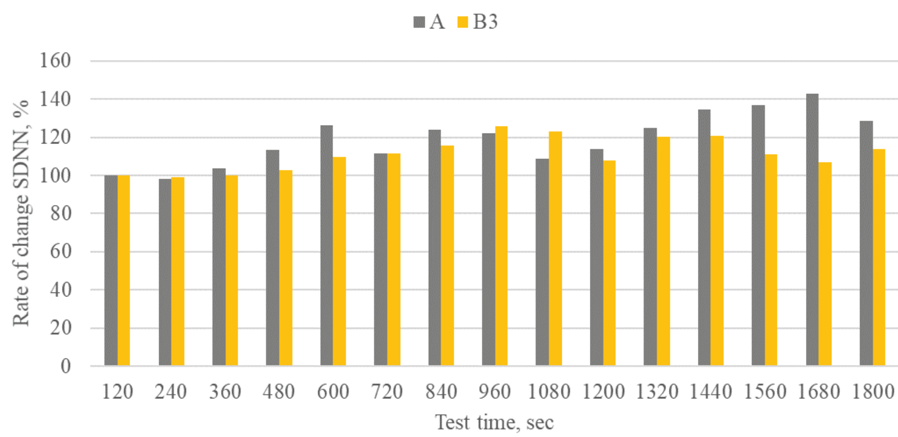


Figure 10a. Comparison of rate of change in SDNN (A vs. B3)

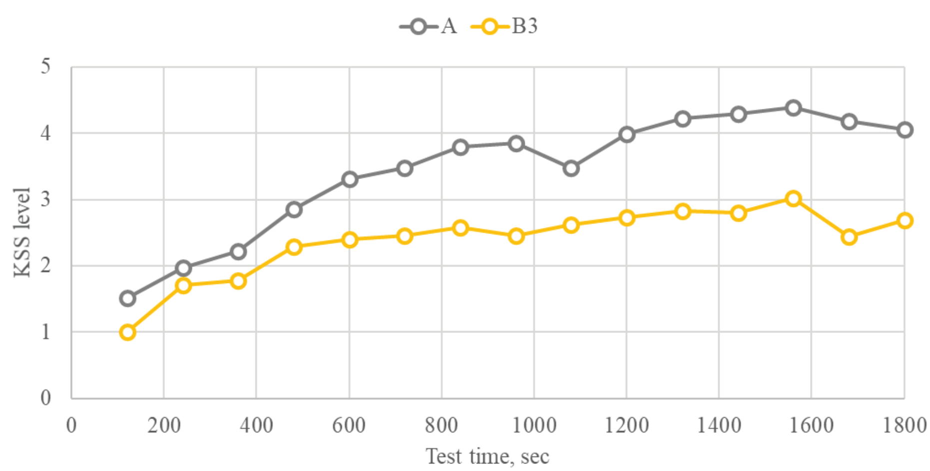


Figure 10b. Comparison of change in average KSS level at 120-sec intervals (A vs. B3)

DISCUSSION

Figure 8a indicates that the increase in the rate of change in SDNN was suppressed for B1 from 480 to 960 sec and 1440 to 1680 sec, respectively, and that can be assumed to be linked to the suppression of rise in sleepiness for B1 shown in Fig. 8b. This suggests that the minimum TTC for A shown in Fig. 6 is trending to the safe side due to the effect of the musical stimulus on the autonomic nervous system. From these results, it can be assumed that the tempo of the music presented in Stimulus Level B1 was somewhat effective in inhibiting loss of attention due to entrainment. In comparing the stimulus methods in scenario B, the average KSS level in B2 and B3, as shown in Figs. 9b and 10b, remained low compared to B1 staying at Level 3 or below. In addition, it was found that the length of time for which test subjects were aware of being sleepy also decreased compared to B1, as shown in Fig. 7. As the KSS level remained low, there was no radical drop in minimum TTC as seen in B1, and it can be assumed that B2 and B3 tended to the safe side. The stimuli in B2 and B3, which this study investigated, focused on enhancing bass and beat stimuli to affect the autonomic nervous system [10] and stimulate the sensorimotor cortex [11]. In particular, the rate of change in SDNN for B2 remained almost unchanged until the end of the test, as shown in Fig. 9a. This suggests that by considering an amplified bass stimulus as well as a musical tempo, researchers were able to inhibit acceleration of the parasympathetic nerves through the recognized effect on the autonomic nervous system, and maintain the state of attention participants had at the start of the test. On the other hand, no clear decreasing trend in the rate of change in SDNN was observed for B3 when compared to B1. One possible reason for this may be differences in how the stimulus was felt by each subject. Concerning the stimulus in B2, some test subjects did not feel any difference compared to the original sound source. As there were no negative comments in particular, it can be assumed that the intensity of the stimulus did not go beyond what was acceptable. However, after the Stimulus Level B3 test, some participants commented that while they clearly felt the vibration of the seatback when the stimulus began, they stopped noticing it after a while. It is thought that this is due to the haptic characteristic of humans known as acclimatization, in which sensitivity to stimuli decreases with time. This could be why the effectiveness of the stimulus decreased. In addition, some participants commented that the stimulus was annoying, and this may have impaired the stability of the physiological index. These differences in the acceptability of the stimulus may have played a part in the failure to inhibit acceleration of the parasympathetic nerves to as great an extent as in B2. For future stimuli, in order to keep the stimulus subliminal, its strengths and weaknesses should be considered. Also, setting personalized stimuli that do not annoy the driver can be expected to stabilize their physiological index and effectively increase acceptability.

CONCLUSION

This study sought to inhibit the loss of attention brought about by careless driving. To this end, the researchers obtained indicators consisting of TTC for driving behavior, SDNN as a physiological index, and KSS for subjective sleepiness, and evaluated the change in each indicator brought about by each stimulus method. It was found that three stimulus conditions were able to move driving behavior toward the safe side: simple musical presentation, music with amplified bass, and music with superimposed vibration. In particular, music with amplified bass was found to suppress the acceleration of the parasympathetic nerves from the standpoint of the physiological index. Also, it is thought that it may be possible to maintain attention using this stimulus. On the other hand, while it is possible to maintain attention by superimposing vibrations, it was not acceptable to some test subjects. It is possible that further improvement can be achieved by personalizing the intensity of the stimulus. In addition, it is thought that signs of sleepiness that will trigger the stimulus can be obtained by using a driver monitor camera or wearable device. The above results demonstrate the feasibility of a system that inhibits loss of attention due to careless driving through natural intervention using music.

REFERENCE

[1] ITARDA-Institute for Traffic Accident Research and Data Analysis., 2020. “MAJOR ACTIVITIES: ITARDA- JAPAN TRAFFIC ACCIDENTS DATABASES.” https://www.itarda.or.jp/english/activities

[2] NHTSA-National Highway Traffic Safety Administration., 2017. “Traffic Safety Facts – Drowsy Driving 2015.” DOT HS 812 446

[3] Lavie, P. 1986. “Ultrashort sleep–waking schedule. III. ‘Gates’and‘forbidden zones’for sleep.” Electroencephalogr. Clin Neurophysiol, 63, 414–425

[4] Phillips-Silver, J., Aktipis, C. A. & Bryant, G. A. 2010. “The ecology of entrainment: Foundations of coordinated rhythmic movement.” Music Percept. 28, 3–14

[5] Bernardi, L., Porta, C., Sleight, P. 2006. “Cardiovascular, cerebrovascular,and respiratory changes induced by different types of music in musicians and non-musicians: the importance of silence.” Heart, 92, 4 ,445-452

[6] Haas, F., Distenfeld, S., Axen, K. 1986. “Effects of perceived musical rhythm on respiratory pattern.” Journal of applied physiology, 61(3), 1185-91

[7] Etani T, Marui A, Kawase S, Keller PE. 2018. “Optimal tempo for groove: Its relation to directions of body movement and Japanese nori.” Frontiers in Psychology, 9, 462

[8] Dillman Carpentier, F. R., & Potter, R. F. 2007. “Effects of music on physiological arousal: Explorations into tempo and genre.” Media Psychology, 10, 339-363

[9] Neher, A. 1962. “A physiological explanation of unusual behavior in ceremonies involving drums.” Human

biology, 34, 2, 151-160

[10] Bowling, D. L., Graf Ancochea, P., Hove, M. J., & Fitch, W. T. 2018. “Pupillometry of Groove: Evidence for Noradrenergic Arousal in the Link Between Music and Movement.” Front Neurosci, 12, 1039-1039

[11] Lenc, T., Keller, P. E., Varlet, M. & Nozaradan, S. 2018. “Neural tracking of the musical beat is enhanced by low-frequency sounds.” Proc. Natl. Acad. Sci. U.S.A. 115, 8221–8226

[12] Åkerstedt T, Gillberg M. 1990. “Subjective and objective sleepiness in the active individual.” Int J Neurosc, 52, 29–37.

[13] Ohsuga M. 2008. “Estimation of driver's arousal states.” Journal of Society of Automotive Engineers of Japan 62 (12), 46-52

[14] NHTSA-National Highway Traffic Safety Administration., 2013. “Forward collision warning system confirmation test.” https//www.regulations.gov. Docket No. NHTSA-2006-26555-0134

BENEFITS OF TACTILE WARNING AND ALERTING OF THE DRIVER THROUGH AN ACTIVE SEAT BELT SYSTEM

Frank Laakmann
Martin Seyffert
Thomas Herpich
ZF Friedrichshafen AG
Germany

Lotte Saupp
Stefan Ladwig
RWTH Aachen University, Institute for Automotive Engineering (ika)
Germany

Mirko Kugelmeier
fka GmbH, Aachen
Germany

Mark Vollrath
Technische Universität Braunschweig, Institute of Psychology, Engineering and Traffic Psychology
Germany

Paper Number 23-0216

Abstract

Research Question/Objective: Strengthening human ability to perform the driving task in emergency situations is a key ambition of vehicle design. This work aims to further improve driver reaction time by utilizing the tactile sensory channel in multi-modal warning concepts. A specific emphasis of the work is the evaluation of unique characteristics of tactile warning through an active seat belt system in contrast to other modalities.

Methods and Data Sources: Two complementary user studies were conducted by two independent research facilities with two dynamic driving simulators. With 87 participants in total there was the aim for statistical relevance of the measurements. The setup included alternative driver warning concepts for both drivers, during manual and assisted driving, and drivers engaged in another task during conditional driving automation. The tactile warning by the active seat belt system consisted of a series of retract pulses on low force levels. The assumption is that drivers will benefit from a high exclusivity of the modality in comparison to a tactile seat, steering wheel or pedal.

Results: In a first setup (manual driving, undistracted), the replacement of the acoustic/auditory warning by a tactile warning, when combined with a visual signal, resulted in an improvement of reaction time of 250 milliseconds for brake initiation. In a second setup (AD SAE Level 3) the driver took over vehicle control 1.0 second earlier with a combination of auditory, tactile and visual warning compared to a warning without vibrotactile alerting.

Discussion: Until now, only a few studies existed aimed to evaluate a tactile warning provided by a seat belt system. The work may support, within the limitations of these studies, the initial assumption that a seat belt system providing vibrotactile stimuli to the torso – specifically chest and shoulder – shows some unique benefits. The exclusivity of the sensory channel and a low interference with other signals in the vehicle lead to high degrees of detectability, discriminability, and intelligibility.

Limitations and outlook: Is the use of a tactile warning always positive, or what are effects of training or habituation? A differentiated semantic design of such tactile stimuli, the incorporation in escalating and multi-modal warning concepts, and the combination with a holistic occupant monitoring are seen as levers for improvement and subjects of further investigations.

Conclusion: This research has found that tactile warning of the driver, through an active seat belt system, can contribute significantly to improved warning effectiveness and can help to improve the driver's ability to react in vehicles equipped with Advanced Driver Assistance Systems (ADAS) and Automated Driving Systems (ADS). Functions like take-over request (TOR) or forward collision warning (FCW) may benefit by more robust alerting of the driver as part of the emergency warning procedure.

INTRODUCTION

Why is there an investigation of tactile warning and alerting through an Active Seat Belt system?

The ability for the vehicle to communicate to the driver and other passengers is one of the most active fields of automotive developments and a key lever for current brand differentiation.

In the context of occupant safety, the design of the vehicle's human machine interface (HMI) aims to contribute to safer driving by an enablement of the driver to fulfill the dynamic driving task. Well-performing HMI systems reduce unnecessary distraction and adapt to specific situations or driving modes. In-vehicle HMI systems help to build trust with other in-vehicle systems and raise the acceptance of safety features through relevant warnings and alerts [1]. On the other hand, HMI functionalities may cause distraction from the driving task by using infotainment features or through confusion of the driver by poor warning and alerting experiences.

With higher levels of driving automation, expectations to functionalities of the HMI raise as well, because the driver still needs to be involved and needs to interact with the vehicle in most cases. The presence of visual and auditory modalities across all levels of importance and urgency is developing. ADAS (ADS up to SAE Level 2) systems *help* to control the vehicle but don't cover the whole operational range. With ADS SAE Level 2 and below, drivers have the driving responsibility and therefore need to know such system limits, need to be aware of them in every moment of assisted driving and are to be prepared to *contribute* to vehicle control or *take it over*. ADAS systems require '*<u>cooperative</u>*' vehicle operation – driver *<u>and</u>* vehicle together. For flawless vehicle operation drivers need to perfectly anticipate and perceive the (limited) actions of the vehicle function and complete it with their own *contribution* – a complex operation requiring additional 'brain power'.

An ADS SAE Level 3 (AD L3) with conditional driving automation also has system limits. However, during an active AD L3 drivers do not contribute to the vehicle control. When reaching the system limit, a TOR will be sent to the driver. After a verified take-over process, the conditional driving automation ends, and the ride will be continued by manual control. This leads to an '*<u>alternating</u>*' vehicle control: either the vehicle function *<u>or</u>* the driver. During conditional driving automation drivers do not need any 'brain power' regarding vehicle control. However, in case of the automation system reaching its limits, they need to be prepared for a takeover process. Both *cooperative* and *alternating* vehicle operations challenge drivers with processes beyond unassisted driving in close interaction between human and machine. Thus, ADS functions are to be based on an effective communication. HMIs and warnings – as a crucial part of the communication between vehicle and driver – are safety relevant building blocks in today's vehicle design. HMI functionalities of ADS are subject to technical standards or regulations, and they are addressed by safety performance protocols of consumer rating organizations.

Amongst others tactile warning is being applied for vehicle safety systems – i.e., a force feedback of brake pedals when Electronic Stability Control (ESC) is in action or a vibrating steering wheel in the case of a Lane Departure Warning (LDW). In newer applications, vibrotactile features of the seat are used to provide driver warnings as well as for LDW but also for Park Distance Control (PDC) or other ADS functions.

TERMS

> **<u>Tactile/Haptic</u>**: Here, the term *tactile* describes the investigated solutions in a clearer way than only *haptic*, which is defined as *active* exploration of objects or surfaces
>
> **<u>Automated driving systems (ADS)</u>**: For ADS or Advanced Driver Assist Systems (ADAS) SAE J3016 [2] differentiates the six levels of driving automation. ADS interface with the in-vehicle infotainment (*IVI*) or in-car entertainment (*ICE*) to perform human machine communication
>
> **<u>Alerting and Warning</u>** are essential ADS features. *Alerting* can be defined as the notification or perception of an approaching danger. According to Ayres et al. [3] *warning* can be defined as 'any information that has the potential to change behavior and prevent accidents'
>
> **<u>Warning Concept</u>**: In this paper a *warning concept* is understood as HMI solution that first alerts drivers and second transmits warning information by using appropriate human sensory channels. Thus, as an example one warning concept could be '*auditory alerting*' combined with '*visual warning information*' and will be abbreviated in this text as '*A-V*'

BASIC PRINCIPLES

This work aims to contribute to the ongoing investigations of improved performance of the warning and alerting by an in-vehicle HMI. This paper outlines investigations if and how the tactile modality can support and improve the performance (effectiveness) of driver warnings. The specific context is safety-relevant communication and the utilization of tactile warning for ADS functions and related traffic scenarios.

Vehicle warnings are generated as an initiating process step for safe vehicle operation by the driver targeting a safe state in the vehicle. The situations in which warnings are to be processed by drivers are influenced by human factors.

Situations causing an emergency warning

Warnings are required in situations where the occurred or predicted problem is of safety-relevant importance and urgency. *Importance* describes the *relevance* for driving and safety, while *urgency* describes the *need* for action under the aspect of time. Driver *warnings* are rooted in both, importance and urgency, as e.g., in a forward collision warning (*FCW*). The situation is relevant for driving and safety (importance) and needs immediate driver action (urgency).

Driver mental state of mind and the need for alerting

Drivers are expected to always be prepared for every situation. However, drivers might be in a mental state affecting their 'readiness' to appropriately react on an emergency in a traffic situation. They might be *drowsy* instead of *wide awake*, *indifferent* instead of *attentive*, or *busy with a non-driving task (distracted)* instead of *focused on driving*. An inappropriate mental state may cause a dangerous traffic situation needing a warning, but also needs an *appropriate alerting* designed to *arouse* the driver.

What is emergency warning? A closer look to the process of warning and alerting

A detailed look to the warning and alerting procedure is shown by *Figure 1* and the discussion of three elementary steps is important for the design of an efficient emergency warning:

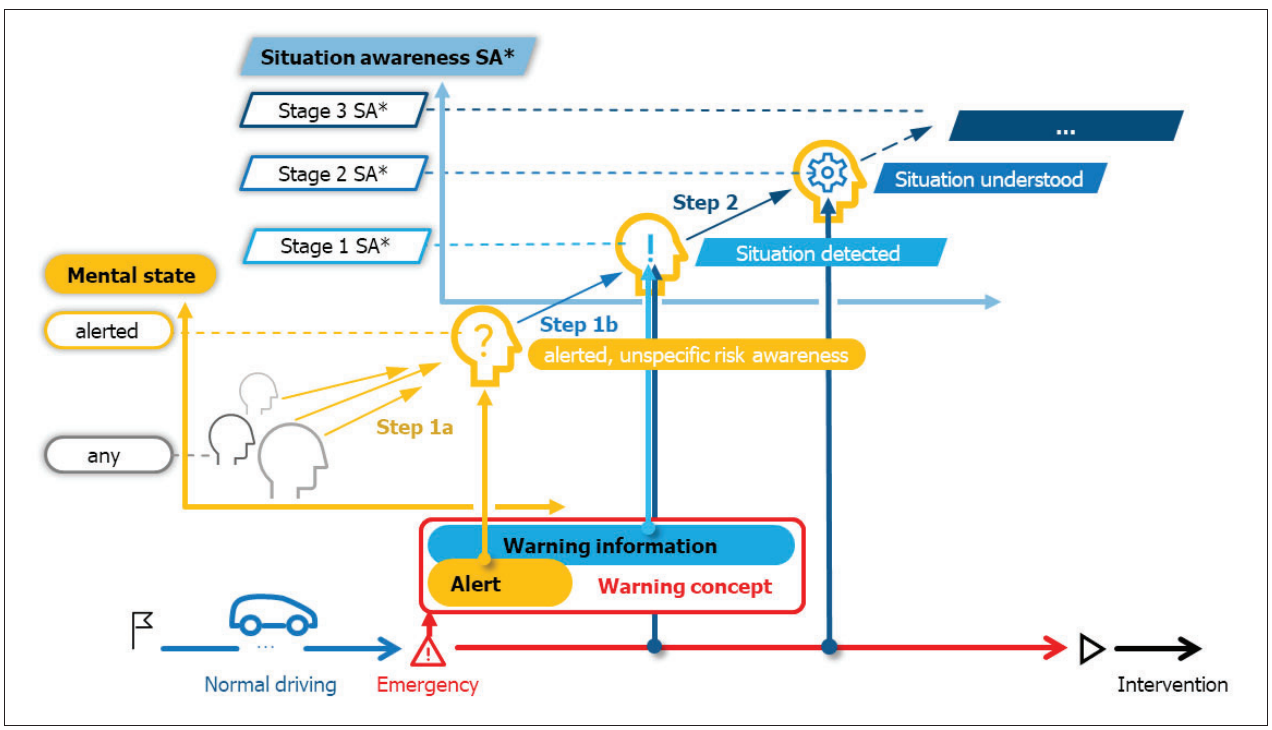


Figure 1 Detailed steps of the warning and alerting process prior to intervention
** with reference to Endsley's situation awareness (SA) concept* ***[4]***

Alerted/alarmed: Since drivers might be in mental states affecting cognition processes an alerting process is required as a first step (*Figure 1*, *'Step 1a'*) to ensure the following:

- an *aroused state* (awake, attentive, and focused) to enable processing of new information,
- an *unspecific risk awareness*, i.e., the perception that there is *any* safety-relevant situation needing his immediate intervention. With this perception the driver is *alerted/alarmed* – still not knowing the details of the emergency.

Situation Detected: In a following step (*Figure 1*, *'Step 1b'*) the alerted driver must detect the *specific threatening character* of the situation. For example, the driver detects that the vehicle ahead brakes suddenly.

Endsley [4] proposed a three-stage model of situation awareness (SA) with a 'Stage 1 SA' to *perceive* the situation, a 'Stage 2 SA' to *comprehend* and a 'Stage 3 SA' to *project* the situation into the future. The 'Steps 1a' and '1b' described in *Figure 1* are a closer look into Endsley's 'Stage 1 SA' resulting in the situation detection.

Situation understood: In 'Stage 2 SA' of Endsley's three-stage model the driver *comprehends* the significance and meaning of warning and situation. For example, the drivers understand that they need to brake immediately to avoid a collision (*Figure 1*, '*Step 2*').

The *stimulus* for 'Step1a' in *Figure 1* to create an *unspecific risk awareness* is labeled as an '*alert*' or '*alarm*' in this paper, depending on the level of *'criticality'* in the current situation. This stimulus does not need complex informational content. Its only purpose is to arouse the driver and to create an unspecific risk awareness. Therefore, main requirements are to be salient, short, and simple. An alarm tone in the event of an emergency is a good example. It shall be *salient* to penetrate the consciousness of a person and to be *noticed*. It also should be *crisp* and *simple* to be *quick*.

Creating *unspecific awareness* is not only a preparational step for *warnings* but used for other information transfers as well. Mobile devices, for example, create a notification sound accompanied with an incoming message. Some vehicles generate a soft gong sound, this causes the driver to look up more information on the instrument cluster or central display, e.g., in this case the fuel tank is at the reserve limit. Hence, a preparational signal to give notice of available information is an established *HMI concept* not only used in vehicles and not only for warnings.

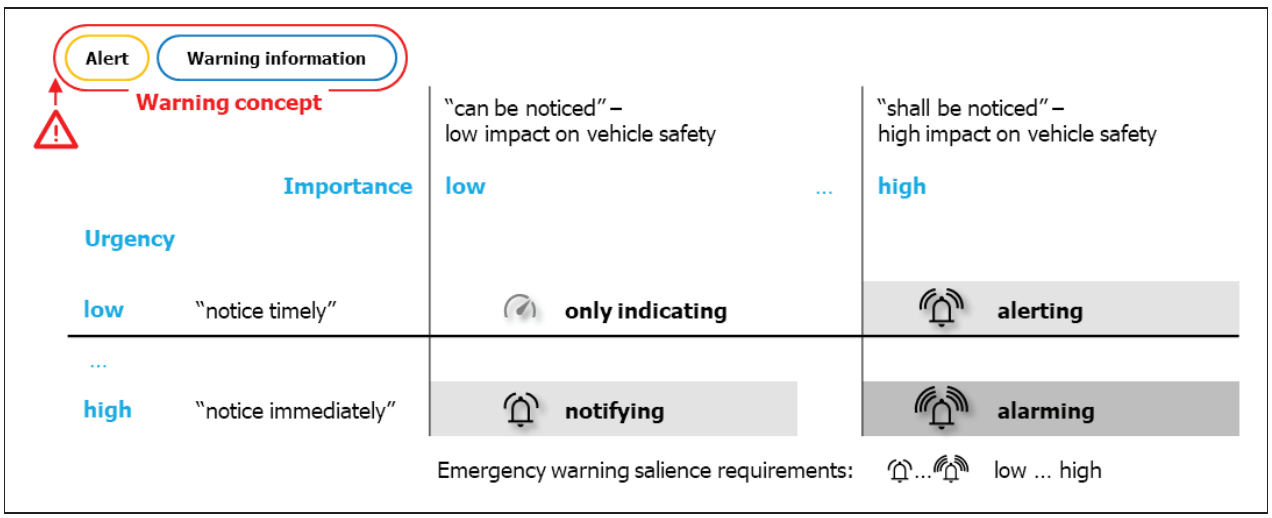


Figure 2 Criticality of an emergency warning and salience requirements of alarming stimulus

Therefore, when applying this concept in vehicles for *warnings* the related alerts/alarms must differentiate from notifications with less critical information. The stimulus for alerting or alarming must be *salient* but does not need to be pleasant or comfortable. Furthermore, since there are situations of different *criticalities*, it might be desirable to differentiate between lower and higher criticalities.

Figure 2 shows a proposal for alerting stimuli with a salience adapted to the criticality of the situation, as a preparational step to enable an ensured and efficient processing of warning information.

The criticality is driven by importance and urgency of a situation while '*importance*' describes the *relevance* for driving and safety and '*urgency*' describes the need for action under the aspect of time. Both, importance and urgency, might be low or high, by this creating a two-by-two matrix with different criticality levels as shown in *Figure 2*. If a situation is of low importance and low urgency a notification is not needed. In the case of a personal or general information as e.g., the exterior temperature '*notification*' – as a pleasant form of an alert – might be appropriate. In driving and safety relevant situations a salient alert would be justified and in case of a life-threatening situation an alarm is the strongest and most effective form of an alert that is required.

Aroused drivers are *prepared* for the following *cognitive* 'Step 1b' to *detect* the dangerous situation. As indicated in *Figure 1* there are two options to detect details of the emergency:

- by directly looking into the related direction (*Figure 1, red arrow*),
- or first, spotting the warning information (*Figure 1, blue arrow*) and then looking into the direction of the related critical situation (*Figure 1, red arrow*)

The first way – also described as 'bottom up' processing – may be faster since there are no additional cognitive processes being involved but requires that the driver immediately looks in the direction of the problem. The second way – 'top down' – needs processing time to redirect the gaze to the instrument cluster, detect the related warning information, understand the meaning of the symbol, redirect the gaze to the related dangerous situation and then detect it.

Which processing path is being used, strongly depends on the nature of the alerting stimuli and warning information received. Alerts of high salience are likely to trigger bottom-up processes whereas information enriched with expectancies and value are likely to trigger top-down processes [5].

In case of an *unspecific alerting stimulus* the driver is expected to direct his gaze in the driving direction allowing for the detection of critical situations as e.g., imminent forward collision or lane departure situations. In cases of *specific* emergencies *out of* the normal viewing area. E.g., a vehicle is in the blind spot, and it might be beneficial to alert the driver with a *specific* stimulus guiding his attention/gaze into the relevant direction. The flashing light of a Blind Spot Detection (*BSD*) system in the exterior mirror draws the driver's gaze to the side facilitating the driver to notice the vehicle nearby. An *alerting stimulus* works *intuitively* when the driver's attention is guided to the related critical situation and the driver can detect it. In this case additional *warning information* might not be necessary and the speed of situation detection is high.

However, drivers need further assistance in case they don't recognize the threatening situation immediately. Therefore, *specific threat warning information* helps show what might happen, e.g., a forward collision – but hopefully is not needed. Therefore, for fast and effective situation detection, the warning design should *explain* the character of the critical situation, e.g., an imminent forward collision. Warning information – usually visual symbols – should be clear and simple for rapid understanding.

However, alarms are limited to unique events with both, highest importance, and highest urgency, and are assumed to be very seldom. Therefore, training effects cannot be presupposed for alarms by this emphasizing the requirement for being intuitive.

Warning design requires sensory channel selection for both, alerting and informing, to create an effective (multi-) modal concept. The visual, auditory, and tactile modalities or any combination of them are eligible options. But to better understand the potentials of these options it needs a closer look into the nature of the three sensory channels.

MODALITIES FOR WARNING AND ALERTING

What are distinguishing characteristics of the relevant modalities?

Warning systems in today's vehicles mostly address those human modalities known for their high resolution and relevance for the driving task, hence, the visual and auditory modality. *Figure 3* outlines a qualitative assessment of the three modalities visual, auditory, and tactile in the context of this work.

The **visual sense** is switchable 'on/off' by the eyes, perceives the half space in gaze direction while recognition of detailed structures with complex informational content is restricted to a small area around the visual center. The visual sense has its limits to *alert/alarm* drivers, since drivers might not look in the direction of the alerting/alarming signal or have their eyes closed for a moment. However, *warning information* can be relayed very well, and contents can be displayed simultaneously. A reading process of several content elements usually follows a sequential order – influenced by their salience values.

Furthermore, the visual sense is known for its dominance in the field of sensorimotor transformations in tool use. Driving a car meets the definition of a sensorimotor transformation as driving is being guided by the information perceived from the outside. In other words, information sensed by the different human modalities (*visual*, *auditory*, *haptic/tactile/kinesthetic*, *vestibular*, etc.) are being cognitively processed and transformed into an event file with certain features [6] which is relevant for the motor action to be performed.

In sensorimotor transformations, the visual modality is known for its dominance in terms of guiding motor actions which is supported by various experiments in the field of tool use [7], [8].

In contrast, the **auditory sense** is always 'switched on'. It is not directed and can receive either simple or complex information sequentially. It is also highly involved in the driving task and therefore this modality is permanently exposed to a broad variety of stimuli. While visual messages still can be looked up a moment later, vocal warning information is transient and is lost when not captured instantaneously.

While driving a car, the driver is confronted with stimuli originating from infotainment, further passengers, environment and from ADS – all of them competing in the prevailing noisy situation. These stimuli need to be either filtered out or further processed and selected by the human processing system regarding their relevance, which in turn underlies a certain limitation of capacities [9]. The merely cognitive process is highly demanding and exhausting as it has the potential to bind a high level of cognitive resources, especially when the message is not simple (e.g., '1 bit') but complex, of lower salience or takes longer (word or sentence).

Therefore, the auditory sense works for *alerts* with *simple* and *salient* stimuli and in fact is used frequently for this purpose. Complex information can be transmitted as well but competes with noisy backgrounds, needs cognitive resources, a certain time for transmission and is to be understood immediately during transmission.

From the nature of sound and auditory sense it follows that acoustic information will be received not only by the driver but all other occupants in the interior as well. This could be a negative attribute for user acceptance and might lead to switched off ADS systems.

Tactile sensing by mechanoreceptors is distributed across the entire human body. It requires physical contact, the tactile sense is always 'switched on', its eligibility for transmission of complex information is very poor and the signals are transient (to be captured when sent, otherwise lost). Tactile signals can be transmitted in a discrete way without disturbing other passengers.

	Indicative characteristics of modalities for automotive applications					
	Direction of signal *Transmission characteristics of signal*	**Transience of signal** *Availability of signal over time*	**Effort for reception** *Recipient senses with or without additional effort*	**Suitable for alarming** *Ability to raise or redirect attention of recipient*	**Message content** *Ability to transmit low or also rich information*	***Warning concept*** *Main purpose for an emergency warning concept*
Modality						
visual	unidirectional	depends	with effort (gaze)	lower	low ... complex	*Warning information*
auditory	omnidirectional	only during transmission	no additional effort	higher	low ... complex	*Alerting*
tactile	dedicated body parts	only during transmission	no additional effort	higher	low	

Figure 3 Modalities of automotive applications and characteristics

What makes Active Seat Belts unique compared to other means of tactile warning?

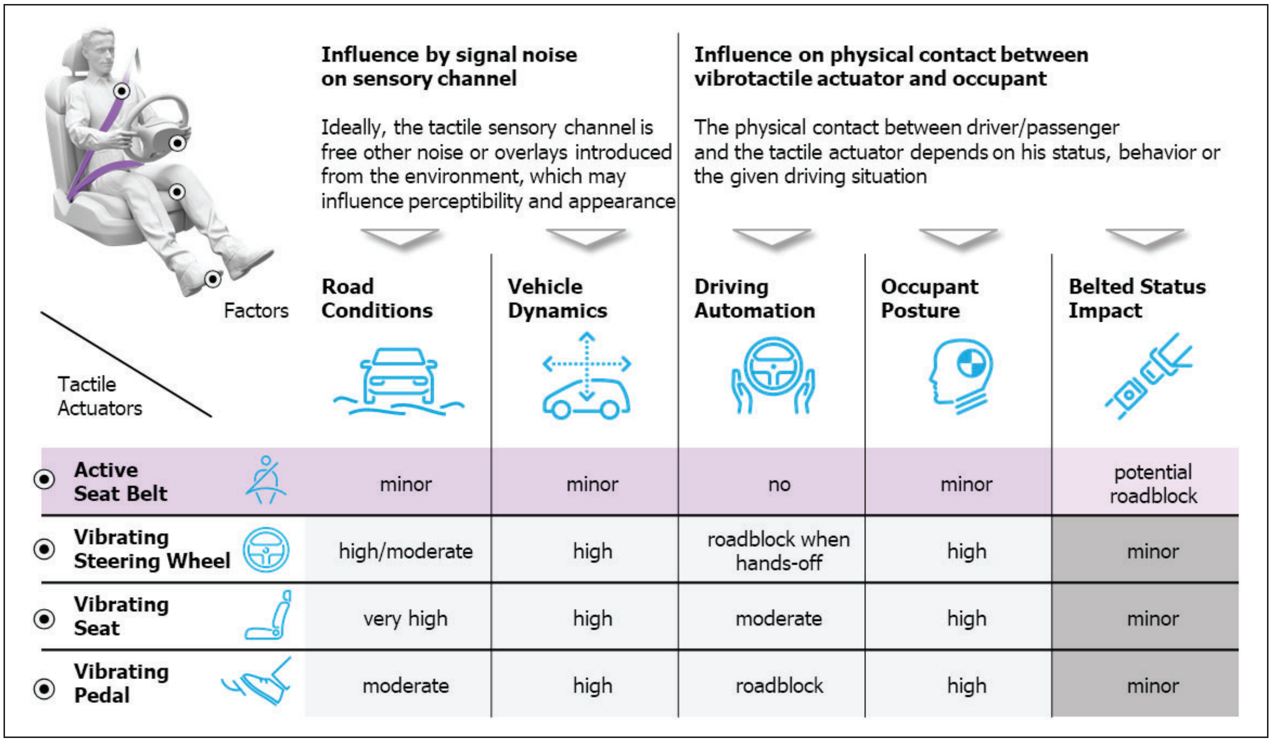
The conclusion from above is that tactile stimuli are effective to arouse driver's attention in case of critical situations. This work aims to investigate, how tactile warnings provided by an Active Seat Belt system can help raise the performance of driver warning in safety-critical situations.

Figure 4 indicates the key differences between the Active Seat Belt and other tactile options like seat-integrated and steering wheel-integrated tactile actuators, and vibrating pedals.

As the impact of those factors is minor on the seat belt there is a potential roadblock due to missing norms for the fastening of seat belts. Still, certain regions do not mandate the use of seat belts and therefore seat belts cannot be utilized for tactile warning.

Wickens et.al [5] outlined factors that determine the *noticing probability* of visible events. Applying the findings to the tactile modality it can be derived that a high noticing probability can be achieved when first a stimulus is tangible by the mechanoreceptors e.g., by contact and second the tactile stimulus shows high *salience*, meaning the ability for easy discrimination from other competing signals, or 'background' noise, coming from the environment.

Seat, seat belt, steering wheel and pedals are potential interfaces to the driver, see *Figure 6*. In vehicles with ADS systems and higher levels of driving automation the physical contact between the driver and the pedals or the steering wheel is not present. Therefore, warnings related to safety-critical emergencies or ADS functions provided by these actuators would not be recognizable in certain situations. Also, an increased variation of an occupant's posture may impact the steady contact to these tactile actuators.



Factors / Tactile Actuators	Road Conditions	Vehicle Dynamics	Driving Automation	Occupant Posture	Belted Status Impact
Active Seat Belt	minor	minor	no	minor	potential roadblock
Vibrating Steering Wheel	high/moderate	high	roadblock when hands-off	high	minor
Vibrating Seat	very high	high	moderate	high	minor
Vibrating Pedal	moderate	high	roadblock	high	minor

Figure 4 Assessment of alternative tactile actuators within the automotive interior

As discussed above, the second condition for a high noticing probability is high salience. The signal must be recognized and separated from this superimposed background. Seat surface and backrest are interfacing with the human body and take the largest share of the body weight. The interaction of body weight, seat surface, backrest and human body creates local compressive forces constantly stimulating certain areas of the body. Driving accelerations and vibrations emerging from the road surface are transmitted to the mechanoreceptors of those areas. They build a tactile 'background' noise when an additional tactile signal for warning and alerting is added to that interface.

In contrast, seat belt systems are designed to provide the highest possible wearing comfort and minimal perceptibility. A transmission of forces and vibrations from the vehicle chassis to the driver is hardly possible since there is just a slight contact between driver/occupant and the flexible seat belt. Without tactile noise background this channel transmits the tactile signals exclusively. Hence, a high noticing probability and fast perception is expected making the seat belt-based tactile modality unique.

As discussed above, driver warning at first needs a 'Step 1a' initiated by an alert/alarm generating an (unspecific) risk awareness. For this the driver is to be aroused, mental state to be raised to an *awake* and *attentive* state with a focus on driving. Both the auditory and the tactile modality fulfill the requirements for a stimulus to alert/alarm drivers and can be understood as *alternatives* for alerting and alarming while the visual modality fits best for providing more detailed information.

Finally, it would be desirable to generate the tactile stimulus with components already existing in the car. Active Seat Belt systems for reversible seat belt pretensioning are widely used in today's vehicles and allow the generation of tactile stimuli simply by adding the function by software change.

	Receivability Tactile actuator and driver body (receptor) have physical contact	**Noticeability** Noticeability of alert through the tactile stimulus ('Step 1a', see *Figure 1*) estimated for ADS functions Level ≥ 1
Active Seat Belt	given across all ADS modes (if belted)	++
Vibrating Steering Wheel	not given in hands-off situations (if lateral controlled by ADS)	-
Vibrating Seat	ensured, however variations due to passenger posture	+
Vibrating Pedal	only given if feet have contact with pedals (manual driving only)	-

Figure 5 Utilization of alternative tactile actuators for warning concepts of ADS functions

How is tactile warning adding to the other Active Seat Belt use cases?

Since its first application two decades ago, Active Seat Belts are widely used across automotive brands, vehicle segments and regions. Usually, seat belts are part of the Passive Safety domain. The initial purpose of Active Seat Belts was improved effectiveness of the occupant safety system by adding proactive seat belt pretensioning in emergency situations. Today, Active Seat Belts enable use cases of further domains like ADS/ADAS, HMI and passenger ergonomics/comfort for a superior user experience.

Active Seat Belt key functions

Active Seat Belts are equipped with an additional mechatronic drive unit to a apply a broad range of force to the seat belt. In contrast, standard seat belt pretensioners are equipped with a pyrotechnic pretensioning unit that is only intended to operate once during single crash event. Active Seat Belts serve passengers and the driver along multiple driving phases:

- Driver Warning and Notification – A *TOR* or an emergency warning of an ADS/ADAS (e.g., *FCW*) can trigger a vibrotactile pulse by Active Seat Belts being part of a multimodal warning or notification scenario
- ADS/ADAS Maneuver – Interventions in longitudinal (e.g., Autonomous Emergency Braking, *AEB*) and lateral (e.g., Autonomous Emergency Steering, *AES*) directions may impact the occupant position. Active Seat Belts may help to stabilize the occupant towards the nominal position
- Passenger Ergonomics – For buckling up and during normal driving, passengers experience a high freedom of movement as Active Seat Belts can enable a reduced seat belt retraction force
- Driving Experience – If the driver and the passengers intend to experience a dynamic and agile driving experience Active Seat Belts help stabilize the torso and mimic the experience of a 4-point belt system
- Belt Slack Removal – Shortly after the start of driving Active Seat Belts aim for a tight fit to improve the effectiveness of the seat belts in the event of a crash. It also contributes to a habit for buckling up and positive experience of the vehicle's safety systems
- Emergency prior to a potential crash event – Occupant in-position: Reversible pre-pretensioning of the seat belt in severe driving conditions to ensure that the driver is in a nominal position using strong, reversible pre-pretensioning of the seat belt prior to a likely crash event

Design aspects of vibrotactile warning with Active Seat Belts

The extended functionality makes Active Seat Belts relevant for the ADAS/ADS warning system and the vehicle HMI/UI. A salient tactile stimulus by an Active Seat Belt system can be created by an initial explicit pre-tensioning followed by a phase of consecutive pulls or vibration (see *Figure 6*). The vibrotactile pulse to the driver is customizable for different types of notifications, confirmations of operations or warnings. This allows variations depending on type of notification or warning, priority, occupant/ driver state, responsiveness of occupant/driver, context and driving situation. Frequency, force level, amplitude and duration are adjustment parameters. Cascading warning concepts or escalation strategies with a higher level of responsiveness are possible.

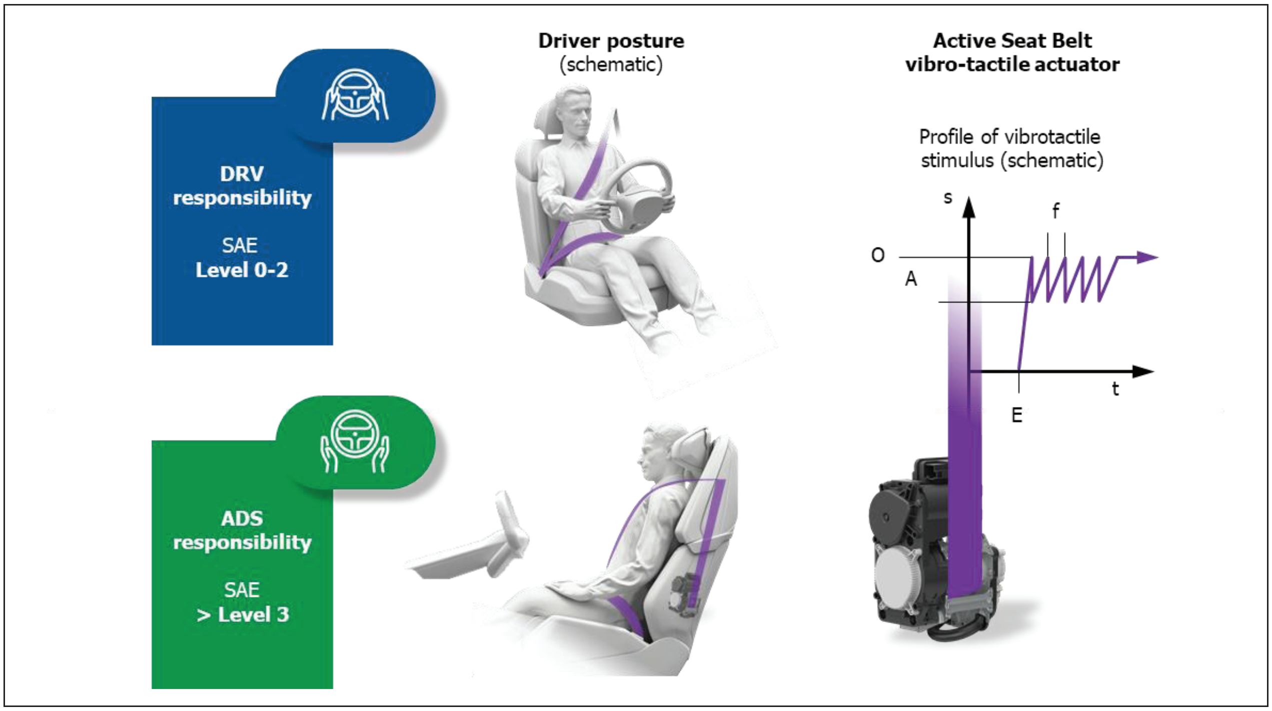


Figure 6 Vibro-tactile warning provided by an Active Seat Belt system

Application of vibrotactile warning for ADS/ADAS functions

Today, the Euro NCAP protocol considers tactile/haptic warning as an option for a multimodal warning in case of driver distraction (short/long distraction, phone usage) and driver fatigue (drowsiness, microsleep, sleep) [10]. The ADAS Performance Testing Program of U.S. NCAP by NHTSA considers tactile/haptic warning also as an option for multimodal warning concepts for different ADAS functions, e.g., Lane Departure Warning (*LDW*). The current protocol mentions tactile/haptic alerts through steering wheel or seat vibrations to alert the driver.

Previous research

Research on the impact of alternative bi- or multi-modal cues and the influence workload shows that tactile stimuli result in a more effective warning compared to the visual and/or auditory cues, especially when a certain workload must be processed [11], [12]. In general, multi-modal redundant warning concepts show faster reaction times in both situations – easy and complex – compared to a warning concept with a single cue [13], [14], [15].

Further research on automotive use cases shows that the tactile stimulus can facilitate the response to a visual information when monitoring the traffic, especially when tactile information as a spatial (location-specific) warning signal can be synchronized with the visual information [16], [17], [18].

A relevant benefit of the tactile cue for warning concepts is the robustness and noticeability under various conditions. It seems to be more difficult to miss a tactile stimulus compared to visual or auditory information [19].

The application of tactile stimuli in Automotive applications for ADS functions had been assessed for different tactile actuators like steering wheel [20], brake and accelerator pedals [21], [22], or seat [23], [24], [25], [26]. See also the survey of Petermeijer et.al [27] on driver support systems with tactile warning concepts.

The use of the seat belt system to transmit vibrotactile stimuli to the driver had been investigated already in the past [28], [29]. Alternative solutions for the tactile actuator had been investigated for certain ADS functions, e.g., tactile seat belt vs. vibrating steering wheel [30]. However, it can be assumed that the detailed technical solution for the vibrotactile stimulus is different to the Active Seat Belt with vibrotactile feature used for the studies of this work.

RESEARCH QUESTIONS AND METHODOLOGY

The nature of the tactile modality, as part of a warning concept by an ADS system and especially the use of Active Seat Belt systems as the tactile actuator, have not yet been sufficiently investigated and therefore are in the scope of these investigations.

The tactile modality is considered as an option for multi-modal warning concepts of ADS in case of emergencies. Multiple combinations of modalities and specific configurations of stimuli – also along an escalation path – are possible. The performance of tactile warnings should be investigated across different driving automation levels in a controlled environment and across a significant number of users. This leads to the following research questions:

Research question 1

What, if any, are the advantages of a tactile stimulus for warning concepts of ADS systems?

For safety-relevant emergencies, the most relevant parameter is *reaction time* of the user. In the focus of this work are warning concepts related to ADS functions for different levels of driving automation. In a first test setup, the driver's reaction time should be measured after the occurrence of an emergency. Reaction time is the time from the beginning of an event (i.e., cut-in vehicle begins to leave its own lane) to the measurable reaction of the driver (i.e., time until brake is pressed). In a second test setup the reaction time of a driver will be measured in a TOR during conditional driving automation of an ADS. See below for a more detailed description of the selected test scenarios.

Research question 2

Which conclusions can be derived from user studies regarding further aspects of efficient warnings?

To better understand the impact on warning concepts by an ADS system, further data was collected during the series of tests with altered warning concept modality combinations.

Decision quality was evaluated when comparing the multi-modal warning concepts. In the first user study the driver could choose between braking and evasive steering. In the second user study driving parameters like speed or lateral offset were measured to compare the impact of the specific warning concept on drivers' perception of the situation or reaction.

Other evaluation aspects are related to the specific alerting or information function within a warning concept. How does a tactile stimulus compare to the auditory stimulus to arouse the driver within the first part of the warning process (impact initial mental state of mind)? What is the influence of the driver's mental load during the emergency warning? How robust are the different warning concepts in different types of emergencies or different mental states of mind?

Research question 3

How do the users respond to tactile stimulus being part of ADS warning concepts?

Today, drivers are familiar with visual and auditory modalities applied for warning concepts and may have experienced a few applications of tactile/haptic modalities during emergency warnings. However, the seat belt is currently not known for tactile warning capabilities by consumers. Therefore, personal user feedback was collected through structured interviews during and after the simulated test drives.

Research question 4

How does an Active Seat Belt system with tactile features compare to other tactile actuator options?

The user interface increasingly needs to respond to diverging conditions and requirements of ADS driving automation levels. As a tactile actuator typically requires physical contact to the human body, options are limited. The findings of the user studies should be utilized to discuss tactile stimuli provided by an Active Seat Belt system with other options of tactile actuators.

Design of test series

It is the ambition of this work to assess the impact of vibrotactile stimulus as part of multi-modal warning concepts across a broad range of driving conditions, and different conditions of the drivers during driving with or without the presence of driving automation.

Therefore, the design of the user studies and the selection of the specific test series configurations should cover both conditions and driving modes (see *Figure 7*):

- first: the driver is responsible for the driving task: SAE Level 0 to 2
- second: ADS is responsible for the dynamic driving task: especially SAE Level 3

With the driver being responsible for the driving task and the driver being in an alert (attentive) mental state of mind, the most challenging environment for the comparison of alternative multi-modal warning concepts was selected. This condition can be considered as the nominal state for SAE Level 0 to 2.

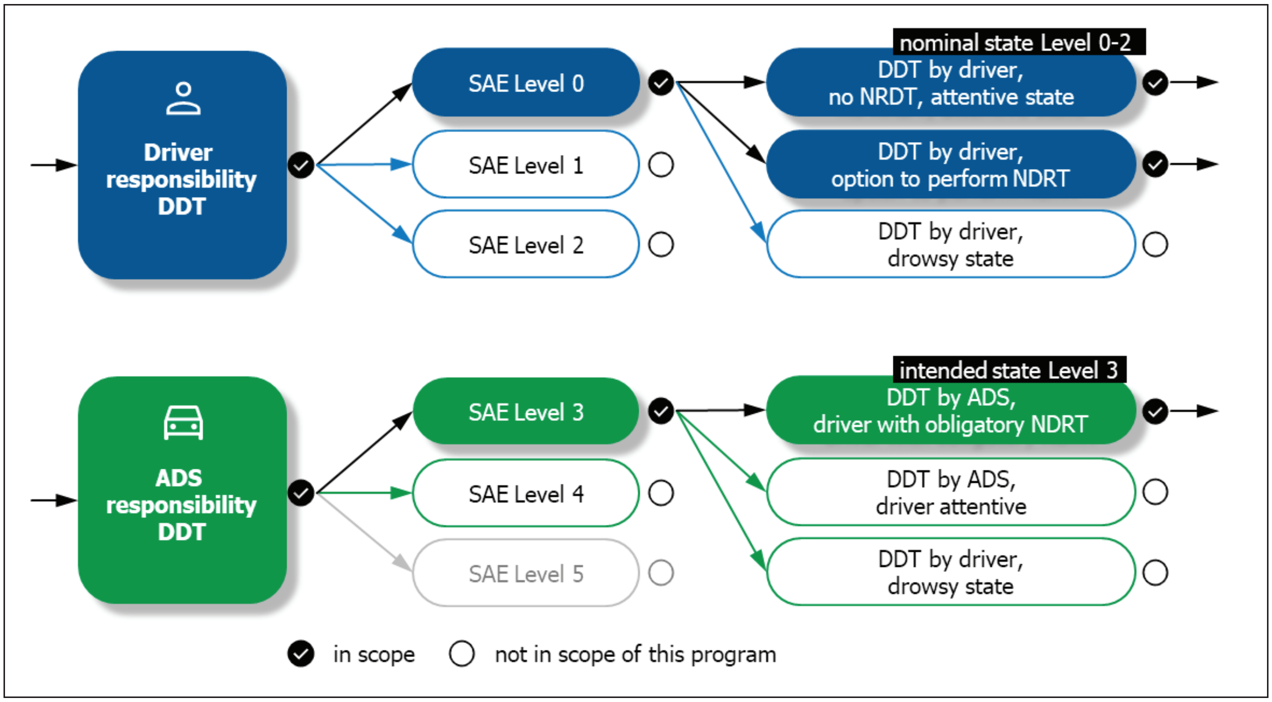


Figure 7 Proband study program

Due to the nature of vibrotactile alerting, compared to an auditory cue, the impact on the performance of the warning concepts might be even higher in situations with a driver being not attentive or distracted especially during partial driving automation (SAE Level 2) with longitudinal and lateral vehicle motion control.

For driving scenarios with responsibility of the dynamic driving task transferred to the vehicle's ADS, an emergency resulting in a take-over request was selected to compare alternative multi-modal warning concepts. Here, the driver will be occupied with non-driving related tasks as the intended use case during conditional driving automation. In real driving situations, the driver might be listening to music, watching videos, or having conversations with other passengers. Otherwise, drivers might require a recovery phase or are in a less attentive state. Haptic signals may be advantageous in these situations for alerting the driver compared to other cues.

The two studies were designed following alternative approaches. In the first study, the sample size was divided in three groups so that one participant received one type of warning concept. Here, the aim was to avoid potential training effects due to repeated test situations. The second study followed a repeated measurement design. All participants received all three warning concepts with the different configurations of alerting stimuli and warning information. Here, the aim was to assess the statistical significance of the different variation parameters (warning concept, emergencies, sequence/learning effect) across the entire sample size.

USER STUDY TACTILE WARNING DURING DRIVING WITHOUT DRIVING AUTOMATION

A first user study on benefits of tactile warning through an Active Seat Belt system had been carried out by RWTH Aachen University, Institute for Automotive Engineering (ika) and fka GmbH, Aachen. The driving simulator test series, comparing three variations of warnings, were conducted in the ika's high fidelity driving simulator. Here, test scenarios focused on emergencies during a driving scenario without driving automation (SAE Level 0).

Method and procedure

The dynamic driving simulator at ika is utilized for the assessment of driver assistance systems, future steering and control concepts, and the analysis of driver behaviour (see *Figure 8*). The dome's diameter is 7.0 m and the field of view is 360° x 45°. The driving scenarios were created using the driving simulation software Virtual Test Drive by Vires. A mock-up of a BMW i3 served as the test vehicle.

The user study was carried out over six days of testing during the period between November 23 and 30, 2019. Participants started off by completing a socio-demographic questionnaire and then were introduced to the driving simulator. In a short familiarization drive participants were able to get accustomed to the simulator.

Figure 8 ***Test setup for user studies with 'No Driving Automation' (SAE Level 0) scenario***

Following the familiarization drive, participants were asked to maintain a target speed of 120 km/h, causing them to be in the left of two lanes, passing slower traffic on the right. For an illustration of the test scenario see *Figure 9*. Each participant experienced two test scenarios in random order, which were designed identically, but differed in the demand put on the driver due to the presence or the absence of a non-driving related task (NDRT).

Participants drove on the left lane of a two-lane highway at a target speed of v = 120 km/h, passing slower traffic on the right. Suddenly a vehicle from the right lane, moving with a speed of v = 100 km/h, cuts in front of the test vehicle (time-to-collision TTC = 5 seconds). After 0.8 seconds the cut-in vehicle contacts the lane and the multi-modal warning of the driver takes place ($TTC_{warn} \approx 4.2$ seconds) according to the experimental conditions. The only correct reaction of the driver, to handle the situation, was braking as the highway is equipped with crash barriers and therefore swerving is not an option.

After experiencing the situation, participants were asked to stop in the emergency lane and participate in a short question and answer session. The participants were then again instructed to proceed driving and maintain a speed of 120 km/h. Before participants experienced the described test situation a second time, a dummy event followed aimed at the prevention of learning effects. The dummy event consisted of driving through a construction site with a (uncritical) braking event with a preceding vehicle. After experiencing the cut-in event a second time, the questioning was repeated.

During one of two test drives, participants were instructed to carry out a non-driving related task (NDRT) as quickly and accurately as possible. The chosen task was the *Surrogate Reference Task* (*SuRT*) introduced by Mattes et.al [31], which is an artificial motor-visual task that allows distraction of the drivers under controlled conditions.

Multimodal warning system and stimuli

Participants were presented one of the three multimodal collision warning systems considered in the study. Today, requirements for the design of warning concepts already exist and mostly consider auditory alerts combined with visual information. To assess the impact of the tactile stimulus it replaced the auditory cue in one configuration and was added as a second alerting cue. The design of the stimulus modalities was identical across the combinations:

- *Acoustic and visual (A-V)*: A standard acoustic stimulus using a standard warning tone ('beep') and a visual sign ('hands-on-wheel' with a 'general caution' symbol) was presented in a heads-up display
- *Tactile, and visual (T-V)*: A vibrotactile stimulus by the Active Seat Belt replacing the acoustic stimulus was combined with a visual sign ('hands-on-wheel' with a 'general caution' symbol) presented in the heads-up display of the car
- *Acoustic, tactile, and visual (A/T-V)*: Three modalities as combination of acoustic, tactile, and visual stimulus for the take-over request

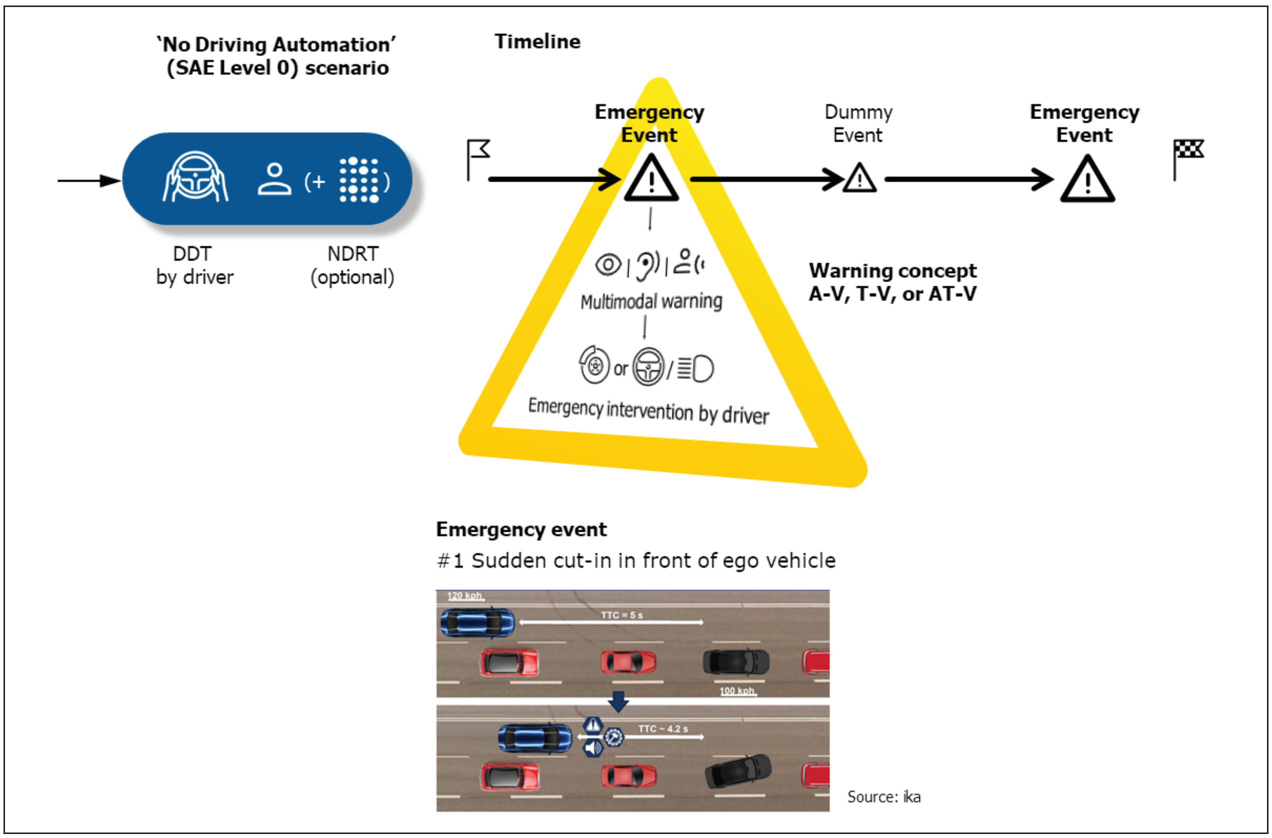


Figure 9 Characteristics of test scenario without driving automation (SAE Level 0)

Sample

A total of $N = 42$ drivers (17 female, 25 male) participated in the study. The participants' mean age was $M = 40.5$ years, ranging from 20 to 66 years ($SD = 14.6$). On average, the drivers had a driving experience of 22.6 (M = 20). The average annual mileage as a driver was $M = 15{,}295$ km ($SD = 15{,}418$ km).

Subjects were asked about their practical experience with driver assistance systems. 31% reported using a Cruise Control System (*CCS*) on a regular basis and 36% used it at least several times before. The regular use of Dynamic Cruise Control (DCC) and of a Lane Keep Assistant (LKA) was reported by 5% and 10% of the drivers.

Results

Objective data

Measures: To describe the driver's response to the critical situation and, accordingly, the warning concepts, the *reaction time* was of particular interest. In the simulated emergency this was the time from the beginning of the event (cut-in vehicle begins to leave its own lane) to a measurable reaction of the driver, i.e., brake pedal deployed.

As braking is considered the only correct response, the time until the brake is applied (brake time) is of particular importance. To investigate the effect of a tactile or acoustic cue, each in combination with visual information, brake times when using the acoustic-visual and tactile-visual warning strategies were compared.

The data was analysed using the statistics software IBM SPSS Statistics, Version 27. Abbreviations in all tables are to be understood as follows: *SS* = total sample size, *N* = analysed samples, *M* = mean, *SD* = standard deviation. Inferential statistical analysis was carried out by means of one-way ANalysis Of VAriance (*ANOVA*).

Data suggests that the response time is shorter when using the tactile-visual system variation rather than using the acoustic visual system variation ($F_{1,26} = 5.275$, $p = 0.03$, $\eta^2 = .169$). Please see *Figure 10* for the results.

In addition, it was investigated whether the addition of a tactile cue to a system variation consisting of an acoustic stimulus with visual information (A/T-V) adds further value in terms of the warning effect (operationalized here via braking time). Although the mean braking time for the acoustic-tactile-visual system variation appears to be descriptively shorter compared to the acoustic-visual system variation, inferential analysis shows no significant differences.

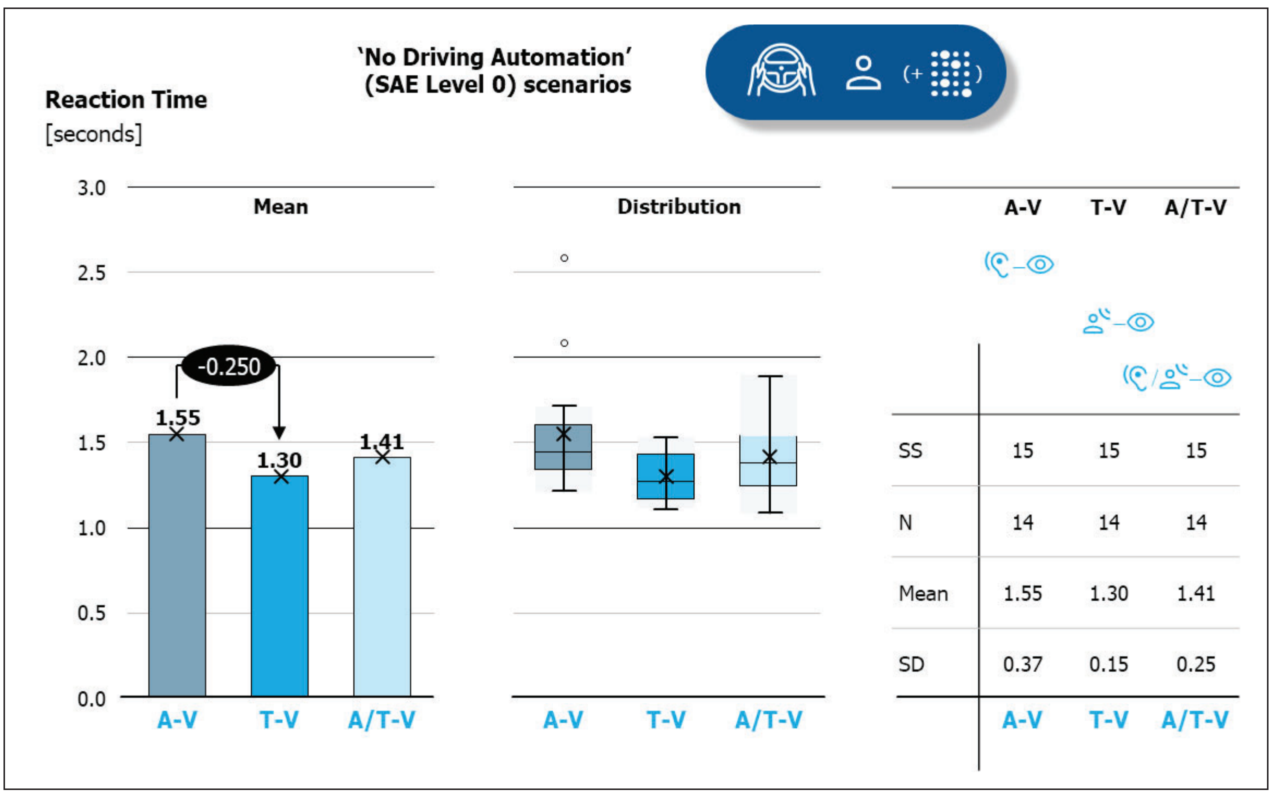


Figure 10 ***Reaction time for alternative warning concepts A-V, T-V and A/T-V after emergencies / SAE Level 0 scenarios***

Subjective data

Following the critical driving scenario, subjective data was collected by means of a questionnaire.

Questionnaires: The response of the participants to the different warning strategies was examined on a subjective level regarding the *cognitive demand*, the *perceived efficiency* of the warning effect, as well as with regard to *acceptance values* and *usability criteria*. A selection of the instruments and items used is described below.

Cognitive demand of the participants during the tests was assessed by a 5-staged response scale from 'not demanding at all' ('-2') to 'very demanding' ('+2').

Perceived efficiency was assessed by means of two Likert scales referring to the support provided for choosing the right reaction ('The warning system helped me to react *correctly*') and to the resolution of a fast response ('The warning system helped me to react *quickly*'). A 5-staged response scale was used for both items from 'do not agree at all ('-2') to fully agree ('+2').

The Van der Laan Acceptance Scale [32] was used to assess *system acceptance* by applying the two scales usefulness and effective satisfaction. Both scales are collected by means of nine items in total. Each item consists of an opposite pair of adjectives and the participants must indicate how far the adjectives apply (score reaches from -2 to +2) to the experienced system.

The *cognitive demand* in the situation does not seem to be influenced by the variation of modalities in the warning strategies. In both system variations acoustic-visual (A-V) and tactile-visual (T-V) users rated the cognitive demand almost identically on the 5-staged rating scale ranging from -2 (*'not demanding at all'*) to +2 (*'very demanding'*) (A-V: $M = 0.36$, $SD = 0.88$; T-V: $M = 0.38$, $SD = 0.82$). Also, within the experimental condition with the system variation including all three modalities the demand in the situation does not appear changed ($M = 0.46$, $SD = 1.12$).

Regarding the *subjectively perceived efficiency* of the system variations there seems to be a slight tendency towards the tactile-visual (T-V) option on a descriptive level. On average, participants who have experienced the tactile-visual (T-V) system are slightly more likely to agree with the statement that the warning system helps them to react correctly than participants who have experienced the acoustic-visual (A-V) option on the 5-staged rating scale from -2 (*'do not agree at all'*) to +2 (*'fully agree'*) (A-V: $M = -0.08$, $SD = 0.88$; T-V: $M = 0.46$, $SD = 0.89$; A/T-V: $M = 0.23$, $SD = 1.41$). Still, there is no clear (dis)agreement with the statement for either of the three options for multimodal warning. The same applies to the agreement with the statement that the respective warning strategy helps reaction time. Within the context of no clear (dis)agreement, on a descriptive level it appears highest for the tactile-visual system (A-V: $M = 0.00$, $SD = 0.92$; T-V: $M = 0.62$, $SD = 0.88$; A/T-V: $M = 0.15$, $SD = 1.20$).

System acceptance ratings collected via Van der Laan Acceptance Scale [32] were very similar for all three options with a slight benefit on a descriptive level for the system tactile-visual (T-V) regarding the factor *usefulness*. Please see *Figure 11* for results.

	Acoustic-visual A-V		Tactile-visual T-V		Acoustic-tactile-visual A/T-V	
	Usefulness	Satisfaction	Usefulness	Satisfaction	Usefulness	Satisfaction
N	13*	12	13	13	12*	12*
Mean	0.41	0.5	0.95	0.48	0.66	0.55
SD	0.98	0.83	0.65	0.71	0.91	0.80

Note: * = mean n in the case of missing values in single items

Figure 11 System acceptance ratings collected via Van der Laan Acceptance Scale

Further questionnaire responses indicate that the tactile warning cue is rarely not perceived (T-V alert: $n = 2$, A/T-V alert: n = 3). The visual warning cue is the most frequently not perceived across all system variations (A-V alert: $n = 6$; A-V alert: $n = 6$).

USER STUDY TACTILE WARNING DURING CONDITIONAL DRIVING AUTOMATION

A second user study on the assessment of tactile warning through an Active Seat Belt system had been carried out with the focus on the take-over request (TOR) during conditional driving automation.

Thus, a driving simulator study was conducted by the Department of Engineering and Traffic Psychology of TU Braunschweig in the dynamic driving simulator of the NFF (Niedersächsisches Forschungszentrum Fahrzeugtechnik; Automotive Research Centre Niedersachsen) comparing three types of warning signals in order to evaluate the benefits of haptic signals for take-over requests in automated driving at SAE Level 3.

At SAE Level 3, the car driver becomes a passenger for part of the journey and may engage in other activities like reading, using the smartphone or laptop, watching videos etc. However, the SAE Level 3 system has limits where the driver has to take over. About 10 to 15 seconds before these limits are reached, the driver is warned and requested to take over control of the car (TOR). At that moment, acoustic and visual warnings are used to issue this TOR. However, when the driver listens to loud music or videos, this warning may take some time to perceive and react to. Additionally, there may be other signals or warnings which may prevent the driver from understanding that this current warning is a take-over request. Haptic signals may be advantageous in this situation.

Method and procedure

The study was conducted in the dynamic driving simulator at the NFF (see *Figure 12*). The front part of a car is situated on a moving platform. Six screens are used to present the scenery. The platform including the car and the screens are moved to simulate a realistic driving feeling including the typical forces on the driver within a vehicle.

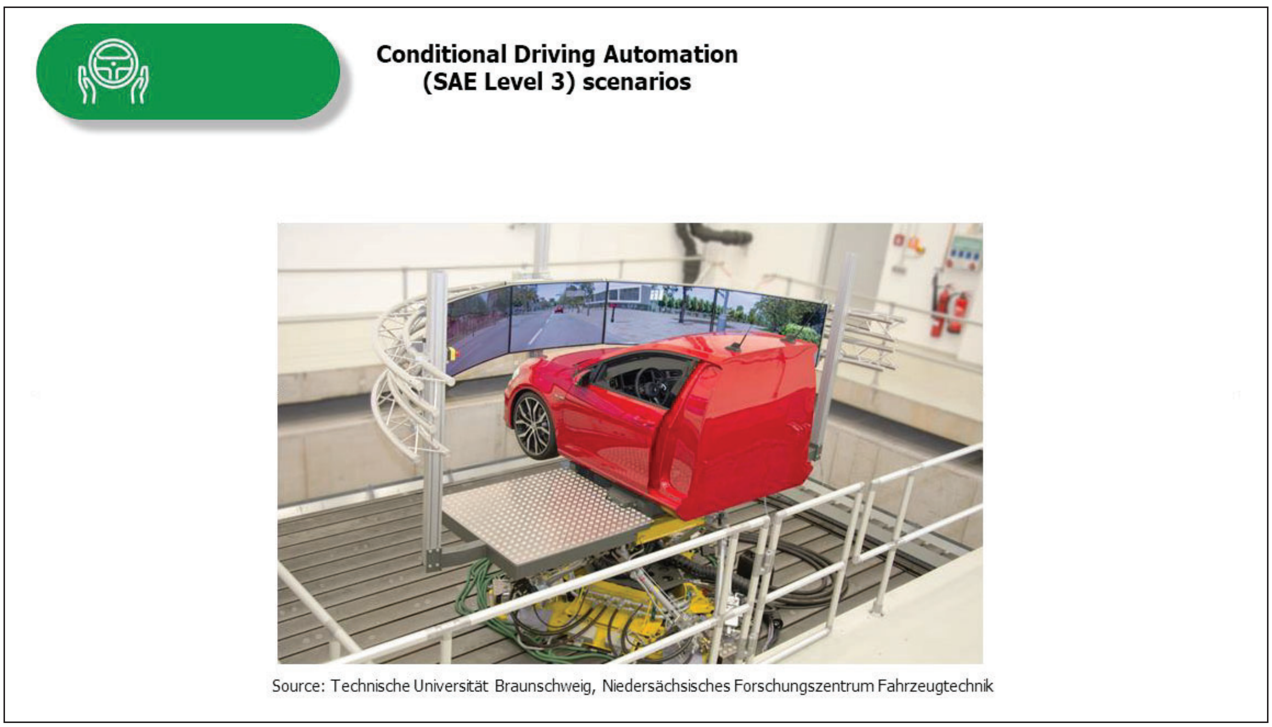


***Figure 12** Test setup for user studies with '<u>Conditional Driving Automation</u>' (SAE Level 3) scenario*

On a simulated, two-lane highway, drivers started handing over the driving task and continued in a conditional driving automation mode (SAE Level 3). During the simulated rides, three different emergencies were presented to the users resulting in a take-over-request:

- The first emergency scenario consisted of a TOR due to missing lane markings on the road ahead (see *Figure 13*, *#1*). About 10 seconds before the vehicle reached that location, the TOR was given. After take-over, the driver manually controlled the vehicle through this stretch of the road until the markings reappeared and the automation could take over again.

- In the second emergency scenario, a crash had happened and one of the crashed cars was partially blocking the right lane (see *Figure 13*, *#2*). A police car was present at the right shoulder and a hazard triangle had been placed on the road. Here, the drivers had to take over, had to change the lane and then had to pass the crash before re-engaging the conditional driving automation mode.
- The third emergency scenario was suddenly starting heavy rain (see *Figure 13*, *#3*) with the vehicle requesting the driver to take over due to limitations of environmental sensing under these circumstances. Later, the drivers re-engaged the conditional driving automation mode.

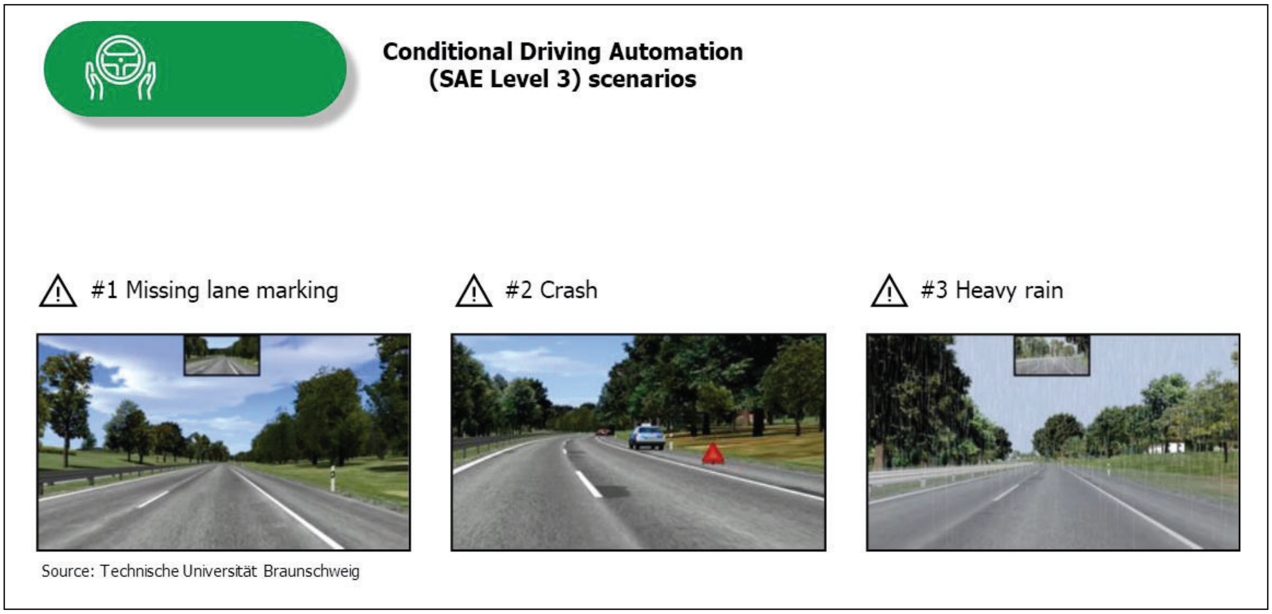


Figure 13 Emergencies resulting in a take-over-request

Each driver experienced one type of take-over request in these three situations as one trip. Three different types of take-over requests were compared by having each driver do three trips. The order of the three situations was different in each of the three trips. Moreover, the order of the three types of take-over requests was also varied in three different groups of drivers in order to control for time effects. The three kinds of take-over requests were the following:

- A standard acoustic and visual request, using a standard warning tone (‘beep’) and a symbol in the head-up display (A-V – acoustic-visual) which was shown in a simulated head-up display.
- A tactile and visual request, replacing the warning tone with a jerking movement of the seat belt (T-V – tactile-visual)
- A combined tactile, acoustic and visual take-over request thus adding a haptic signal to the standard acoustic-visual take-over request (A/T-V – acoustic-tactile-visual)

The vibrotactile stimulus was at a frequency of 9 Hz at a force level of approximately 50 N. This configuration was chosen based on previous experiences aiming for a salient but acceptable stimulus.

After receiving this take-over request, the driver could turn back to manual driving by simply braking or by using the headlight flasher.

During the trip, take-over times and driving behavior after taking over were recorded. After each trip, drivers were questioned with regard to usability and subjective acceptance of the system. At the end of the study, each driver was asked to select the kind of feedback that he or she would like most to have.

45 drivers (15 females and 30 males) were recruited from the subject pool of the NFF ranging from 20 to 73 years of age in order to have a heterogeneous sample with regard to age and sex. The mean age was 37 years (sd = 15 years). On average, the drivers had obtained their driving license 19 years ago. About half of the sample drove less than 9,000 km per year but there was also one driver going more than 50,000 km per year. Subjects were asked about

their attitude towards driving assistance system (without qualifying this further) and about 75% of them had a positive or very positive attitude. About 70% reported being very interested in new technologies. Thus, the sample was a wide mix of different drivers, however, with a somewhat positive bias towards new technologies and driver assistance systems.

Results

Driving behavior

The most relevant parameter is take-over time. This was measured from the start of the take-over signal to the time-point of deactivating the automation. As the experimental design was a repeated measurements design, loss of data due to technical problems in even one situation led to exclusion of the subject. Data from 34 drivers was fully complete and used for the analyses of driving behavior. For the analysis, a 3x3 (type of situation x type of warning) repeated measures ANOVA was conducted.

For take-over time, a significant main effect was found for type of warning ($F_{2,66} = 9.7$, $p = 0.000$), no effect of the driving situation ($F_{2,66} = 0.1$, $p = 0.863$) and no interaction ($F_{4,132} = 0.1$, $p = 0.988$). Thus, independent of the driving situation, there was a clear effect of the type of warning, which is shown in *Figure 14 (left)*. Take-over time with the typical acoustic-visual warning was on average 5.4 seconds. Replacing the acoustic warning with the haptic warning, take-over time was reduced to 4.6 seconds, thus 0.8 seconds faster with the haptic as compared to the acoustic warning. Adding the haptic component to the acoustic-visual warning resulted in the fastest take-over time of 4.4 seconds, thus one second gain by the additional haptic signal.

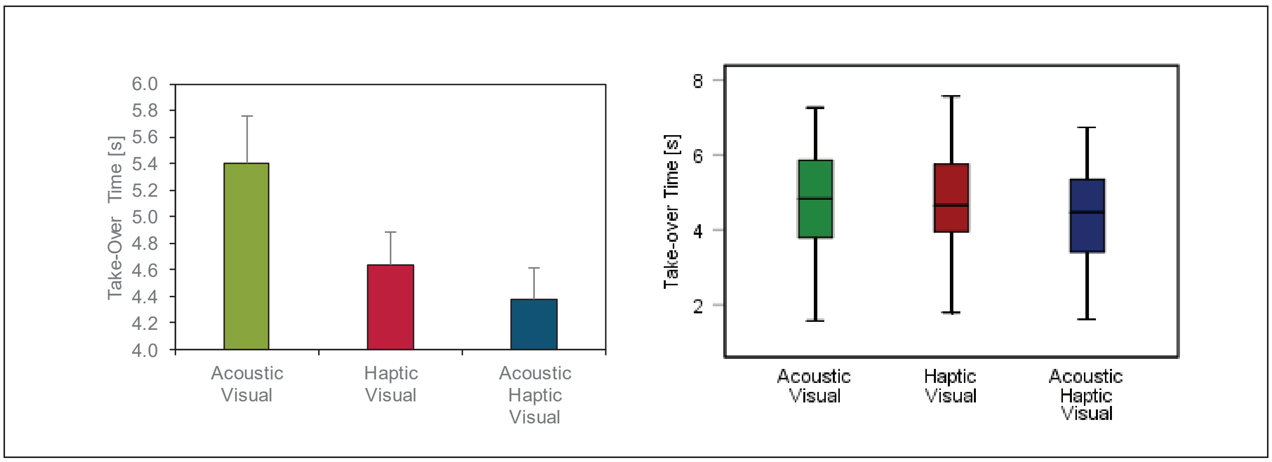


Figure 14 ***Average take-over time in seconds for the different warning signals. The left gives mean and standard error. At the right, the same data are given as boxplots showing the distribution of the take-over times***

Figure 14 (right) shows the same data as boxplots. In every condition, there is a large variation of take-over times reflecting the individual take-over behavior of the drivers, ranging from 1.6 to 7.6 seconds as averaged for each driver over the three driving situations with each warning signal. On an individual level, comparing individual take-over times with the haptic warning instead of the acoustic warning, 66% of the take-over times in the different situations were shorter with the haptic warning. With the haptic warning added to the acoustic-visual warning, 75% of all take-over situations benefited from this additional component with a faster take-over time.

With regard to the driving behavior after taking over manual control of the car, the maximum speed when driving manually was taken as an indication of adaptation to the situation as having a lower maximum speed indicates a better adaptation to the take-over situation with going more slowly being more careful. For maximum speed, there was a main effect for type of warning ($F_{2,66} = 24.9$, $p = 0.000$) and a main effect of the driving situation ($F_{2,66} = 21.8$, $p = 0.000$) but also an interaction ($F_{4,132} = 10.6$, $p = 0.000$).

First of all, the adaptation of speed depends on the criticality of the situation after take-over with the smallest reduction in speed at the missing lane markings, followed by the crash and then heavy rain. The amount of adaptation is clearly different for the three warning types. The least adaptation is shown with an acoustic-visual warning. The tactile-visual warning increases the reduction in speed with the crash and the heavy rain. Adding the tactile warning to the acoustic-visual warning leads to the strongest adaptation, especially when heavy rain occurred.

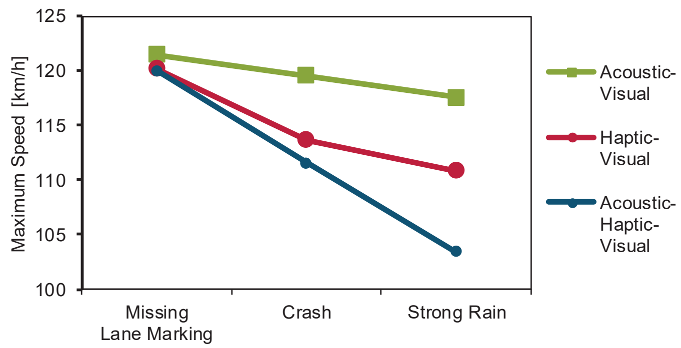


Figure 15 Maximum speed in the three take-over situations with the three different warning signals

With regard to lateral control of the car, the standard deviation of the lane position was computed (SDLP). For this parameter, there was neither a main effect for type of warning ($F_{2,66} = 0.7$, $p = 0.492$) nor an interaction ($F_{4,132} = 0.6$, $p = 0.656$), but only a main effect of the driving situation ($F_{2,66} = 3.7$, $p = 0.000$). The SDLP is largest in the crash situation. This is due to the lane changes which had to be done to avoid colliding with the crashed car. Otherwise, the three different warning signals did not influence the lateral control of the car.

Subjective evaluation

After each trip drivers were asked to evaluate - depending on the warning signal condition - the visual warning symbol (included in every warning type), the acoustic warning (in the acoustic-visual and acoustic-tactile-visual condition) and the tactile warning (in the tactile-visual and acoustic-tactile-visual condition) separately with regard to the signal being clear, urgent and annoying. A repeated measures ANOVA was conducted for each of these questions (data from 43 drivers). The warning components did not differ with regard to being clear ($F_{2,84} = 1.5$, $p = 0.227$) or being annoying ($F_{2,84} = 1.4$, $p = 0.251$), but in being urgent ($F_{2,84} = 6.6$, $p = 0.002$).

All components were equally clear (between the category 'somewhat clear' and 'clear' and medium annoying. The visual symbol was rated less urgent than the tone and the tactile vibration. However, all three were within the category 'urgent'.

With regard to the design of the tactile component, all ratings lie in the middle to positive range of the scales. The vibration is not really pleasant but also not unpleasant. It can very well be noticed, is not too weak but also not too strong. Time-point and duration are good to very good.

When directly asking afterward, which component of the system was most helpful, 69% of the drivers indicated the tactile vibration, 49% the tone and 13% the visual symbol in the head-up display (drivers could also indicate two or even three components).

Finally, drivers were asked which combination they would buy if they ever bought a car with SAE Level 3 automation. The typical acoustic-visual HMI would be bought by 22% of the drivers. 29% would prefer the tactile-visual system and 42% the acoustic-tactile-visual warning system. 7% would like an acoustic-tactile system (which was not included in the study).

SUMMARY OF MAIN RESULTS FROM BOTH STUDIES

The warning concept with *acoustic alerting* and *visual information* ('A-V') was used as a reference configuration for both studies (manual driving and conditional driving automation) as it is the usual warning concept for critical driving situations in vehicles today. The efficiency of the warning concept was measured by comparing the driver's reaction time between the reference warning concept ('A-V') and two alternatives – one with *additional tactile* alerting ('A/T-V') and the other one with *tactile alerting* replacing the *acoustic cue* ('T-V').

In the first test setup '**manual driving**' undistracted drivers (as the nominal driver condition) had been exposed to a sudden situation requiring them to brake.

- Measures: The reaction time with the tactile alert ('T-V') was **reduced by 250 ms** (mean) in comparison to the reference warning concept ('A-V').
- Calculated safety effect: The impact of this faster reaction time is a reduction of a stopping distance of 6.6 m for a vehicle at a speed of 100 kph. According to a report on real world accident data and the impact analysis of active safety technologies [33] this reduction of reaction time may result in 27.5% avoided collisions.

Warning concepts	'No Driving Automation' (SAE Level 0) scenarios – Measures: Reaction time emergency scenarios (mean)	'No Driving Automation' – Measures: Difference reaction time compared to ref. warning concept (A-V)	'No Driving Automation' – Calculated safety effect: Stopping distance reduction from 100 kph (calculated)	'No Driving Automation' – Calculated safety effect: Collisions avoidance potential (calculated) [33]	Conditional Driving Automation (SAE Level 3) scenarios – Measures: Reaction time Take over request (TOR) (mean)	Conditional Driving Automation – Measures: Difference reaction time compared to ref. warning concept (A-V)	Conditional Driving Automation – Calculated safety effect: Stopping distance reduction from 100 kph (calculated)
Acoustic-visual (A, V)	1,55 s	reference	reference	reference	5,4 s	reference	reference
Tactile-visual (T, V)	(1,41 s)	*measures statistically not significant*			4,4 s	**-1 s** (-18%)	**-28 m** (18,5%)
Acoustic-tactile-visual (A, T, V)	1,3 s	**-250 ms** (-16%)	**-6,6 m** (-17%)	**-27,5 %**	4,65 s	**-750 ms** (-14%)	**-21 m** (14%)

Use of modalities for the warning concepts: Alert | Warning information

Figure 16 Summary of measures and calculated safety effects between warning concepts

The reaction time with an additional tactile alert ('A/T-V') was reduced by 140 ms, but the difference was not significant. Mental overload during the demanding emergency situation combined with a redundant alerting (acoustic and tactile) and the additional visual information could be a reason for this result. Such effect has been detected before [14] but may need to be investigated in further studies.

In the second test setup '**conditional driving automation**' drivers were focused on a non-driving-related task (NDRT) when different emergency situations occurred. They had to react with a take-over request (*TOR*).

- Measures: The reaction time after the warning with the tactile alert ('T-V') was **reduced by 750 ms** in comparison to the reference warning concept ('A-V'). The reaction time with additional tactile alert ('A/T-V') was reduced by **1.0 second**.
- Safety effect: For the 'T-V' warning the safety effect of this faster reaction time is a reduction of a driven distance at 100 kph by 21 m compared to the 'A-V' warning. The safety effect of the faster reaction time of the 'A/T-V' warning is a reduced stopping distance of 28 m for a vehicle at a speed of 100 kph.

Here, the addition of a second alerting stimulus resulted in an additional improvement of the reaction time and the related safety effects. An explanation could be that the mental load of drivers in the automated driving mode (but with NDRT) might be lower as the mental load of drivers in a demanding emergency. And a lower mental load could allow for better processing of the 'A/T-V' warning concept using three sensory channels simultaneously.

LIMITATIONS

First, some limitations are rooted in the study design. Both studies have been implemented with dynamic drive simulators and with 87 users in total: 42 (manual driving) and 45 (AD Level 3). However, due to the variations of test parameters as described above in the chapter *Methodologies*, analyzable sample data sets are based on 14 (manual driving) and 34 (AD Level 3) users. The specific implementation of the tactile and auditory alerting stimuli and visual information may have a further influence on the measurements. Each vehicle's HMI system may have differing characteristics, and therefore influence the individual performance of warning and alerting in emergency situations. For example, tactile alerting stimuli could be realized by seat vibration and visual warning information could be displayed in a head-up display.

Second, the *driving scenarios* had been limited to *highways*. Users had a main task, either DDT or an NDRT, were *not distracted from this task.* In addition, required sensory channels were *not* loaded with *additional* noise. Further studies could help to extend the range of driving scenarios and environmental influences.

Third, the *dynamic simulator-based* study results should be compared to investigations of *real-world driving* studies with natural conditions or events influencing the emergency situations. However, the authors are aware of comparable investigations based on real-world driving that already support the findings of this work.

Fourth, the *mental states (awake, attentive, focused)* of the users have not been varied during the simulations. A detailed analysis of the influence of the mental states on the effectiveness of warning concepts would be desirable. Additionally, users might perceive their participation during the study as a *test situation* and study data might be influenced by this.

Fifth, *warning concepts* are part of the human machine communication. The *concept* found in this paper *'alert and inform'* with its multimodal options as e.g. 'A-V', 'T-V' or 'A/T-V' can be expanded to non-critical situations leading to the concept variant '*notify and inform*'. This should be the subject of further investigations. Further concept options as e.g., 'alert only' (e.g., 'T- ') or 'inform only' (e.g. ' - V') or any other have not been investigated and could be subject of further work to better understand how to support human alerting and perception (situation awareness) with technologies available in vehicles.

Sixth, since warning is just a crucial step to achieve safe driving, it would be desirable to investigate the effectiveness of warnings also in the context of the whole process chain. In this study vehicle control parameters after the warnings have been measured, also to measure whether the reaction of the warning was correct or leading to undesired reactions (not reported here in detail) but detailed investigations of vehicle control and threat control performance in different driving and threat scenarios would be desirable to allow an assessment of how intuitive various warning concepts are in different threatening situations.

Seventh, the intent of this study was to employ prototypical examples of the concepts ('A-V', 'T-V' and 'A/T-V') to provide a general comparison of driver response to each. Therefore, the conclusions provided in this section emphasize general principles and relative effectiveness of general concepts rather than detailed design specifications.

CONCLUSION

This work answers the four research questions (see *Research Questions and Methodology*) in the following way:

1: Advantages of a tactile stimulus for a warning concept of an ADS system
The work presented in this paper found that the *replacement* of the *auditory alerting stimulus by a tactile seat belt alert* in warning concepts with additional visual warning information can significantly contribute to an improved effectiveness of driver warnings. Furthermore, the study data shows that an *additional tactile alerting stimulus* by an Active Seat Belt in a common warning concept with auditory alerting and visual warning information may improve reaction time and effectiveness of driver warning – depending on the selected driving mode and related workload level of the drivers.

2: Conclusions regarding further aspects of efficient warning
The results also *suggest* that during *demanding driving situations*, as e.g., manual driving in dense traffic, *single alerts* with visual warning information can be more effective than *complex* warning concepts with *dual alerts* (acoustic with tactile) and visual information.

3: User feedback to tactile stimulus
Tactile alerting stimuli are well accepted by the study participants in both driving modes, manual driving and automated driving as well. The positive perception includes the usefulness and the preference to buy a vehicle with this feature.

4: Comparison of tactile stimulus of an Active Seat Belt vs. other options of tactile actuators
Active Seat Belt with a vibrotactile alert feature can help to improve the driver's ability to react faster than with alternative multi-modal warning and alerting scenarios during manual driving and in driving scenarios with conditional driving automation (SAE Level 3).

The vibrotactile alerting through an Active Seat Belt system is unique compared to other means of alerting (e.g., seat, steering wheel, pedal) as it is free of signal noise and therefore easy to identify by the users. Especially significant warning and alerting scenarios for ADS functions, e.g., Forward Collision Warning (FCW), Take-over requests (TOR), may also benefit from the Active Seat Belt.

The study provides a clear recommendation for tactile alerting provided by an Active seat Belt for ADS functions like FCW or TOR and should be considered for future HMI systems.

Collaboration and Funding

This paper and the related research were executed by the three parties:

- ZF Group, ZF Automotive GmbH, Germany with a focus on the system function, the tactile Active Seat Belt system, and the coordination of the user studies
- RWTH Aachen University, Institute for Automotive Engineering (ika) together with fka GmbH, Aachen Germany with a focus on the implementation of the user study 'Tactile Warning during driving without driving automation'
- Technische Universität Braunschweig, Institute of Psychology, Engineering and Traffic Psychology Germany with a focus on the implementation of the user study 'Tactile Warning during Conditional Driving Automation'

The development and execution of the two user studies were funded by ZF Automotive GmbH through 2019-2021.

REFERENCES

[1] “Euro NCAP Vision 2030“, European New Car Assessment Programme, 2022.

[2] “Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles”, Society of Automotive Engineers, Standard-No: J3016, 2021.

[3] T. J. Ayres, M. M. Gross, C. T. Wood, D. P. Horst, R. R. Beyer and J. N. Robinson, "What is a Warning and When Will it Work?", Proceedings of the Human Factors Society Annual Meeting, vol. 33, no. 6, p. 426–430, October 1989.

[4] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," Human Factors, vol. 37, no. 1, p. 32–64, March 1995.

[5] C. D. Wickens, J. Helleberg, J. Goh, X. Xu and W. J. Horrey, Pilot Task Management: Testing an Attentional Expected Value Model of Visual Scanning, 2001.

[5] C. D. Wickens, J. Helleberg, J. Goh, X. Xu and W. J. Horrey, “Pilot Task Management: Testing an Attentional Expected Value Model of Visual Scanning”, 2001.

[6] B. Hommel, J. Müsseler, G. Aschersleben and W. Prinz, “The Theory of Event Coding (TEC): A framework for perception and action planning”, Behavioral and Brain Sciences, vol. 24, no. 5, 2001.

[7] S. Ladwig, C. Sutter and J. Müsseler, “Crosstalk between proximal and distal action effects during tool use”, Zeitschrift fur Psychologie / Journal of Psychology, vol. 220, no. 1, p. 10–15, January 2012.

[8] S. Ladwig, C. Sutter and J. Müsseler, “Intra- and intermodal integration of discrepant visual and proprioceptive action effects”, Experimental Brain Research, vol. 231, no. 4, p. 457–468, October 2013.

[9] D. Kahneman, “Attention and Effort”, Prentice-Hall, 1973.

[10] “Assessment Protocol - Safety Assist - Implementation 2023”, European New Car Assessment Programme, 2022.

[11] J. C. Brill, M. Mouloua and S. D. Hendricks, “Compensatory strategies for managing the workload demands of a multimodal reserve capacity task”, 2009.

[12] M. J. Enriquez, K. E. MacLean and H. K. Neilson, “Interplay of Tactile and Visual Guidance Cues under Multimodal Workload”, 2007.

[13] F. Biondi, D. L. Strayer, R. Rossi, M. Gastaldi and C. Mulatti, “Advanced driver assistance systems: Using multimodal redundant warnings to enhance road safety”, Applied Ergonomics, vol. 58, p. 238–244, January 2017.

[14] I. Politis, S. Brewster and F. Pollick, “Evaluating multimodal driver displays under varying situational urgency”, 2014.

[15] R. S. Nickerson, “Intersensory facilitation of reaction time: Energy summation or preparation enhancement?”, Psychological Review, vol. 80, no. 6, p. 489–509, 1973.

[16] M. K. Ngo and C. Spence, “Auditory, tactile, and multisensory cues facilitate search for dynamic visual stimuli”, Attention, Perception, and Psychophysics, vol. 72, no. 6, 2010.

[17] C. Spence and C. Ho, “Tactile and multisensory spatial warning signals for drivers”, IEEE Transactions on Haptics, vol. 1, no. 2, 2008.

[18] S. Kennett, C. Spence and J. Driver, “Visuo-tactile links in covert exogenous spatial attention remap across changes in unseen hand posture”, Perception and Psychophysics, vol. 64, no. 7, 2002.

[19] C. Spence, M. E. R. Nicholls and J. Driver, “The cost of expecting events in the wrong sensory modality”, Perception and Psychophysics, vol. 63, no. 2, 2001.

[20] F. Beruscha, K. Augsburg and D. Manstetten, “Haptic Warning Signals at the Steering Wheel: A Literature Survey Regarding Lane Departure Warning Systems (Short Paper)”, 2011.

[21] S. B. Brown, S. E. Lee, M. A. Perez, Z. R. Doerzaph, V. L. Neale and T. A. Dingus, “Effects of haptic brake pulse warnings on driver behavior during an intersection approach”, 2005.

[22] H. de Rosario, M. Louredo, I. Díaz, A. Soler, J. J. Gil, J. S. Solaz and J. Jornet, “Efficacy and feeling of a vibrotactile Frontal Collision Warning implemented in a haptic pedal”, Transportation Research Part F: Traffic Psychology and Behaviour, vol. 13, no. 2, p. 80–91, March 2010.

[23] W. Chang, W. Hwang and Y. G. Ji, “Haptic seat interfaces for driver information and warning systems”, International Journal of Human-Computer Interaction, vol. 27, no. 12, p. 1119–1132, December 2011.

[24] L. M. Stanley, “Haptic and Auditory Cues for Lane Departure Warnings”, Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 50, no. 22, pp. 2405-2408, 2006.

[25] T. B. Sayer, J. R. Sayer and J. M.H. Devonshire, “Assessment of a Driver Interface for Lateral Drift and Curve Speed Warning Systems: Mixed Results for Auditory and Haptic Warnings”, 2005.

[26] J. D. Lee, J. D. Hoffman and E. Hayes, “Collision warning design to mitigate driver distraction”, 2004.

[27] S. M. Petermeijer, D. A. Abbink, M. Mulder and J. C. F. de Winter, “The Effect of Haptic Support Systems on Driver Performance: A Literature Survey”, IEEE Transactions on Haptics, vol. 8, no. 4, pp. 467-479, October 2015.

[28] J. J. Scott and R. Gray, “A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving”, Human Factors, vol. 50, no. 2, 2008.

[29] E. Adell, A. Várhelyi, M. D. Fontana and L. Bruel, “Test of HMI Alternatives for Driver Support to Keep Safe Speed and Safe Distance - A Simulator Study”, The Open Transportation Journal, vol. 2, pp. 53-64, October 2008.

[30] J. Chun, I. Lee, G. Park, J. Seo, S. Choi and S. H. Han, “Efficacy of haptic blind spot warnings applied through a steering wheel or a seatbelt”, Transportation Research Part F: Traffic Psychology and Behaviour, vol. 21, pp. 231-241, November 2013.

[31] S. Mattes and A. Hallén, “Surrogate Distraction Measurement Techniques: The Lane Change Test”, 2009.

[32] J. D. V. D. Laan, A. Heino and D. D. Waard, “A simple procedure for the assessment of acceptance of advanced transport telematics”, Transportation Research Part C: Emerging Technologies, vol. 5, no. 1, 1997.

[33] H. Schittenhelm, “Advanced Brake Assist – Real World effectiveness of current implementations and next generation enlargements by Mercedes-Benz”, 2013.

[34] L. Petersen, L. Robert, X. J. Yang and D. Tilbury, “Situational Awareness, Driver’s Trust in Automated Driving Systems and Secondary Task Performance”, SAE International Journal of Connected and Automated Vehicles, vol. 2, no. 2, 2019.

[35] N. Lerner, J. Jenness, E. Robinson, T. Brown, C. Baldwin and R. E. Llaneras, “Crash Warning Interface Metrics: Final Report”, United States, Department of Transportation, National Highway Traffic Safety Administration Report NHTSA, No.: DOT HS 811 470a, Rockville, MD, 2011.

A HUMAN MACHINE INTERFACE SUGGESTED FROM NEUROSCIENTIFIC ANALYSIS OF HUMAN FACTOR

Shin'ichi Murakami
Koji Dan
Toshiaki Seo
Takaya Yamazaki
Miki Cho
Minoru Higuchi
Honda R&D Co., Ltd.
Japan

Daisuke Matsuyoshi
Ryota Kanai
Araya Inc.
Japan

Yasunori Aizawa
Makiko Yamada
National Institutes for Quantum Science and Technology
Japan

Paper Number 23-0232

ABSTRACT

Honda aims for zero traffic collision fatalities involving Honda motorcycles and automobiles globally by 2050. To realize a zero traffic-incident society, we need to minimize human driver errors. Improper processing of information should trigger human errors during driving, however, despite its importance for our society, neural mechanisms during driving that can lead to catastrophic traffic consequences remain unclear. To clarify these, we have researched the relationships among drivers' manipulation, gaze, and brain activation. In particular, we have focused on the human eye gaze because it is not only a passive input organ, but also reflects dynamic information processing in the brain. To investigate the human brain mechanisms involved in safe and secure driving, we scanned the human brain using functional magnetic resonance imaging(fMRI) while driving in an MRI-compatible driving-simulator. We introduce one of the experiments showing differential brain activation between safe drivers and control drivers manipulating a vehicle in ordinary traffic conditions. In this experiment, participants were healthy adults, and they manipulated the driving-simulator in the MRI scanner, while their

driving manipulation and gaze were monitored. The participants encountered risk factors in the driving scenarios. We extracted the difference in the brain activation at gazing some risks between the safe drivers and control drivers, then the differences in brain activity between safe drivers and others were found in the precuneus, V1, and SMA. Then we constructed a human-machine interface (HMI) that aimed to complement and enhance the cognitive processing which is necessary for safe driving. To verify the efficacy of our HMI, we conducted experiments by using driving-simulator composed of the front part of N-BOX(Honda) and 5 displays (65 inches), the original system. As a result, the suggested HMI could have effect on early noticing and avoiding high-risk object. It is possible, therefore, that general drivers began to drive more safely with a safe driver-inspired information processing assistance system. Our findings will help elaborate the specification of devices for ADAS and ADS.

INTORODUCTION

Traffic incidents are one of the critical causes of death worldwide approximately 1.3 million people lose their lives on the road. Traffic injuries lead to some serious losses not only to individuals and their families but also to nations as a whole. It is said that road traffic crashes cost most countries 3 % of their gross domestic product [1]. To serve people worldwide with the "joy of expanding their life's potential", Honda is having the effort to realize "zero traffic collision fatalities involving Honda motorcycles and automobiles globally by 2050." That said, no matter how much we improve safety technologies for motorcycles and automobiles, these alone will not eliminate traffic collision fatalities completely. We need to work on making people's driving behavior safe by eliminating human errors. Thus far, several studies examined the association between traffic collision fatalities and human errors and suggested that inattention such as failure to notice risky objects is a critical factor underlaying them [2] [3] [4] [5]. In order to understand the underlying factors of human errors, the present study focused on the relationship between driving behaviors leading to human error and human brain activity. In previous studies, Oba et al. reported that frontoparietal control network activity positively correlated with better lane-keeping [6]. Schweizer et al. studied the tasks in which participants had to drive at simple or complex intersection. They found that in the complex condition, brain activation shifted from the posterior visuospatial attentional system to the frontal system related to multi-tasking and divided attention [7]. In this study, we focused on hard braking caused by the driver's failure to notice high-risk as a human error. Our purpose is to propose a methodological instance for constructing a neuroscience-inspired driving support system while having two primary aims: 1) To investigate cognitive function related to safe and secure driving, especially noticing risky objects, form a neuroscientific point of view and 2) to ascertain the effects of prototyped HMI based on these neuroscientific findings.

NEUROSCIENTIFIC EXPERIMENT

Devices

In order to investigate the cognitive processes related to the safe and secure driving, a driving-simulator system (MRI compatible, shown in Figure 1) was used for experiment on fMRI (functional magnetic resonance imaging). The fMRI is the main non-invasive method used to research some cognition mechanisms from brain activation. All fMRI data were acquired with a Siemens Verio 3T MRI system (Siemens, Germany) located on the National Institutes for Quantum Science and Technology (QST). The manipulation part of the driving simulator consisted of an acceleration pedal and a braking pedal and a steering wheel (Driving System, CURRENT DESIGNS, USA) added two buttons input device (CURRENT DESIGNS, USA) as a turn signal. Participant's gaze was measured by the Live Track AV for fMRI (Cambridge Research Systems, UK).

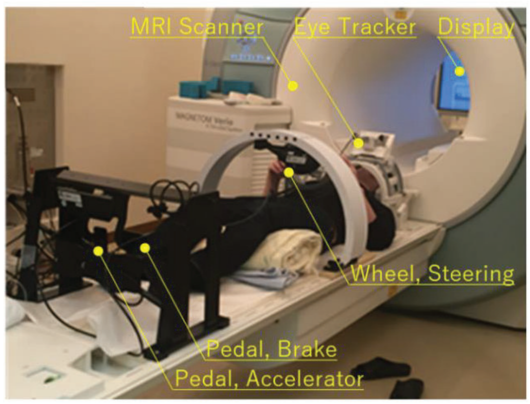


Figure 1. Overview of experimental driving-simulator composed of MRI-compatible manipulations and eye tracking device.

Procedure

The runs presented on the driving simulator were made by UC-win/Road (FORUM8 Co., Ltd., Japan). A run had some "blocks" composed of the preparing part that took about 5 seconds and the driving part that took about 30 seconds. In the preparing part, an instruction for changing lane such as "Right", "Left" or "Straight" was displayed and participant would be asked to follow the message shown in Figure 2(a). In the running part, participant drove at 40 km/h as initial velocity in some scenes which were modeled after some frequently road-accident areas in Japan. Figure 2(b) shows some samples of the scene used in this experiment. A traffic light located in the latter half of block was designed. There were three conditions about blocks in this experiment. First, in the target condition, participants were asked to change lane for right and stop at the traffic light. Second, in the dummy condition, they were required to change lane for left or right or go straight, sometimes not to stop at the traffic light. Third, the baseline condition which participants drove on no others in the scene was prepared.

A run had 8 target blocks, 6 dummy blocks and 3 baseline blocks that shown on the display randomly. In the target and dummy blocks, risky traffic objects were shown with slight difference (body color, velocity, type of car, etc.) in each block. Four runs were conducted per a participant, it took about 45 minutes totally including rest time as shown in Figure 3. Prior to commencing this experiment, all participants had a practical driving during about

5 minutes for getting used to manipulate this driving simulator in the MRI system.

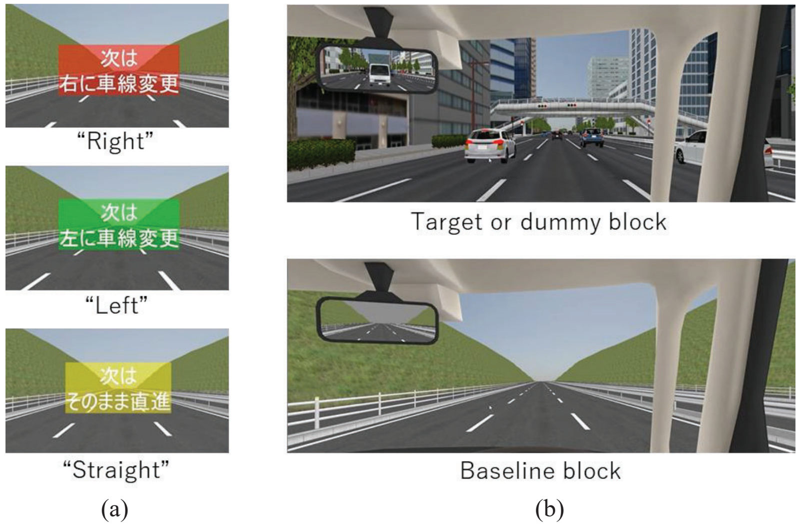


Figure 2. Contents used in the MRI experiment. (a) indication in the preparing part, (b) scenes shown in the driving part.

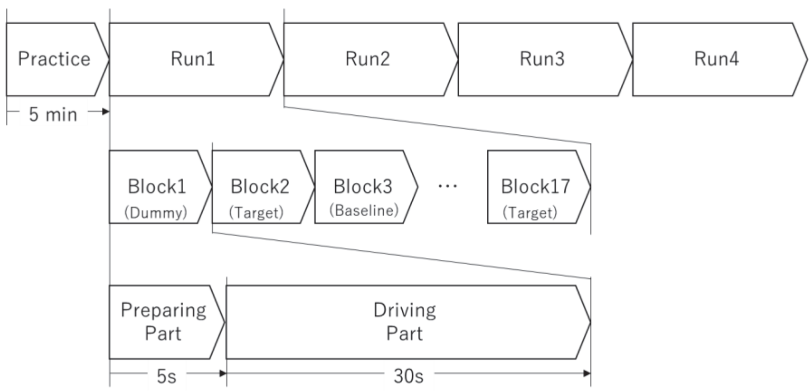


Figure 3. Schema of the MRI experiment protocol showing the timing and duration.

The MRI system was used to obtain T2*-weighted echo-planar imaging (repetition time [TR] = 1000 ms, echo time [TE] = 30 ms, slice number = 42 (interleaved), slice thickness = 4.0 mm, matrixes = 64 × 64, 600 volumes) with bold oxygen level-dependent (BOLD) contrasts and structural T1 image (TR = 2300 ms, TE = 1.98 ms, slice number = 176, slice thickness = 1 mm, matrixes = 512 × 512)

Participants

Fourteen healthy participants (mean age ± standard deviation, 37 ± 6 years) in this study were recruited from public and HONDA’s test drivers. All participants had no history of neurological and psychiatric disorders and were not taking any medication, had a driver’s license and were driving on a daily basis. Prior to the start of the experiment, all participants received an explanation of the contents and the risk of the experiment, their rights, and voluntarily signed a participation agreement. This study was approved by the Ethical Committee of HONDA MOTOR CO., LTD., Japan and QST.

Analysis

The participants were divided into two groups based on their performance on the maximum magnitude of deceleration for stopping at the traffic light in the target blocks. The participants, who decelerated more moderately than the average, were assigned to the gentle braking group, and all others were assigned to the hard braking group. The mean maximum deceleration of the gentle braking group (mean age ± standard deviation, 37 ± 6 years) was 2.5 ± 0.5 m/s^2 (mean ± standard deviation), and 4.0 ± 0.4 m/s^2 for the hard braking group (37 ± 8 years).

We used SPM12 (The Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, UK) software implemented in MATLAB (R2018b, Mathworks, USA) for preprocessing and statistical parametric mapping analysis. The preprocessing pipeline included realignment of the functional volumes, spatial normalization of the images to the Montreal Neurological Institute (MNI) space, and spatial smoothing using a 3D 8 mm full width at half maximum Gaussian kernel. We performed subject-level denoising using a regression model with the six motion parameters to eliminate head motion artifacts from the time course of the blood oxygenation level-dependent (BOLD) signal.

To extract cognitive process associated with attention for risk, we focused on brain activation at the moment to gaze the vehicle closing behind via right sideview mirror was adopted as the behavioral feature associated with attention for risk.

Result & Discussion

Figure 4 shows the differences between the gentle and hard braking groups in brain activity at the moment of gazing at risky object at the thresholds of uncorrected $p < .001$ at voxel level. The gentle braking group showed greater activations in the precuneus, SMA, and V1 than in the hard braking group. Several reports have shown that the precuneus is associated with cognitive functions such as episodic memory, visual imaging [8], and spatial judgement [9]. We speculate that these functions may play a role in linking cognition activities such as noticing risk objects, with motor execution, such as avoiding risky objects with smooth braking.

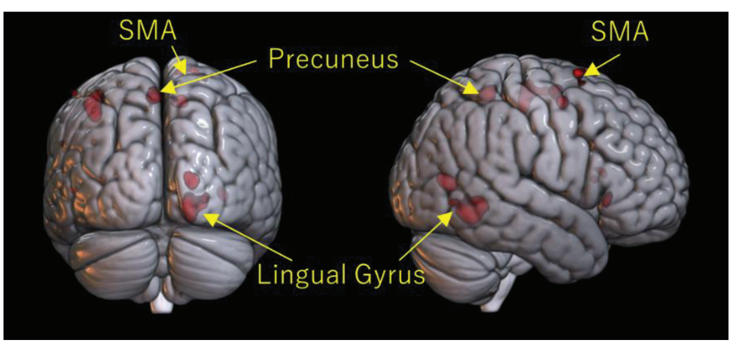


Figure 4. Significant brain activation on the contrast of the gentle braking group > the hard braking group at the threshold of uncorrected p < .001 at the voxel level.

A SUGGENSTION OF HMI

According to this neuroscientific finding, we considered a specification of a head up display (HUD) as Human Machine Interface (HMI) for supporting the spatial cognition for early noticing then urging to avoid some risky objects as shown in Figure 5. Its outline was designed to assist the spatial cognition that was one of the precuneus' main function by showing overhead view centered on ego-car. Not to distract the participants' perception, ego-car and other running objects were represented by simple symbolic icons.

To establish whether this HMI's effectiveness that had a driver notice and then avoid risky objects on a road, the validation experiment was conducted by using a driving simulator. The experimental HUD (size: 116 mm x 49 mm) was created by the software (UC-win/Road) controlling the simulator and drawn on an area where a mass-produced HUD would be projected on an actual vehicle.

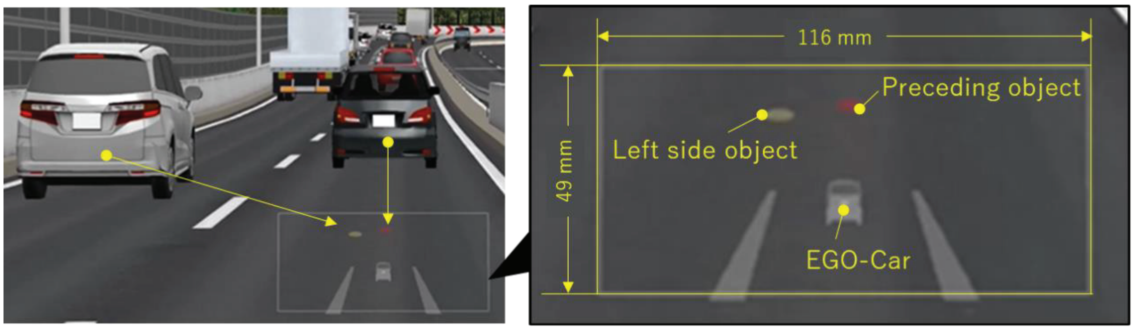


Figure 5. The experimental HUD drawn on the front display of the driving simulator.

HMI VERIFICATION EXPERIMENT

Devices

The driving simulator shown in Figure 6 used in this experiment to validate the suggested HMI was constructed originally based on a front half of an actual vehicle (N-BOX, Honda) with the driving simulator for training for safety driving [10] and the eye tracking system (Seeing Machines).

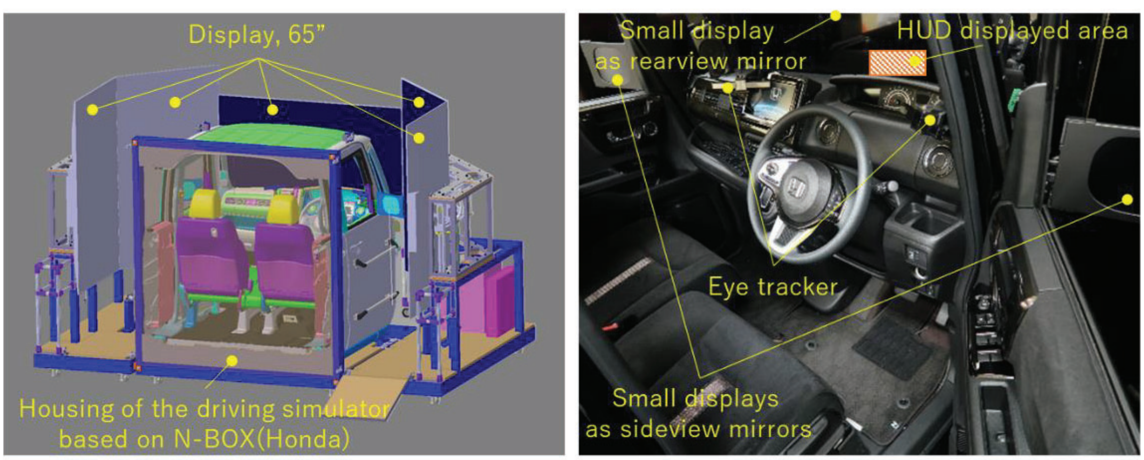


Figure 6. Overview of the prototyped driving simulator used in this experiment

Procedure

The Runs executed in this experiment were made by UC-win/Road. A run had 20 blocks, and each composed of the preparing part that took about 3 seconds and the driving part that took about 30 seconds. In the preparing part, two messages were set for, one was specifying direction of driving lane and the other was described “Follow the voice guidance” and either was displayed in the preparing part. In the case of former, an instruction for changing lane such as “Stay the lane” or “Change lane to right” was shown. In the other cases, a voice guidance message was spoken in the middle of the driving part such like “please stay the lane”. In any cases, participant would be asked to follow them. In the driving part, participant drove at 65 km/h as initial velocity in some scenes which were modeled after some parts of the Metropolitan Expressway in Japan.

In the target blocks, the vehicle overtaking on left lane intended to cut in the front of the car ahead was set as the risky object to notice and the car ahead was set as the risky to avoid.

(a) (b)

Figure 7. Contents used in the experiment for validation of the suggested HMI.

(a) indication in the preparing part, (b) sample of scenes shown in the driving part.

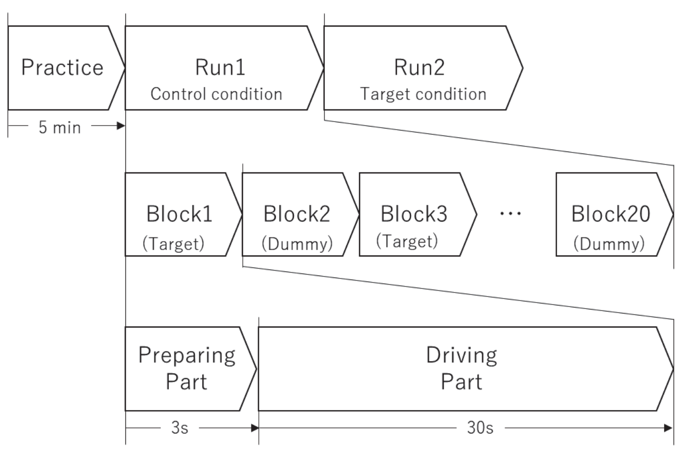


Figure 8. Schema of the experiment protocol for verification of HMI showing the timing and duration.

A run had 7 target blocks and 13 dummy blocks shown on the display randomly and it took about 12 minutes.

Traffic objects were shown with slight difference (body color, velocity, type of car, etc.) in each block. Participants asked to drive with no assistant tools as the control condition, then with the HMI as the target condition. In order to familiarize participants with this simulator, they were asked to drive practically for about five minutes before the experiment began.

Participants

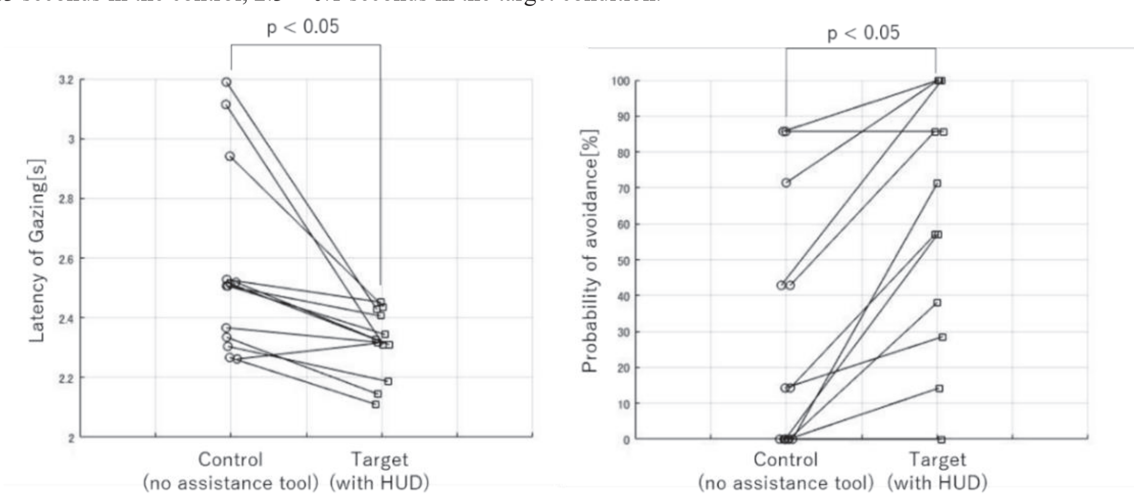
In this study, 16 healthy participants were recruited from public apart from MRI experiment, consisted of 4 from each 20's, 50's, 60's and 2 from each 20's, 40's, and the gender ratio was even. 13 individuals were cited for the analysis because of their gaze data availability. Prior to the start of the experiment, all participants received an explanation of the contents and the risk of the experiment, their rights, and voluntarily signed a participation agreement. This study was approved by the Ethical Committee of HONDA MOTOR CO., LTD., Japan.

Analysis

The performance of the early noticing the risky was measured by a time from the moment when the left running vehicle that would be the risky on side by side to when the driver's gazed the vehicle. The index of avoidance of the risky was set as an appearance probability of behaviors such as releasing the acceleration pedal or putting on the braking pedal, to avoid the risky before the ahead vehicle became too close. Wilcoxon signed-rank test was used to compare these values between the control condition and target condition. Data management and analysis were performed using the Statistics and Machine Learning Toolbox of MATLAB (2018b).

Result & Discussion

Figure 9 (a) shows a data of the average of latency gazing risky object, the median ± standard deviation were 2.5 ± 0.3 seconds in the control, 2.3 ± 0.1 seconds in the target condition.



(a) (b)

Figure 9. Results of verification on HMI suggested in this study

The difference between two conditions was significant ($p < .05$). The 12 participants out of 13 declined the latency, we thought that the suggested HMI could have effect on early noticing risky objects. Figure 9 (b) indicates the probability of avoidance, 14.3 ± 34.2 % in the control and 56.8 ± 37.5 % in the target condition then there was a significant difference ($p < .05$). This result implies the HMI could have effect on avoiding risky objects. But three participants would not increase the probability unfortunately, especially the two were zero to zero despite of their latency improved. It seems that the two participants could notice the risky but could not come up with what action they did next by using information that displayed on the HUD.

LIMITATIONS

There were some limitations to this research. In particular, the MRI experiment required participants to lie down on the bed of MRI machine instead of sitting in a seat, an environment different from real-life driving. However, even with this limitation, it may be possible to examine some aspects of visual cognitive functions for traffic risk. The specification of the suggested HMI has contained only neuroscientific findings, but also other views of well-known knowledges. To evaluate only neuroscientific effects, well-targeted experiments are needed. There are some kinds of human errors, then just one of them such as "failure of noticing" was picked up in this study. For realizing our vision, other human errors should be dealt with, and further neuroscientific, psychophysics and praxeological validation would be needed for future.

CONCLUSION

This study investigated differences in brain activation between the gentle and hard braking groups at the moment of gazing at a risky object while driving. The results showed significantly increased precuneus, SMA, and V1 response in the gentle braking group (uncorrected $p < .001$). Then, an experimental HUD for supporting safe and secure driving designed according to these neuroscientific findings, especially precuneus' main function of spatial cognition, by showing overhead view centered on ego-car. The suggested HMI was validated by participants recruited publicly, the risk detection time was significantly reduced ($p < .05$) and the probability of behavior for safety was increased ($p < .05$). These results might indicate that some drivers' abilities for noticing and avoiding the risky object during driving were improved by showing the environment centered on ego-car on the suggested HMI which was made by utilizing neuroscientific findings. These findings were expected to be applied to consider and consolidate devices development for mass production.

REFERENCES

[1] Road Traffic Injuries. (2022). World Health Organization.
https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries
[2] Chen, C., Yiik, D.W., and Xuesong, W. (2017). Safety evaluation of driver cognitive failures and driving errors on right-turn filtering movement at signalized road intersections based on Fuzzy Cellular Automata (FCA) model.

Accident; Analysis and Prevention, 104(8):156-164
[3] Thomas, A. (2008). Driver Distraction: A Review of the Current State-of-Knowledge. Washington, DC: National Highway Traffic Safety Administration.
[4] Ruben, D. L., Silvana, A. M., Fernando, M. P., and Maria, F. L. (2010). Individual Differences in Driver Inattention: The Attention-Related Driving Errors Scale. *Traffic Injury Prevention*,11:142–150
[5] Klauer, S.G., Dingus, T. A., Neale, V. L., Sudweeks, J.D., and Ramsey, D.J. (2006) The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data. Washington, DC: National Highway Traffic Safety Administration.
[6] Oba, K., Hamad, K., Tanabe-Ishibashi, A., Murase, F., Hirose, M., Kawashima,R., and Sugiura M. (2022). Neural Correlates Predicting Lane-Keeping and Hazard Detection: An fMRI Study Featuring a Pedestrian-Rich Simulator Environment. *Frontiers in Human Neuroscience*, Volume 16, Article 754379
[7] Tom, A. S., Karen, K., Yuwen, H., Fred, T., Gary, N., and Simon, J.G. (2013). Brain activity during driving with distraction: an immersive fMRI study. *Frontiers in Human Neuroscience*, Volume 7, Article 53
[8] Andrea, E. C., and Michel, R. T. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129, 564-583
[9] Chadwick, J. M., Jolly, E., Amos, P. D., Hassabis, D. and Spiers, J. H. (2015). A Goal Direction Signal in the Human Entorhinal/Subicular Region. *Current Biology*, 25, 87-92
[10] Honda Begins Sales of All-new Automobile Driving Simulator Developed for Traffic Safety Education, (2010). https://global.honda/newsroom/news/2010/c100302eng.html

KEY HUMAN FACTORS DEVELOPMENT PRINCIPLES FOR DMS ENHANCED COLLISION AVOIDANCE SYSTEM DEVELOPMENT

Mikael Ljung Aust

Emma Tivesten

Pär Gustavsson

Christian Applehult

Volvo Cars

Sweden

Paper number 23-0263

ABSTRACT

Driver monitoring systems (DMS) can enhance Collision Avoidance Systems (CAS) in numerous ways, for instance by adjusting warnings or interventions when drivers are inattentive or in other ways disengaged or impaired. However, the driver interaction principles applied when using DMS to enhance CAS must be based on State-of-the-Art Human Factors research and have a clear focus on understanding driver needs and in what way assistance should be provided to be appreciated by the driver. Otherwise, one risks implementing interactions that either do not make sense or are perceived as disturbing, both of which degrade the CAS's safety potential.

Some of these interaction principles may not be fully intuitive unless your background is in behavioral psychology. For example, it may be surprising that DMS is best used to delay certain collision avoidance warnings rather than supply them earlier. It may also not be fully intuitive that DMS is best used for detection of generic degradations in the behavioral patterns that define normal driving rather than for diagnosis of specific impaired states.

To use a CAS properly, you need to interact with it regularly to learn what its outputs mean. However, current accident and mileage statistics suggest that driving conflicts where a CAS could save you from an unrecoverable error that otherwise would have resulted in a high severity crash are rare; maybe as infrequent as once in a decade or lifetime depending on how one does the calculation.

From a design perspective, CAS are therefore best approached as lifetime driving companions. You may only need them once, but they still need to be interacted with regularly to work as intended. Hence, the conversation between driver and CAS should adhere to the same principles as applied between humans. For example, if your colleague is busy, you only interrupt for good reason, and if you interrupt regularly, both of you must agree its relevant and the message must be clear (though not necessarily loud) so the other person quickly can decide whether to interrupt the current task.

In this paper, first a general framework and corresponding design approach for CAS is formulated based on accident statistics, driving mileage and CAS interaction frequency analysis. Next, three specific development principles for DMS enhanced CAS are described to illustrate what the outcome is when the framework and design approach are applied in practice. These include how DMS enhancement can be used to avoid "cry wolf" effects in CAS interactions, how DMS enhancement can be used to get CAS timing right for both distracted and aware drivers and finally, how DMS offers a more efficient way than specific state diagnosis when tackling driver impairment.

By explicitly describing these fundamental Human Factors development principles for DMS enhanced CAS to the traffic safety engineering community, one may avoid unnecessary development pitfalls that could counteract DMS enhanced CAS deployment.

DEFINING A GENERAL FRAMEWORK

To define a general framework for the interaction principles for CAS, one has to consider three key facts. The first is that even though CAS come in many forms, they all have one thing in common which is that to use them properly, the driver has to know how they work. So, as long as a driver uses the same car, s/he needs to interact with the CAS regularly to first learn what the systems do, and then remember what their outputs mean.

Second, for the CAS to achieve its intended safety benefit, drivers need to trust their inputs and actions (i.e., the warnings and interventions provided). If driver trust in, and understanding of, the CAS can be established in a good way, chances are good that drivers will cooperate with the CAS in the intended way and perceive the CAS actions as beneficial. However, if what the CAS does is perceived as unintelligible, pointless or perhaps even scary, it follows that drivers will neither use the CAS as intended by the designer, nor spend money on CAS features in their next car purchase.

Third, when one studies how often people end up in traffic accidents and compare that to the distances travelled or hours driven without crashing, one can conclude that driving conflicts where an CAS would save you from unrecoverable error are rare. For example, in 2020 the US average was 1,33 deaths in 100 million vehicle miles travelled [1]. At the same time, the average mileage per car was about 13 500 miles per year [2]. This means that there are about 5500 years of successful driving per fatal crash.

Of course, non-fatal crashes are far more common, but the key point here is that crashes, and particularly severe crashes, are very infrequent compared to all the driving we do; maybe as infrequent as once in a decade or lifetime depending on which crash type and outcome one is looking at. This means that the average driver exposure to critical situations where the safety margins are so small that severe injury or death is imminent if not handled correctly is very low.

Putting these three key facts together, it becomes clear that we cannot rely on exposure to a particular set of truly critical situations with potential high crash severity for drivers to understand and learn how a particular CAS works. This insight provides the foundation for the general framework proposed in this paper. This framework states that each CAS must be capable of providing a frequent enough driver interaction also outside of truly critical situations to provide a sustained, acceptable and trustworthy learning experience. The reason for this is that it's that experience which in turn secures an adequate driver response on those rare occasions when a truly critical situation occurs, and the CAS can provide a real safety benefit. This general framework forms the basis for the rest of the paper.

INCREASING MARGINAL TRUST - A FRAMEWORK COMPATIBLE APPROACH TOWARD CAS DESIGN

As shown in the general framework discussion above, to maintain an interaction frequency that is sufficiently high to create and maintain learning in the driver, most CAS interactions need to happen in situations that are not truly critical, since the latter are too infrequent to provide a robust exposure rate for learning.

Still, for the interaction with the CAS to make sense and provide that learning experience, there must be a clear connection between the CAS output and the type of critical situation which the CAS is meant to provide protection from or create awareness of. If that connection is unclear, it will be very hard for the driver to understand what the purpose of the CAS is.

From this follows an interesting design constraint for CAS interactions. Since most CAS are intended to act on, or make drivers aware of, critically small safety margins, the learning experience needs to happen in situations where the safety margins either are shrinking in an obvious manner toward a point where they would be considered too small (things would have gone bad if this state had been allowed to continue), or where they are sufficiently small[1] to create an intuitive learning experience once the driver becomes aware of them.

[1] Note though that they cannot be too small, since if they were, the CAS would not provide the relevant safety benefit in the situations where it is actually needed.

Moreover, since the designer of the CAS interaction will not be riding along to explain what s/he meant when shaping the system, the driver is the sole interpreter of both meaning and relevance. The connection between an CAS's outputs and the truly critical situations it intends to prevent or mitigate therefore has to be obvious to all drivers without additional explanatory material being present or read/watched in advance (note that it has been shown that at least 50% of customers learn about CAS through trial and error [3,4,5].)

Now, as many studies in pedagogy will testify, people are different and have different ways of understanding things. However, since most vehicle manufacturers make hundreds of thousands or millions of vehicles, the possibilities for tuning the shape of the CAS connection to suit individual preferences or explanatory models are limited. Also, even though more computational power and intelligent algorithms are expected to make their way into future cars, deploying CAS where each system gets tuned to each (current) driver is likely a too large and complex undertaking to warrant the effort required.

Thus, to secure an intuitive and relevant CAS interpretation in all drivers, one instead has to find approaches that speak to all drivers, regardless of who they are and where they drive. A successful CAS interaction pedagogy thus needs to be anchored in basic human factors design principles, i.e. principles that one has good reason to believe applies to all drivers in general rather than just a few, to meet with the capabilities of the technology.

Fortunately, this is an area where extensive research has been done over the years and lots of good advice and guidelines exist, both in the scientific community and in the more popular science literature, such as the highly entertaining books "The design of everyday things" [6] and "Things that make us Smart" [7]. Still, the advice on offer here covers many different applications and contexts. It is therefore worthwhile to define an approach that is specifically suited to the automotive domain in general and CAS interaction designs in particular.

Here, we propose that the concept of *Marginal Trust*, i.e. the study of how trust is built or lost based on the acquisition of new information or input [8] provides the approach we need to design CAS interactions. As described above, the success of a CAS in preventing or mitigating a crash depends on the meaning of its outputs being intuitive (what does it want with me?) and perceived as relevant for driving (does that make sense?).

Since every interaction a driver has with a CAS provides information on what the CAS intends to achieve, following the model in [8] one can say that these two questions are revisited in each such interaction, which in turn influences marginal trust in the CAS (i.e. it either increases, decreases or remains unchanged). Aiming to design CAS such that each interaction has a positive influence on marginal trust, or not a detrimental one if trust already is high, can therefore be stated as main design goal if the full safety potential of an CAS is to be realized.

Now, various types of trust have been identified in the literature [9,10]. Often, one makes a distinction between cognition- and affect-based trust [11]. Cognition-based trust is based on a rational evaluation of competence, responsibility, and dependability [10] and rests on logical and rational calculation of likely behavior and outcomes of future collaboration. Affect-based trust on the other hand happens when emotional bonds are created, and this bond can work as a replacement or surrogate for cognition-based trust, enabling people to take a "leap of faith" that trust will be honored [12].

Affective trust is based on beliefs that the one you place your trust in cares about your welfare, will act positively towards it and take care to avoid harming it. However, while clearly a strong influencing factor in the build of marginal trust, affective trust is rather hard to exploit through system interaction design. In the automotive domain, one might rather say that this type of trust instead should be built through brand communication and how one describes the stance of the company towards its customers in various contexts.

Cognition-based trust however fits the bill for CAS design. If viewing an CAS as a lifetime driving companion which you may only really need once but still need to interact with regularly, then the conversation between driver and CAS should strengthen the driver's view of the CAS as a competent, responsible, and dependable agent. This is very similar to conversations between humans. For example, if your colleague is busy (or in this case, if the driver is busy driving) you should only interrupt for good reason, and if you interrupt regularly, both must agree that its relevant, the message must be clear, etc.

Approaching CAS interactions as a design space where one needs to leverage cognition-based trust mechanisms to positively influence marginal trust thus seems to be a both viable and promising way forward. This brings us to the next question: how can DMS facilitate the design of CAS?

SPECIFIC DEVELOPMENT PRINCIPLES WHEN USING DMS TO ENHANCE CAS

So, how can one make use of DMS inputs to when one needs to leverage cognition-based trust mechanisms to positively influence marginal trust in CAS interactions? In the following, three specific principles will be described that illustrate how DMS can be used to support the designers in this endeavor. These principles are not meant to be exhaustive nor definitive; rather they serve to illustrate which knowledge must be acquired and which decisions need to be made when specifically using a DMS system to enhance CAS interactions.

Using DMS to avoid “Cry wolf” effects in CAS interactions

When applying collision warnings to alert the driver to external threats, the CAS doesn’t have to be right every time in the sense that to the driver, a perceivable external threat always must exist to match the warning and hence make it appear relevant. However, the CAS cannot be wrong most of the time either, because then it will be perceived as a system which, in the words of Aesop [13], cries wolf all the time. Or in the context of the framework and approach described above, each false warning decreases marginal trust, so you can’t have too many of those or all trust will be lost.

Some basic rules for how often you need to be right can actually be established. From a purely theoretically standpoint, one can make the argument that the system needs to be right more than half the time, otherwise the ADAS outputs literally appear stochastic in the eyes of the driver (being right every other time implies also being wrong every other time). Simply put, it has to perform better than chance.

Still, the CAS does not have to be right all the time. In an interesting study on behavioral adaptation to Lane Departure Warnings (LDW), Le Noy and Rudin-Brown [14] showed that drivers reported almost as high levels of trust in a flawed LDW system that was programmed to miss one in three true positives and also intermittently added false positives, as in a fully accurate LDW.

What the exact ratio of true to false positives should be to build sufficient driver trust in a CAS remains to be empirically determined. However, it is clear that a key aspect of CAS design is to secure a positive balance toward true positives, that means that the warning is warranted both from a situational and a driver perception perspective and avoid false positives where the warning appears to be given for no real reason.

Here, adding a DMS to the equation opens up for significant improvements of a CAS true to false positive ratio in two ways; one regarding the opportunities to be right and the other regarding the possibility to get the timing right.

To understand the being right part, we must first look at today’s systems. These are often designed around a main conflict scenario with a number of exceptions added. For example, you might design your forward collision warning system (FCW) to warn when Time-to-Collision (TTC) is less than 1.7 s, based on the assumption that drivers would not voluntarily place themselves in a situation with such a small safety margin. However, once that general rule is in place, the developers immediately start adding exceptions for situations where the safety margins can become smaller, but where they believe the driver is in control and does not need a warning. An example of the latter would be overtaking. Here, TTC values can get very low while the driver is still in full control.

Still, even if the list of exceptions and their associated detection criteria is well thought out and tested in development and through customer feedback, it is very difficult to get it right all the time. Developers therefore generally take a conservative approach and avoid letting warnings through when they are unsure about whether the driver really needs the warning, an approach that potentially leads to under- or disuse of the CAS.

By adding a DMS to the vehicle, one can replace that exception list with real time analysis of where the driver’s attention actually is directed when a conflict arises. In a first step, the system can determine whether the driver is looking in the direction of the threat at all. In a second step (though a bit more challenging on the detection side), the system can determine whether someone with his or her visual attention in the right place also is ready to act.

Adding these enhancements to the situation analysis can provide a huge step forward in terms of equipping CAS designers with the confidence they need to let their system warn the driver. They no longer need to make that decision based on assumptions and predictions about whether the driver is attentive or not. Instead, they can use the driver's actual, real time, direction of attention to help arbitrate whether CAS inputs are needed or not.

This opens up for a much more forward, less conservative, approach in terms of getting the CAS dialogue right with the driver. The achieved safety benefit of these systems could thus get a significant boost, since being able to increase the true to false positive ratio would build driver trust and confidence in the CAS actions.

Using DMS to get the timing right for both distracted and aware drivers

Even if a DMS can detect where the driver's attention is directed in a given moment and hence whether a CAS input might be warranted or not, the designer still has to decide on the input timing. Conceptually speaking, a key element in drivers' judgement of situational relevance will depend on what the safety margins are when the CAS activates. For example, consider a situation where you are approaching a lead vehicle and the DMS has determined that your visual attention is directed through the side window. Also, let's say the car is set to warn drivers of a potential lead vehicle collision extremely early, e.g., when TTC is 5 seconds. At 50 kph, this translates to being about 65 meters away from a stationary lead vehicle, which means most drivers would say there is no immediate danger present. Chances are therefore quite high that you as a driver would consider this a false warning when you get it. On the other hand, if the warning is set to come extremely late, e.g., when TTC is 0.5 s, there will not be enough time to react. Today's CAS systems have been developed to strike a balance between these end points. For example, the European New Car Assessment Programme (EuroNCAP9 requires that forward collision warnings be given at TTC > 1.2 s [15].

As described above, a basic assumption when deciding on a warning timing threshold is that the driver would not be in this position voluntarily, which means that there also is an underlying assumption that driver must be distracted or unaware for some other reason. An inherent challenge for this approach is the side effect it has on drivers who have their visual attention in the right place when the warning is given. Looking at what happens a bit more in detail, the time required for visually distracted drivers to move their gaze from e.g. a secondary task to the road scene ahead is typically around 500-700 ms [16]. If we start from a warning timing threshold of say 1.7 s TTC, the critical event will thus have developed to a point where TTC is 1.0-1.2 s before a visually distracted driver is able to assess the forward road scene.

Visually attentive drivers on the other hand do not need additional time to get their eyes back on the road. They will therefore assess the scene at the same time as the warning is given, i.e. when TTC = 1.7 s. The same warning timing thus presents a considerably less critical situation to the visually attentive driver, who therefore less likely to consider the safety margins small and the CAS input relevant. The same warning timing can thus lead to a true positive perception by the visually distracted driver but a false positive perception by the visually attentive driver.

One simple resolution to this problem would be to suppress all CAS inputs for visually attentive drivers. However, that would disregard the possibility that the driver, although visually attentive, for some reason does not realize the need to act and hence still needs the CAS warning. The latter is sometimes referred to as cognitively distracted drivers, who would be labelled as 'failed to look properly' or 'looked but failed to see' cases in British crash causation analysis [17].

Here, the DMS provides a very interesting alternative, since it can be used to *delay* warnings for the visually attentive drivers. Since they don't need those 500-700 ms to get their eyes back on the road, a delay of the warning by 500-700 ms will move the perception of the situation to the same TTC value as that for the visually distracted drivers (1.0-1.2 s in this example). This means that the visually attentive driver will have as much time to respond as the visually distracted driver. At the same time, the CAS is bringing attention to a driving situation that is more critical, in the sense that safety margins are perceivably smaller when the warning is issued. So, adding a delay gives visually inattentive and 'looked but failed to see' drivers equal time to react, but also increases the chance that a driver who is both looking in the right place and aware of the need to act still will judge the CAS input as relevant and meaningful, and hence as a true positive.

Using DMS to detect deviations from the normal driving

In addition to enhancing the precision of CAS warnings and interventions, DMS systems are also expected to have an application in the understanding of when drivers are fit to drive or not. The latter can result from many well-known risk states such as the driver being intoxicated or extremely sleepy [18,19]. While previous efforts largely focused on preventive work such as alcohol interlocks or “don’t drink and drive” campaigns, many new initiatives are being brought forward to promote detection and mitigation of impaired states while driving in both the legislation [20] and the consumer rating [21] arena.

A main, and at first sight reasonable, track in this endeavor to combat impaired driving is to try to turn new and existing vehicle sensors (including DMS systems) into diagnostic tools that can be used to precisely determine impaired states, such as intoxication above a certain level or a sudden drop in blood sugar to name a few. However, under closer scrutiny, this approach has several severe challenges. There are challenges coupled to the drivers’ privacy if the car would read and store medical diagnosis information. Also, there are severe technical gaps to be closed, since precise medical diagnosis that has to rely on non-invasive sensors and be performed in a moving vehicle is very hard. Additionally, if these two challenges are overcome, the result would be a system that most likely has to go through medical certification procedures for the information to be considered valid to act on. As medical certification procedures are, for a good reasons, both rigorous and slow, going down that route would make both development and updating of these features cumbersome and costly.

Now, this problem can be simplified greatly if one reverses the perspective. Instead of aiming to precisely diagnose specific driver impairment states such as a particular alcohol intoxication level, one can instead aim to detect generic degradations in the behavioral patterns that define normal driving and make the vehicle’s CAS act on those. Put differently; knowing that the driver has left normal driving behind is actually enough to give a CAS the freedom it needs to act on an impaired driver state. It does not have to be more precise than that.

The reason why this approach works in practice is because driving is a highly practiced and overlearned skill in most drivers. To travel those average 13500 miles a year [2], one has to spend several hundreds of hours driving, which also means that we train our speed, distance and lane keeping skills for hundreds of hours per year. Furthermore, the possibilities for variability (i.e. the possibility to drive in a very different style compared to how others drive) are limited if you’re to keep the car on the road and disrupt the traffic flow or pattern. And the latter you don’t want to do, since other drivers will give you feedback (basically tell you off) for doing so.

This means that the “normal driving box”, i.e. the control parameter space within which you normally constrain yourself while driving, is 1) small and 2) very similar between drivers. Hence, if we detect significant deviations from that box, we can deduce that you’re no longer driving like people normally do, which means there is now reason to think that you are impaired for some reason.

So, if the goal is to detect significant deviations from normal driving, having a DMS onboard offers the completely new opportunity to look at gaze patterns, in addition to vehicle control patterns. And as previous research has shown [22], gaze patterns can be a very good predictor of whether a driver has mentally checked out from driving, even though s/he is still in the driver’s seat and mostly looking at the road ahead.

Aiming to detect traffic relevant driver impairments by looking for significant deviations from normal driving offers a path toward combatting impaired driving that does not require medical grade detection procedures, which are also likely is more robust in the face of the natural variability that always is associated with large populations. Also, by leveraging DMS to study gaze patterns in additions to vehicle control patterns, one has new opportunity to give the CAS systems ‘license to intervene’ in situations where warnings and/or interventions might otherwise have been suppressed, for instance where pedal or steering wheel use studied by themselves might indicate an alert and aware driver.

CONCLUSIONS

To use an CAS properly, you need to interact with it regularly to learn what its outputs mean. However, current accident and mileage statistics suggest that driving conflicts where an CAS would save you from unrecoverable error are rare; maybe as infrequent as once in a decade or lifetime depending on crash type.

CAS are therefore best approached as lifetime driving companions. You may only need them once, but they still need to be interacted with regularly to work as intended. Hence, the conversation between driver and CAS should adhere to the same principles as applied between humans.

Based on this general approach, a few specific development principles for DMS enhanced CAS can be derived to maximize the benefit one can get by adding DMS assessments of driver state and attention to the CAS threat assessment platform. By explicitly describing these fundamental Human Factors development principles for DMS enhanced CAS to the traffic safety community, the designer may avoid unnecessary development pitfalls that could counteract the deployment of these systems.

REFERENCES

[1] National Center for Statistics and Analysis. (2022, May). Early estimates of motor vehicle traffic fatalities and fatality rate by sub-categories in 2021 (Crash Stats Brief Statistical Summary. Report No. DOT HS 813 298). National Highway Traffic Safety Administration.

[2] Highway Statistics (2019) . Available from https://www.fhwa.dot.gov/policyinformation/statistics/2019/

[3] Lubkowski, S. D., Lewis, B. A., Gawron, V. J., Gaydos, T. L., Campbell, K. C., Kirkpatrick, S. A., et al. (2021). Driver trust in and training for advanced driver assistance systems in real-world driving. Transportation Research Part F: Psychology and Behaviour, 81, 540–556.

[4] Abraham, H., Reimer, B., & Mehler, B. (2018, September). Learning to use in-vehicle technologies: consumer preferences and effects on understanding. In Proceedings of the human factors and ergonomics society annual meeting, 62(1), 1589-1593.

[5] Boelhouwer, A., Van den Beukel, A. P., Van der Voort, M. C., Hottentot, C., De Wit, R. Q., & Martens, M. H. (2020). How are car buyers and car sellers currently informed about ADAS? An investigation among drivers and car sellers in the Netherlands. Transportation Research Interdisciplinary Perspectives, 4, 100103.

[6] Norman, D. (2013). The design of everyday things: Revised and expanded edition. Basic books.

[7] Norman, D. (2014). Things that make us smart: Defending human attributes in the age of the machine. Diversion Books.

[8] White, M. P., & Eiser, J. R. (2006). Marginal trust in risk managers: Building and losing trust following decisions under uncertainty. Risk analysis, 26(5), 1187-1203.

[9] Young, L. (2006). Trust: looking forward and back. Journal of Business & Industrial Marketing. 21(7), 439-445.

[10] Huang, Y., & Wilkinson, I. F. (2013). The dynamics and evolution of trust in business relationships. Industrial Marketing Management, 42(3), 455-465.

[11] McAllister, D. J. (1995). Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. Academy of management journal, 38(1), 24-59.

[12] Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. Social forces, 63(4), 967-985.

[13] Aesop. The Shepherd Boy & the Wolf. Available from https://read.gov/aesop/043.html

[14] Rudin-Brown, C. M., & Noy, Y. I. (2002). Investigation of Behavioral Adaptation to Lane Departure Warnings. Transportation Research Record, 1803(1), 30–37.

[15] European new car assessment programme (euroNCAP) assessment protocol – Safety assist. Collision avoidance - Implementation 2023. Available from: https://cdn.euroncap.com/media/75438/euro-ncap-assessment-protocol-sa-collision-avoidance-v102.pdf

[16] Aust, M. L. (2014). Effects of haptic versus visual modalities when combined with sound in forward collision warnings. In Proceedings of the Driving Simulation Conference 2014.

[17] Department for Transport (2014) Reported Road Casualties Great Britain: 2014 Annual Report. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/463043/rrcgb2014-02.pdf

[18] National Highway Traffic Safety Administration. (2020, October). Quantifying drowsy driving (Traffic Safety Facts Research Note. Report No. DOT HS 813 003).

[19] Thomas, F. D., Darrah, J., Graham, L., Berning, A., Blomberg, R., Finstad, K., Griggs, C., Crandall, M., Schulman, C. Kozar, R., Lai, J., Mohr, N., Chenoweth, J., Cunningham, K., Babu, K., Dorfman, J., Van Heukelom, J., Ehsani, J., Fell, J. & Moore, C. (2022, December). Drug prevalence among seriously or fatally injured road users (Report No. DOT HS 813 399). National Highway Traffic Safety Administration. NHTSA

[20] European Commission (2022) Vehicle General Safety Regulation. Available from: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32019R2144.

[21] European new car assessment programme (euroNCAP) assessment protocol – Safety assist. Safe Driving – Implementation 2023. Available from: https://cdn.euroncap.com/media/75554/euro-ncap-assessment-protocol-sa-safe-driving-v1011.pdf

[22] Tivesten, E., Victor, T. W., Gustavsson, P., Johansson, J., & Ljung Aust, M. (2019). Out-of-the-loop crash prediction: the automation expectation mismatch (AEM) algorithm. IET Intelligent Transport Systems, 13(8), 1231-1240.

DRIVER ALCOHOL DETECTION SYSTEM FOR SAFETY (DADSS) – A VEHICLE SAFETY TECHNOLOGY APPROACH TO REDUCING ALCOHOL-IMPAIRED DRIVING – A STATUS UPDATE.

Abdullatif K. Zaouk
Michael Willis
KEA Technologies, Inc.
U.S.A.

Eric Traube
National Highway Traffic Safety Administration
U.S.A.

Robert Strassburger
Automotive Coalition for Traffic Safety
U.S.A.

Susan Ferguson
U.S.A.

Paper Number 23-0287

ABSTRACT

Alcohol-impaired driving continues to take a significant toll among road users both in the United States and around the world. In 2021, an estimated 42,915 people died in motor vehicle traffic crashes, a 10.5% increase from 2020. The projection is the highest number of fatalities since 2005 and the largest annual percentage increase in the Fatality Analysis Reporting System's (FARS) history. In 2020, in the U.S. alone, motor vehicle fatalities from crashes involving alcohol totaled 11,654, a 14% increase over 2019, which accounts for approximately 30% of all traffic fatalities in the US for the year. To better address this ongoing problem, in 2008 the National Highway Traffic Safety Administration (NHTSA) and the Automotive Coalition for Traffic Safety (ACTS) formed a cooperative research partnership to explore the feasibility, the potential benefits of, and the public policy challenges associated with the widespread use of non-invasive technologies to prevent alcohol-impaired driving. This partnership, known as the Driver Alcohol Detection System for Safety (DADSS) Program has made great strides forward in the development of in-vehicle technologies that will measure blood or breath alcohol and may prevent alcohol-impaired drivers from driving their vehicles. Exploratory research in earlier phases of the program established the feasibility of two sensor approaches, breath- and touch-based, for in-vehicle use. The sensors have since been refined, in terms of both hardware and software, as the program strives to meet the performance specifications required for unobtrusive and reliable alcohol measurement.

In late 2021 the program announced that the first zero-tolerance breath alcohol sensor product equipped with new alcohol detection technology will be available for open licensing in commercial vehicles. "Open licensing" means that the technology, which measures a driver's breath alcohol concentration, will be made available to any product integrator for preparation into fleet vehicles. The breath sensor is designed for fleet operators implementing a zero-tolerance alcohol policy for their drivers, staff or employees. It requires a directed puff of breath and provides a "pass/fail" reading of the driver's breath alcohol concentration.

Currently the DADSS program is focused on transitioning the latest generations of consumer breath and touch sensors from research to product development. Numerous parallel research programs continue including sensor development, development of calibration processes, materials and instrumentation that will verify the technologies are meeting these elevated performance specifications, human subject testing in conditions that replicate those likely to be experienced in the real world, and real-world field trials in diverse settings. The goal for DADSS technologies is commercialization. This paper will outline the technological approaches and the status of the various DADSS research programs.

INTRODUCTION

Alcohol-impaired driving (defined as driving at or above the legal per se limit of 0.08 grams per deciliter (g/dL) or 0.08 percent in all U.S. States except in Utah where the limit is 0.05 g/dl) is one of the primary factors in motor vehicle fatalities on U.S. roads every year. Although strong laws and enforcement have led to fewer alcohol-impaired deaths on the roadways (Ferguson, 2012), in 2020 alone, crashes involving at least one driver with a blood alcohol concentration (BAC) of 0.08 g/dl resulted in 11,654 deaths of U.S. road users (National Highway Traffic Safety Administration (NHTSA), 2020). In 2008, the NHTSA and the Automotive Coalition for Traffic Safety (ACTS)[1] began research to develop in-vehicle solutions to address alcohol-impaired driving. The alcohol sensors under development are required to be seamless with the driving task, that is, passive, accurate, fast, reliable, durable, and require little or no maintenance. Ultimately, DADSS technology seeks to restrict the motive power when the device registers that the driver's blood alcohol concentration (BAC) exceeds the legal per se limit, although other limits could be programmable. This cooperative research partnership, known as the Driver Alcohol Detection System for Safety (DADSS) Program, has been developing both breath-based and touch-based non-invasive technologies that will be able to prevent alcohol-impaired driving, (Ferguson et al., 2009, Ferguson et al., 2010, Ferguson et al., 2011, Zaouk et al., 2015, Zaouk et al., 2017).

To effectively measure blood and breath alcohol in real time with negligible misclassification errors, stringent performance specifications have been developed that provide a template to guide the overall research effort. The ability to calibrate the performance of each generation of sensor prototypes is a critical component of the development process. In addition, SAE led the development of the SAE J3214 standard, Breath-Based Alcohol Detection System, finalized in January 2021, which was specifically developed to provide the testing specifications adopted for breath sensors in fleet vehicles. The standard has been published by SAE International and is available through the sae.org website. Additional standards are in development, as outlined later in this paper. These elevated standards, especially those for accuracy and precision, have necessitated the development of innovative approaches that will enable measurement of the technologies' performance on an ongoing basis. Specifically, calibration processes, materials, methodologies and instrumentation have been the subject of extensive cutting-edge research to enable the requisite testing.

Research vehicles have been equipped for Field Operational Testing (FOT) with the latest versions of the breath sensors seamlessly integrated within the vehicle interiors. Instrumentation packages also have been developed that will provide a myriad of data on sensor performance under challenging real-world driving conditions. Along with determining whether the DADSS sensors are working as anticipated, the FOT data collection effort will allow the identification of areas for system improvement.

A comprehensive program of human subject research is being carried out, starting with the laboratory environment where better control of conditions can be exerted, and in the vehicle where the sensors can be tested in the environment in which they will be used. This research aims to establish that alcohol measurements made with diluted breath as well as capilary blood and interstitial fluid (in the tissue of the finger and thenar region of the hand) exceed or are comparable to the well-accepted standards of venous blood and deep-lung air widely used in traffic law enforcement.

At the same time, media coverage and consumer sentiments are being monitored in anticipation of a future launch of the technology. Consumer acceptance is an ongoing consideration and is a critical element and factor in realizing full implementation.

The purpose of this paper is to provide a status update on the following key DADSS program areas:

- Performance Specification Development
- DADSS Sensor Development and Subsystem Technological Research
    - Breath Sensor
    - Touch Sensor
- DADSS Verification and Validation
    - Laboratory - Standard Calibration Devices (SCD)
    - Human Subject Laboratory Testing (HST)
        - In-House Bood Ethanol Analysis

---

[1] Members of ACTS comprise motor vehicle manufacturers representing approximately 99 percent of light vehicle sales in the U.S.

- Field Testing
  - Controlled Human Subjects On-Road Driving Tests
  - Naturalistic Human Subjects On-Road Driving Tests
    - James River Transportaion
    - Schneider National

PROGRAM PROGRESS

Performance Specification Development, SAE J3214, and Future Standards

The purpose of the DADSS Performance Specifications document is to establish the DADSS Subsystem Performance Specifications for passenger motor vehicles. In addition to specifications that detail the sensor's speed of measurement, accuracy, and precision, reliability specifications have been identified that conform to the automobile industry accepted level of reliability, thus minimizing the potential for system failure. International Organization for Standardization (ISO) standards, provided in the DADSS Performance Specifications, are also followed to ensure that materials, products, and processes are acceptable for their purpose.

In addition to the DADSS Performance Specifications, performance specifications for the zero-tolerance breath sensor (referred to as Gen 3.3) were initiated in October 2019. This device is intended for use-in motor vehicle fleet applications and will determine if the driver is registering any breath alcohol, otherwise known as zero-tolerance. The draft specifications define the accessories' technology performance as it relates to accuracy, precision, speed of measurement, influence of the environment, issues related to user acceptance (such as instructions for use), long-term reliability, and system maintenance requirements. Access to the data memory or the ability to set operational parameters, including the setting of Breath Alcohol Concentration (BrAC) thresholds will be designed to deter unauthorized or inadvertent tampering.

SAE led the development of the SAE J3214 standard which was specifically developed to provide the testing specifications adopted for fleet vehicle breath sensors and was approved in January 2021. SAE International has published the standard and it is available through the sae.org website. The DADSS laboratory received ISO17025 accreditation to the SAE J3214 standard in September 2021 and is currently the only laboratory accredited to the standard. Unlike alcohol ignition interlocks, this fleet device operates without a mouthpiece and measures diluted breath samples. However, the SAE J3214 standard is applicable to systems with and without mouthpieces. The fleet device is designed to meet international specifications and standards for alcohol measurement devices currently in place in the United States, Canada, and Europe, but has more stringent requirements, especially with respect to the calibration curve and test gases. Since its release there have been several changes to improve the document structure and flow, as well as better define the testing requirements, including improvement of requirements for electrostatic discharge, electromagnetic compatibility, and interference. The updated version is expected to be released in the fourth quarter of 2022.

In addition, an SAE working committee has been established to create SAE standards for passive breath alcohol systems in consumer vehicles, and touch-based capillary blood alcohol measurement in-vehicle systems. This committee published the SAE J3214 (Breath Alcohol Detection System Standard) in January of 2021. It is anticipated that the passive breath standard should be released by the second half of 2023, and the touch-based standard by the second half of 2024.

DADSS Sensor Development & Subsystems Technological Research

The two technologies, breath and touch are being pursued for measuring driver BrAC and Blood Alcohol Content (BAC) non-invasively within the vehicle. Progress has been made in the development of both technologies. Two different devices are being pursued for each technology for use in vehicles – the Gen 3.3 breath sensor and Aglow touch sensor devices designed to prevent the vehicle from being driven if any alcohol is detected (<= 0.02 g/dL), and passenger vehicles devices, Gen 4.0 breath sensor and Radiant touch sensor, designed to prevent the vehicle from moving if the driver is at or above the legal per se limit for alcohol (typically 0.08 g/dL in the U.S.). Table 1 shows the target dates by which the DADSS sensors will be ready to license to product integrators. The length of time required to integrate a DADSS sensor into a motor vehicle will vary depending upon the type of product and the length of time needed to conduct validation and verification testing at the vehicle level.

Table 1. Breath and Touch Sensors Derivative.

CHARACTERISTIC	GEN 3.3 Breath	AGLOW Touch	GEN 4.0 Breath	RADIANT Touch
Estimated Commercialization*	**2021**	**2023**	**2024**	**2025**
Market Application	Fleet vehicles & accessory sales		Consumer vehicles	
Vehicle Integration	After mass production (Upfitter or dealer installed)		During mass production	
Alcohol (Ethanol) Set Point	0.02%		0.05 or 0.08%	
Operating Characteristics	Contactless, Directed–breath	Passive operation, up to 4 tunable lasers, single board electronics	Contactless, Passive–breath	Passive operation, up to 2 widely tunable lasers, ASIC–level electronics

** Gen 3.3 Breath Sensor Reference Design released for open licensing for use in commercial vehicle in December 2021*

Breath Sensor

The sensor technology under development by Senseair and its partners uses infrared (IR) spectroscopy, which is stable over the full product lifetime, eliminating the need for recurrent calibrations. The challenge in measuring breath alcohol within the vehicle cabin is that the breath is diluted with the cabin air. The breath-based approach uses sensors to measure the concentration of alcohol and carbon dioxide (CO_2) in diluted breath simultaneously. The use of CO_2 in human breath as a tracer chemical allows it to be used as an indicator for the degree of breath dilution, and thus the dilution of the alcohol concentration in the expired breath. A fan draws diluted breath into a chamber where a detector measures the interaction with the alcohol and CO_2 in the sample (Hök et al., 2006). BrAC is then quickly and accurately calculated (Figure 1).

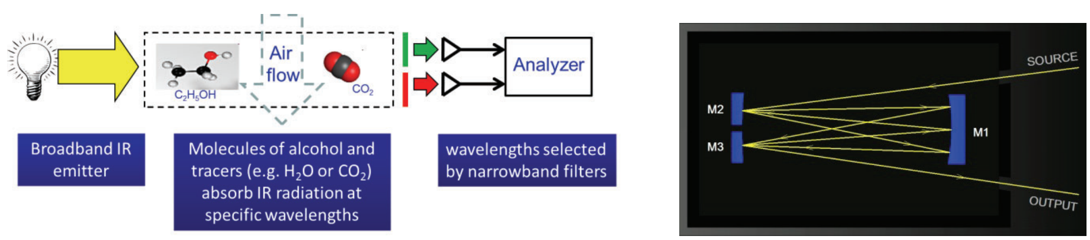


Figure 1. Breath-based sensor block diagram

The goal of the DADSS sensors is to passively measure breath alcohol within the vehicle cabin without directed input from the driver. The challenge is to meet the stringent performance requirements while measuring highly diluted breath. Thus, a significant component of the research is focused on understanding the behavior and flow patterns of the expired breath plume within the vehicle cabin and identifying effective locations for the sensors.

The breath-based sensor was updated with the goal of improving the ability to accurately and passively detect breath alcohol levels. The latest, 3rd Generation version (currently Gen 3.3), underwent a complete re-design to increase sensitivity for measurements of passive samples through improvements in the overall signal-to-noise ratio (SNR), reduce the overall size, and improve performance over the full temperature range of -40°C to +85°C as specified by the DADSS Performance Specifications (Biondo et. al. 2017). A major improvement of the Gen 3.3 sensor is the optical module configuration. Major enhancements were undertaken during the Gen 3.3 sensor development to improve how the sensor detects alcohol. Ethanol detection takes place over the full length of the cavity, whereas CO_2 is detected cross-wise to eliminate systematic timing differences between the two signals. This enables the possibility of passive in-vehicle sensing (Ljungblad, 2017). The Gen 3.3 device has been developed for fleet and accessory application with knowledge gained from Gen 3.2 laboratory studies and human subject trials. There are three product versions available: a vehicle–integrated solution, an aftermarket or accessory solution and a stationary point–of–access solution. This fleet device will be set to detect the presence of alcohol but will also have the flexibility to set the limit up to a BrAC of 0.04 g/dL depending on the company fleet owner's preference. The Gen 4.0 sensor will be suited for wider deployment in passenger vehicles (Figure 2) for a graphic representation of the sensor evolution and SNR improvements. The detector and sensor fan were modified to allow more homogeneous airflow through the system. This resulted in improved sensitivity and increased peak gas levels when measuring breath exhalations at the same distance. This is a critical step for passive breath measurement. Software algorithms for passive detection of breath alcohol levels also have been enhanced, whereby several consecutive signal features can be accumulated to provide sufficient data for reliable measurement.

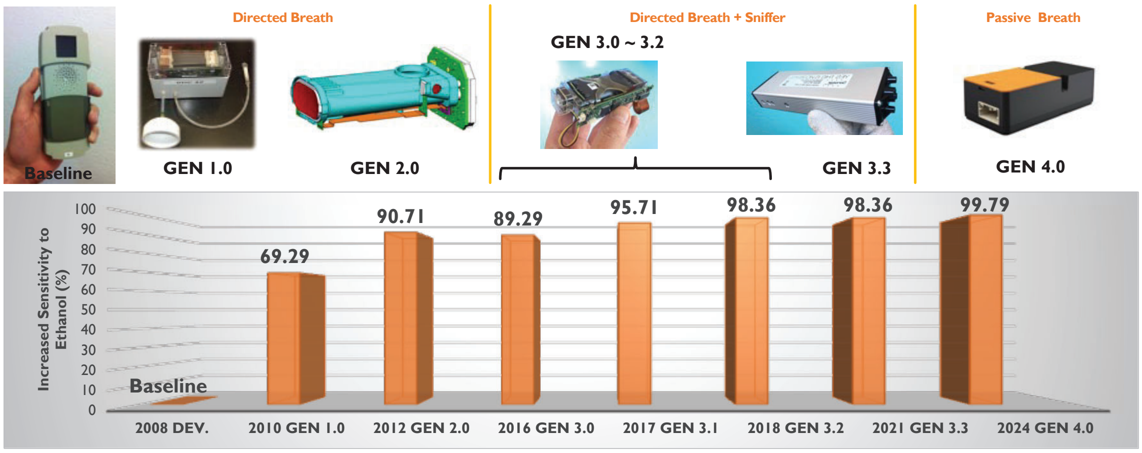


Figure 2. Breath-based DADSS Sensor Evolution and SNR improvements

Further investigations of critical components, including detectors, emitters and mirrors, have identified noteworthy options for more production friendly choices which are intended to be integrated in Gen 4.0. The latest generation 3.3 sensors have undergone rigorous operational and environmental testing aimed to simulate a sensor lifetime of fifteen years.

Touch Sensor

The touch sensor uses near-infrared (NIR) spectrometry - a noninvasive approach that utilizes the near infrared region of the electromagnetic spectrum (from about 0.7 μm to 2.5 μm) to measure substances of interest in bodily tissue (Ferguson et. al., 2010, Ridder et al., 2005). It has been determined that the 1.25-2.5 μm portion of the spectrum provides the highest sensitivity and selectivity for alcohol measurement because the alcohol signal is hundreds of times stronger than the signal in the 0.7-1.25 μm part of the NIR.

As depicted in Figure 3, the measurement begins by illuminating the user's skin with NIR light (like a low power flashlight). The light propagates into the tissue (the skin must be in contact with the device) and a portion of it is diffusely reflected back to the skin's surface where it is collected by an optical touch pad. The light contains information on the unique chemical information and tissue structure of the user. This light is analyzed to determine the alcohol concentration. The challenge is to measure the concentration of alcohol (sensitivity) while ignoring all the other interfering analytes or signals within the skin (selectivity).

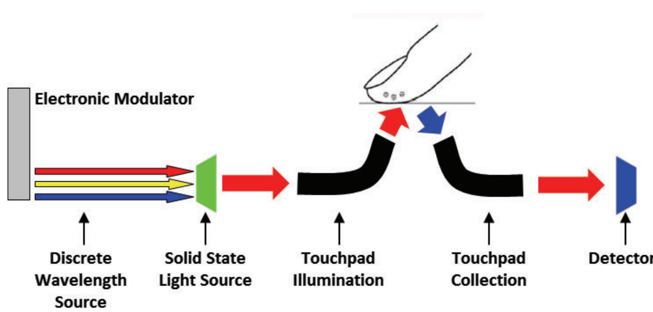


Figure 3. Touch-based subsystem solid-state laser spectrometer approach

The shift from the Proof–of–Concept prototype, which used a traditional Michelson interferometer that utilizes moving parts, to a solid-state laser spectrometer (which is better suited to the automotive environment) has required extensive hardware and software research (Ver Steeg et al., 2017). The key to enabling such innovation is the ability to define an optimized subset of optical wavelengths which will provide the high quality non-invasive alcohol measurement in human tissue. Laser diodes that are tuned for optimal alcohol measurements are used to generate 40 unique wavelengths of light. The laser diode specifications were derived from the comparison and analysis of human subject data and comparative reference data.

Extensive research has been undertaken to develop the requisite laser diodes, many of which have not been previously manufactured, and assemble them in multi-laser packages. The individual laser signals are combined into a broader, diffuse light source in the optical module, which illuminates the finger or palm, and is reflected back to the touch pad's detector, where alcohol measurements are made. After initial work was completed to develop the laser diodes and packaging, a new supplier was selected with greater expertise in these areas. Each stage of the development process has required research and has resulted in multiple patent applications.

Recently, tunable lasers have been developed that are suitable for the touch sensor. Tunable lasers can alter the wavelength of operation in a controlled manner, thus enabling the use of fewer lasers to interrogate the NIR spectrum of interest. This development is expected to have higher measurement sensitivity and perform faster than the prior laser packages. It will enable a smaller sensor footprint, use less power, have better temperature control to prevent measurement drift, and result in simplified optics and electronics. Ultimately, the goal is to use only one to two tunable laser chips to produce these same unique wavelengths. depicts the recent evolution of laser diode development (Figure 4).

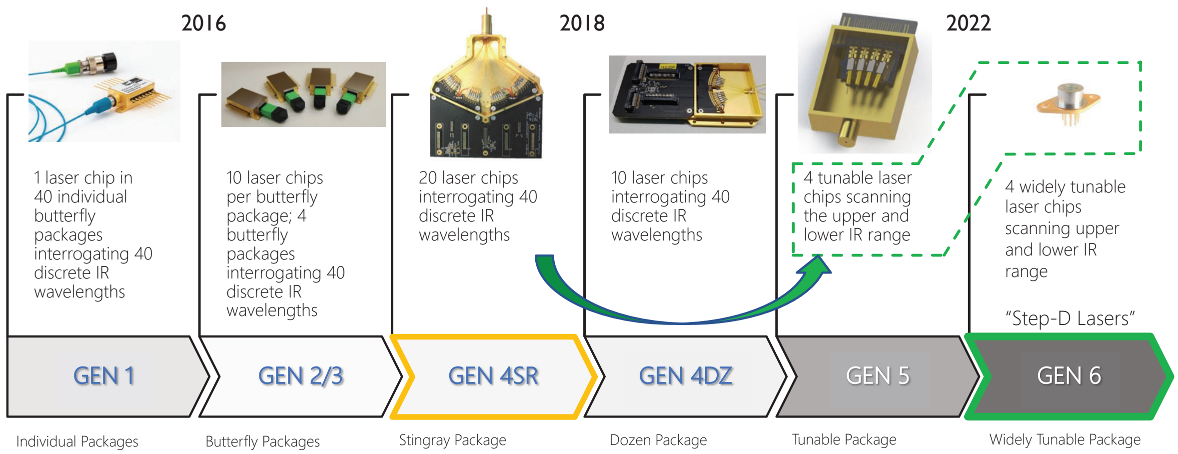


Figure 4. Evolution of laser development.

The touch sensor consists of the laser diodes, the laser guiding system to relay the laser signal into the skin in the prescribed fashion for optimal measurement, the detectors to receive the reflected signal (all of which reside in the driver optical interface), a reference sensor, and the electronics board that controls and guides the system. Each of these design elements will undergo significant enhancements from the current benchtop device. The Aglow sensor availability, suitable for fleet and accessory applications, is targeted for 2023. The consumer version, Radiant sensor, for use in privately-operated vehicles is targeted to be available during 2025.

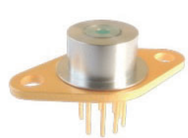


Figure 5. TO66 Tunable Near Infrared Laser

Research and development activity on the touch sensor is currently focused on optimizing the hardware and software for the benchtop optical system used in the laboratory and the development of a compact benchtop systems for field trials. All systems leverage tunable NIR lasers in a TO66 package size (Figure 5) and are primarily designed for in-vivo analysis, more specifically, the illumination and capture of reflected light from the dermal tissue on the thenar palm (Figure 6). The thenar has been selected for initial measurements due to consistency in skin tissue thickness and reduced light scatter compared to the pad of the finger.

The touch sensor functional characteristics critical to accurate and repeatable performance are 1) the laser signal needs to be stable, not drifting or fluctuating; 2) the combined light source from the lasers needs to be homogenized so that the light levels propagating through the tissue are always the same; and 3) levels of background noise need to be low and signal strength sufficiently strong so that the signal can be readily detected when reflected from the tissue.

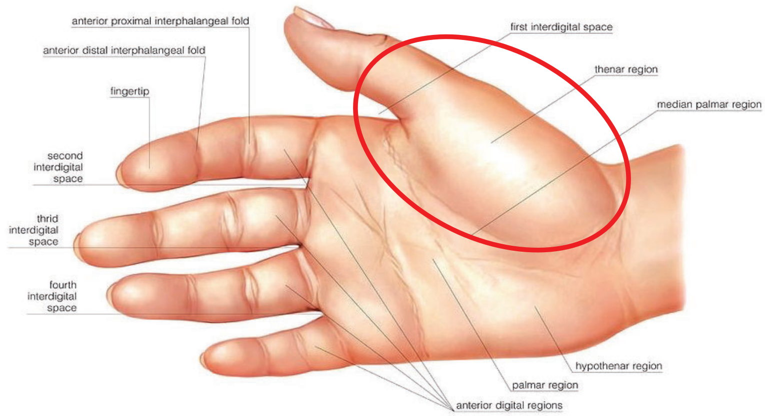


Figure 6. Thenar Palm

As with any innovative technology development, technical difficulties have been experienced along the way. It is to be expected that with each new generation of technology there is a learning curve. Similarly, with the touch system, any time the light sources change, there are new challenges to be addressed. Work continued to address challenges related to sensor performance such as light illumination, assembly, alignment, and straylight.[2] Furthermore, research focused on software development, specifically, ramp profiles for controlling the lasers during measurements. The ramp profile contains the instructions for how the laser is powered in order to create the various wavelengths within its tunable band.

A single laser optical reflectance benchtop system was developed, which is a highly flexible, modular test system with increased configurability and laser light control (see Figure 7). The benchtop system provides enhanced flexibility for varying key optical parameters and allows researchers to perform a matrix-based test plan for collecting optimum system settings and key tissue variable data for improved simulation-based analyses. Utilizing the new benchtop system, testing initially focused on optical system alignment and laser beam positioning as well as tunable laser analyses (light output, temperature stability, wavelength generation). This was followed by system modifications to improve light intensity at the front-end interface window and updates to resolve sources of structured noise associated with the system design.

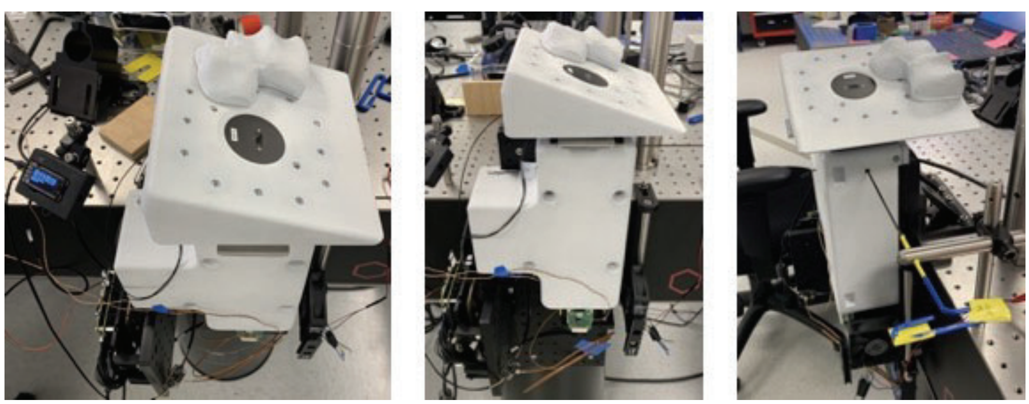


Figure 7. Single Laser Reflectance Benchtop System

The benchtop system is undergoing continuous improvement and is being used to conduct human subject testing (non-dosed and alcohol dosed) to confirm that the unit can detect ethanol non-invasively in the tissue of a dosed human. Lessons learned from this unit are currently being applied to the development of a compact portable unit with improved performance for use in vehicle application.

[2] Stray light is a broad term used to refer to any light within the optical system that cannot be used for the explicit purpose of making spectroscopic measurements.

DADSS Verification and Validation Overview
The DADSS sensors and subsystems undergo rigorous testing to insure they meet the requirements outlined in the DADSS Performance Specifications and SAE J3214 standard. The testing is separated into three categories (Figure 8):

1) Laboratory – Understanding and measuring the performance of new sensors under tightly controlled laboratory conditions.
2) Human Subject Testing – Understanding the performance of sensors in a controlled setting while introducing humans and their variability.
3) Field Testing (Human Subject Driving (HSD), Fleet) – Understanding the performance of the sensors in the real world driving environment with human, sensor, and environmental variability's.

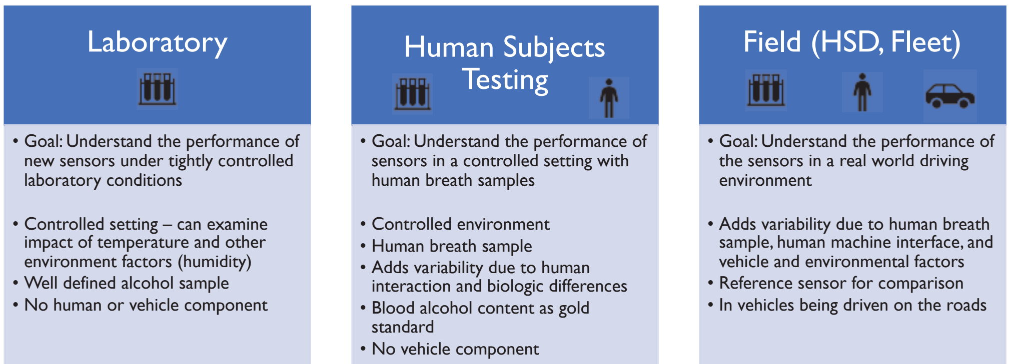


Figure 8. DADSS Verification and Validation Overview

Laboratory - Standard Calibration Devices (SCD)
As sensors evolve and improve, new generations of the breath and touch sensor systems must be evaluated to understand how well they perform. A critical component of the testing process is to develop test methods that can demonstrate in a traceable manner that the breath and touch systems meet the requisite performance specifications. As part of this process the DADSS team must develop both test methods and traceable breath and tissue surrogates for use in the testing. The traceability of these test materials comes from the use of standard reference materials (SRMs) that are produced to a known value. Typically, with the implementation of such materials, the researchers can use them to confirm that they meet the stated specifications. In the United States, such materials are usually traceable to a national standard that is held by the National Institute of Standards and Technology (NIST) or certified by another nation's national laboratory which holds a letter of agreement with NIST regarding the specific material.

To fully address the many aspects of testing, the DADSS Team has undertaken a comprehensive protocol that is based around the research, development, and vetting of apparatus and methodology. The focus of this effort is the development of breath and tissue surrogates capable of evaluating the performance of DADSS sensors against the accuracy and precision specifications. Research efforts are focused on the development of SCD's and methodologies for delivery of the samples to the verification instrumentation and the sensors for analysis. The research and development that has been conducted thus far has resulted in substantial progress, including improvements in calibration samples, measurement procedures, design, and characterization of delivery systems, as well as the characterization of the latest generation sensors.

The breath SCD has been designed to represent a naturally exhaled human breath. Parameters such as volume, pressure, humidity, temperature, and chemical composition must be specifically tailored to represent human physiological conditions. This is made possible by the development of the Alcohol Breath–Based Simulator (ABBS). ABBS, shown in Figure 9, was developed to meet these needs by combining the ethanol gas with stock diluent gases in specific ratios. The ethanol ratio is monitored in real time and automatically adjusted based on a feedback loop to adjust for variation in the ethanol gas. The design intent of ABBS is to allow flexibility in flow rate, ethanol concentration, carbon dioxide concentration, temperature, pressure, and humidity as needed to evaluate the sensors. These variable parameters allow the ABBS unit to produce a simulated human breath to the sensors with a high level of precision. Recently, a second ABBS has been developed to help improve testing throughput, with a significantly modified method to improve the output gas stream with an ideal distribution of gaseous ethanol.

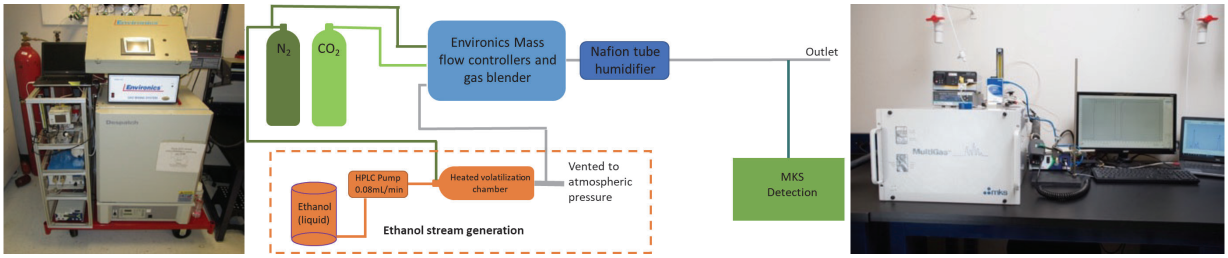


Figure 9. ABBS with the MKS Multigas 2030 FTIR

Once developed, the SRMs composition, accuracy, and precision must be confirmed based on the DADSS specifications. The instrumentation necessary for such verification must meet or exceed the DADSS performance specifications. A worldwide search was conducted for suitable technological approaches and instrumentation that could meet these goals. A comprehensive evaluation of forensic toxicology instrumentation revealed emerging technologies with improved ability to quantify and identify ethanol in SCDs. Various approaches and their methods, such as gas chromatography, liquid chromatography, and infrared spectroscopy were evaluated. A Fourier Transform Infrared Spectroscopy (FTIR) device with the MKS Multi Gas 2030 Continuous Gas Analyzer was selected for the breath samples because of its ability to identify or confirm the chemicals in the sample as well as quantify accuracy and precision at the levels required. For the SAE J3214 Standard, referred to above, there is a requirement that the test gas ethanol concentration has an uncertainty of less than 1.5 percent. In order to achieve this level of uncertainty, the gases are calibrated using the in-house FTIR device with the MKS Multi Gas 2030 Continuous Gas Analyzer using gases in which ethanol concentration is certified to a known concentration with very high accuracy. Gas standards that have enough accuracy to support this calibration are available from VSL and the National Metrology Institute of the Netherlands.

In support of touch sensor validation, work is ongoing on the development of tissue surrogates as a standard reference material, for the Touch sensor as well as delivery systems to introduce a tissue equivalent sample to the sensor. The first Touch SCD developed was as a liquid solution which had a poor shelf life. Currently, research is being conducted to transfer the desired properties of the liquid solution to a different medium, such as a gel or solid. The tissue surrogate must closely represent the properties of a human tissue, so temperature, optical properties, chemical composition, density, hydration levels, elasticity, and conductivity all must be considered. Both the aqueous base and gelatinous base have their respective advantages and challenges. The DADSS Team is working to combine these two approaches either into a hybrid system or develop a methodology which utilizes the advantages of each material.

With insight from the alcohol and environmental testing industries, new methods to improve the tissue solution's accuracy were adopted, including best techniques to weigh, portion, and quantify the ethanol when manufacturing the solutions. In addition, properties of other chemicals were used to quantify the ethanol in the solutions with extreme confidence.

For the tissue surrogate validation device, a Waters Acquity High-Performance Liquid Chromatography (HPLC) device with mass spectrometry, refractive index, and UV-Vis detectors was selected (see Figure 10). The pairing of this unit with an FTIR provides extremely precise measurement and identification of ethanol as well as the other components in the tissue surrogate.

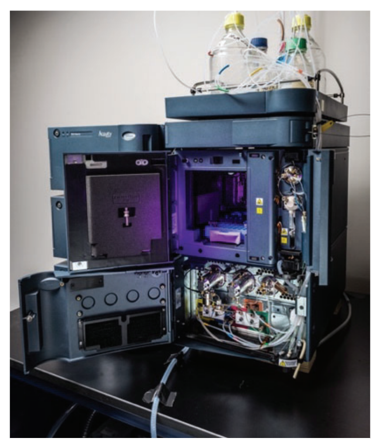


Figure 10. Waters Acquity HPLC in the DADSS research laboratory.

Human Subject Laboratory Testing (HST)

HST, also referred to as in vivo testing, is a critical part of understanding how the DADSS sensors will perform in the real world when confronted with large individual variations in the absorption, distribution, and elimination of alcohol within the human body (i.e., blood, breath, tissue) and across the many factors that can affect alcohol concentration such as age, body mass, race/ethnicity, gender, and medical conditions. Past research has provided a clear understanding of these factors with respect to venous (blood) alcohol and breath-alcohol when samples of deep lung air are used. However, the new alcohol measurement methods being developed under the DADSS program, which determine alcohol concentrations from diluted breath samples and within human tissue, are not well understood. In particular, the rate of distribution of alcohol throughout the various compartments of the body under a variety of scenarios has been the subject of ongoing study.

From the outset, a comprehensive program of human subject research has been carried out to establish that alcohol measurements made with diluted breath and tissue samples are comparable to the well-accepted standards of venous blood and deep-lung air widely used in today's alcohol detection systems. Based on an extensive review of the extant alcohol pharmacokinetics literature, intrinsic and extrinsic factors that can affect alcohol metabolism have been identified. Progress is being made in answering those questions with an ongoing, comprehensive program of human subject research being undertaken by McLean Hospital, a Harvard Medical School affiliate (Lukas et al 2019).

The purpose of human subject laboratory testing is:

- To quantify the rate of distribution of alcohol throughout the various compartments of the body (i.e., blood, breath, tissue) under a variety of real-world scenarios, and across a range of factors that could potentially affect measurement. The key question is whether these various factors have differential effects on the distribution of alcohol within the different compartments.
- To quantify alcohol absorption and elimination curves, both breath and blood, among a wide cross section of individuals of different ages, sex, body mass index and race/ ethnicity and in a variety of scenarios.

Many insights already have been gained regarding the alcohol absorption and elimination curves and maximum BACs/BrACs reached by human subjects in a variety of real-world scenarios (i.e., length of time for alcohol to appear in each compartment, effects of snacking, dining, exercise, and "last call" on alcohol measurements). These studies have confirmed a solid linear relationship between blood, directed breath (using the DADSS breath sensors), and tissue alcohol measurement (using the touch sensor) over a wide range of BACs (0.04-0.12 g/dL) (Lukas et al 2019).

To date, more than 244 clinical human subject studies have been conducted involving more than 15 protocols (see Table 2), generating more than 12,107 blood samples, 15,701 DADSS breath sensor sample, 7,627 comparative breath samples, and 8,212 DADSS touch sensor samples.

Table 2. Clinical Human Subjects Testing Study Protocols Conducted.

Protocol	No. of Studies
Exercise	17
Social Snack	15
Social Brunch	14
LagTime	49
Last Call	18
Cigarette	6
Hand Sanitizer	9
Temperature	30
Energy Drink	29
JUUL	5
Breath	10
Non-Alcohol	10
Beer	20
Wine	11
Marijuana	1
Total Studies	**244**

In-house Blood-Ethanol Analysis

To increase in-house capabilities for the evaluation of breath and touch sensors, a methodology was developed to collect whole blood samples and analyze test subjects blood alcohol levels. The subject's blood is sampled using a finger-prick, which is then processed using a YSI 2900D Biochemistry Analyzer (Figure 11). A lancet is used to pierce the skin and a blood sample is drawn with a heparinized capillary tube (i.e., to prevent blood from coagulating) after which the blood is transferred into a 200 µL vial for analysis (all three steps are illustrated in Figure 12). This low-cost method is accurate to within 3 % of actual blood alcohol levels and can deliver results within 60 seconds after blood collection. Using this method, more than 140 blood samples have been analyzed. The ethanol-containing whole blood samples collected via a finger-prick can remain stable for over 48 hours, allowing for field collection prior to laboratory analysis. Biosafety standards have been instituted in the chemistry laboratory to allow for capillary blood collection and analysis methodology.

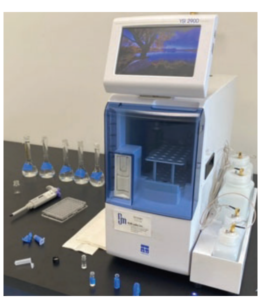


Figure 11. YSI 2900D Analyzer

The HST program continues to assist the HSD program by providing support for subject recruitment, subject safety verification (e.g., negative BrAC for all participants and negative pregnancy test for women), beverage mixing and administration and debriefing/safety testing at the end of the study

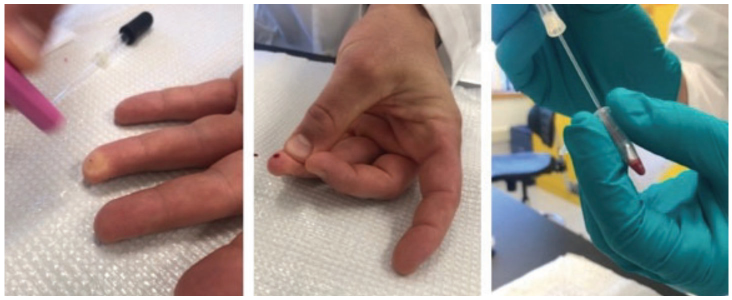


Figure 12. Blood Sampling Process

Field Testing

The goal of human subject driving tests is to conduct basic and applied research to understand the performance of the sensors in the vehicle, across a range of environmental conditions. Such studies are undertaken in more controlled settings by DADSS researchers, and in naturalistic settings in cooperation with the states of Virginia and Maryland as described below.

Controlled Human Subjects On-Road Driving Tests

The purpose of the controlled human subject on-road driving tests (HSD) tests is to conduct basic and applied research to understand the performance of the DADSS sensors in the vehicle physiologically and ergonomically. During the HSD tests, in-vehicle testing is undertaken in a diverse set of geographic and environmental conditions, varying vehicle conditions, and with diverse human subjects to assess the effects of human variability. Routes in New England were chosen to provide varying climactic conditions, such as low and high temperatures, low and high humidity, at varying elevations, and in corrosive environments.

Results from on-road testing is critical in determining the effectiveness of the DADSS sensors in a wide range of conditions including the impact of environmental factors on sensor function over time, the impact of repeated use and vehicle mileage, the impact of vehicle vibration, and user interactions with these devices in a vehicle environment, including driver behavior and user acceptance. The data will also be used to refine the DADSS Performance Specifications, and to improve system design and product development.

Once the breath sensors performed well in laboratory and human subject testing, vehicle trials in real-world driving environments began with the Gen 3.2 breath sensors (known as HSD1) and are currently on-going using the Gen 3.3 breath sensors (knows as HSD2). Once Gen 4.0 breath sensors become available, HSD3 studies will commence. The ultimate goal of the breath sensors is to passively detect drivers' breath alcohol within the ambient air of the vehicle cabin. For the HSD1 and HSD2 studies, subjects are instructed to breathe towards the sensors to assist in data collection. The subjects' directed breath has two types of variability: 1) subjects likely will breathe differently each time they give a sample, for example, the amount of breath, the strength of each breath, and the direction of the breath will vary; and 2) each person breathes slightly differently. One person's version of "provide a breath as if blowing out a candle" may be different than another's. When you factor in variations in subject height, distractibility during the drive, and other variables, these can all contribute to potential variations observed in sensor performance. Variations also arise with changing environmental conditions inside and outside the vehicles. As a result, a large amount of data is needed to fully understand how each of the sensors works across all conditions for every driver.

The controlled HSD tests utilize fully instrumented DADSS Program vehicles equipped with four (4) DADSS breath sensors integrated into the vehicle to measure breath alcohol – two DADSS breath sensors on the driver side and two on the passenger side. On the driver's side the breath sensors are mounted in the steering wheel location and the driver's door. On the passenger side they are mounted in the passenger door and on the dashboard directly in front of the passenger (Figure 13). The DADSS breath sensors can measure directed and passive breath. A comparative sensor, installed on the passenger side, provides a comparison measurement, and requires a deep lung sample of breath delivered through a plastic tube. This sensor is used to assess the DADSS breath sensor's sensitivity (i.e., true positives), validity, and reliability.

Along with the alcohol sensors, the vehicles are equipped with a comprehensive data acquisition system (DAS) with real time data upload to the DADSS data view, a video camera, a web interface[3], data and video storage, and a user interface module (UIM) for use by the test subject and research associate.

Figure 13. Position of the breath sensors and user interface module on the driver (left) and passenger sides of the vehicle.

A comprehensive data view dashboard provides real time data from all sensors as well as vehicle parameters (Figure 14). Elements displayed on the dashboard include the number of reference samples, the number of breath sensor samples collected, the number of protocols run, the total number of subjects, the total study days, and total miles driven. The dashboard also includes the above elements broken down by each protocol.

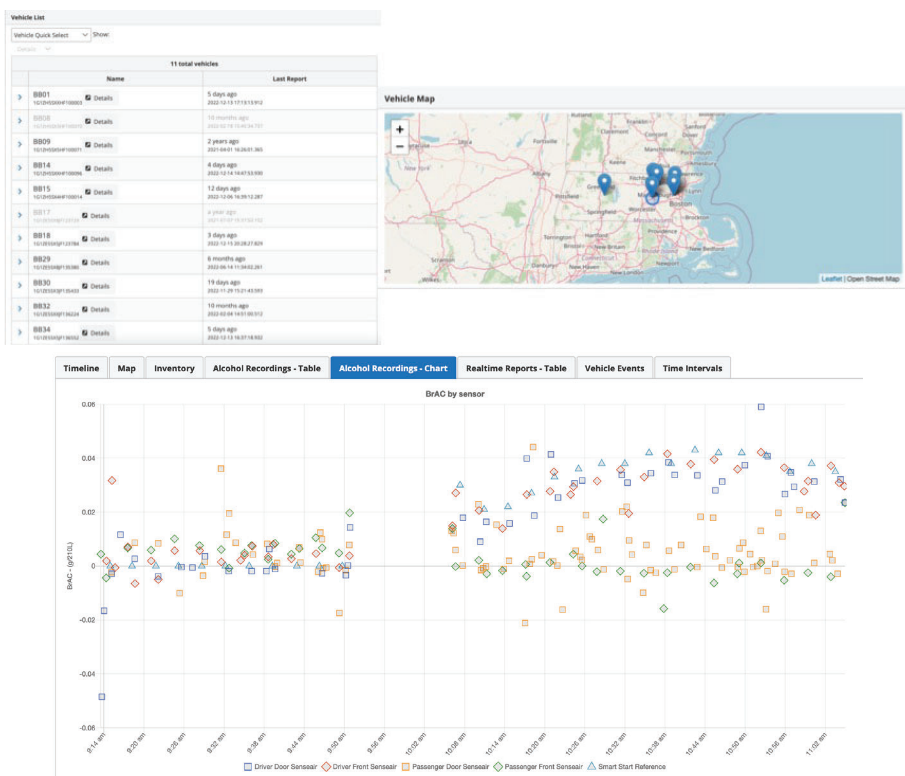


Figure 14. DADSS Data Viewer.

[3] A web interface allows the user to interact with content or software running on a remote server through a web browser.

Recruitment of the test subjects is being conducted at the DADSS laboratory and the McLean Hospital. Many of the test subjects have previously participated in DADSS human subject testing, affording researchers the opportunity to compare subjects' laboratory and in-vehicle data. Participants are brought into either of the laboratories prior to the study. The risks and benefits of the study are explained to them, and if they are comfortable with the study requirements and choose to participate, they sign the informed consent form. Subjects are screened for drug and alcohol presence, and they are familiarized with the vehicle set-up and protocol. After height and weight measurements are taken, they are dosed with the relevant quantities of alcohol over a period of about 10 minutes. The subjects alcohol measurements are collected frequently from the breath and reference sensors. The subjects are instructed by a research assistant in the vehicle to direct their breath in a prescribed sequence toward the DADSS breath sensors in the vehicle. The current methodology permits collection of BrAC on up to four different sensors every 3.5-4 minutes for up to eight hours. The research assistant also monitors the subject's condition.

During vehicle testing the DADSS sensors passively sniff and analyze the vehicle cabin air for the presence of alcohol. Additional vehicle instrumentation tracks environmental conditions and vehicle system data while providing participant videos.

Information collected from HSD1 tests contributed to the development Gen 3.3 breath sensor, including improved accuracy and precision, increased operational temperature range, a faster start-up, reduced cost, improved protection from electromagnetic interference (EMI) protection, and improved start-up behavior. In addition, the trials have functioned as proof of concept studies to identify which variables need to continue to be investigated as new generations of sensors are evaluated.

The Gen 3.3 sensor is now being evaluated in the second round of trials, referred to as HSD2. As noted above, the Gen 3.3 sensor is designed for use in fleet vehicles to detect the presence of alcohol, thus, the initial focus of these studies was on evaluating sober and low-dose alcohol participants. However, studies have also evaluated participants who were dosed to moderate and high doses of alcohol. Overall evaluations were conducted with BrACs ranging from 0.0 – 0.12%.

Table 3 provides a sample count of the DADSS breath sensors and comparative sensor, number of subjects and number of study protocols. Table 4 provides a samples count by sensor for HSD1 and HSD2.

Table 3. HSD1 and HSD2 Metrics.

Study	Total Study Protocols	Total Breath Sensor Samples	Total Comparative Samples	Total Unique Subjects	Total No. of Study Protocols
HSD 1 using Gen 3.2 Sensor	1	78,032	51,368	61	117
HSD 2 using Gen 3.3 Sensor	1	16,179	14,604	17*	245
TOTAL	18	94,211	65,972	67*	362

**11 Unique Subjects participated in both HSD 1 and HSD 2, therefore the total Unique Subjects is 67*

Table 4. HSD Samples Count by Sensor.

Study Protocol	Breath Sensor Sample Count	Comparative Sample Count	Subject Count	Protocol Count
Passenger Side Sober Stationary	3,395	398	17	25
Sober Stationary with Mask	352	118	10	10
Passenger Side Low dose (0.3 g/kg) Stationary	1274	178	12	12
Passenger Side Moderate Dose (0.5 g/kg) Stationary	1070	138	9	9
Low dose (0.3 g/kg) drive	509	131	11	13
Moderate dose (0.5 g/kg) drive	1,049	176	11	15
Sober drive	1,197	201	11	19
Moderate-High Dose (0.7g/kg) Drive	257	69	7	7
High Dose (0.9g/kg) Drive	244	60	5	5
Driver Side Sober Stationary	303	49	4	6
Low Dose Driver Side Stationary	341	61	3	3
Moderate Dose Driver Side Stationary	1,099	188	6	6
Hand Sanitizer - Blowing Across hands	317	0	10	10
Hand Sanitizer - Waving hands	242	28	10	10
Mouthwash	349	0	10	10
Mouthwash with water rinse	331	0	10	10
Mouthwash following UIM	172	26	10	10
Mouth Spray	216	0	10	10
Mouth Spray with water rinse	188	0	10	10
Mouth Spray following UIM	168	28	10	10
Lysol Disinfecting Spray	196	0	5	5
Individually Wrapped Hand Sanitizer Wipes	459	0	7	7
Isopropyl Alcohol Medical Pads	95	0	6	6
Hand Sanitizing Honest Wipes	276	0	5	5
Moderate-High Dose (0.7g/kg) Stationary	1786	278	7	7
High Dose (0.9 g/kg) Stationary	1479	221	5	5

Naturalistic Human Subjects On-Road Driving Tests

In 2018, The Driven to Protect, Powered by DADSS initiative, a partnership with the Virginia Department of Motor Vehicles, Highway Safety Office and the DADSS Program, partnered with James River Transportation (JRT), a transportation company with offices in Richmond and Norfolk, VA, to conduct the first in-vehicle, on-road naturalistic test trials with prototype breath-based alcohol sensors in their vehicles. The partnership was expanded in 2021 to include Schneider National, Inc. to conduct a pilot deployment out of its Virginia-based operations in South Boston, Virginia.

James River Transportation

Initially, four fleet vehicles (2015 Ford Flex airport livery vehicles) were instrumented with Gen 3.1 sensors and underwent field tests. In late 2020, JRT decommissioned the 2015 Ford Flex vehicles and commissioned two 2019 Ford Flex vehicles with the Gen 3.2 breath sensors, and those on-road field tests continue today.

Figure 15. JRT Sensors Location.

At start-up, the JRT driver provides a breath sample toward the Gen 3.2 sensors located in the driver door and on top of the steering column (see Figure 15) and to the aftermarket commercially available interlock sensor used as a reference sensor. If the sensor detects alcohol in the sample breath, an alert goes to authorized personnel. Personnel quickly assess the alert and other relevant data, such as breath readings, directly before and after to determine a potential course(s) of action for JRT consideration.

The JRT pilot deployment project has provided valuable feedback on the driver's experience and interaction with the sensors and allowed troubleshooting anomalous readings or problems with the sensors and data acquisition system. In the initial stages of the project, continuous video surveillance added information on the driver's interaction with the sensor. In addition, two small focus groups were conducted in 2019 with JRT drivers to understand driver receptiveness to the technology, their preference for sensor prompts and other sensor interactions, and their feedback on the program training and driver test plan. Driver feedback showed an initial acclimation period in using the sensors. About a quarter of post-drive surveys in the first two months indicated difficulties with the sensors, including low breath volumes and unfamiliarity with indicator lights on the sensors. These difficulties declined significantly in the following months. Feedback indicated a slight learning curve in providing a breath sample to the sensors, but after a short time, they "got the hang of it." In addition, the post-drive surveys were important for feedback on the functioning of the sensors and data acquisition system. Several drivers reported problems with interference with in-vehicle GPS, radio, and keyless entry (such as EMI). This real-time reporting allowed technical modifications to be made to limit this interference with the addition of sensor shielding.

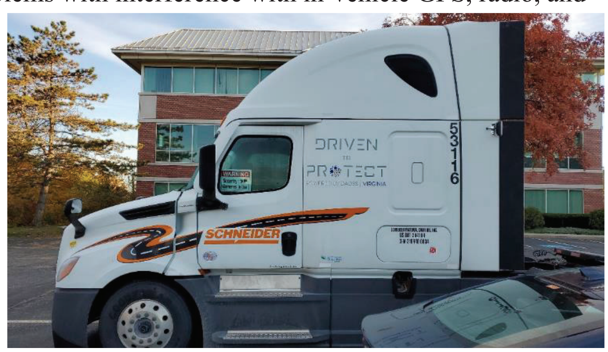


Figure 16. Schneider Freightliner Cascadia Truck.

To date, test vehicles were operational for 29,996 hours and driven 97,746 miles, during which time the system collected 142,388 breath samples.

Schneider National, Inc

The Schneider pilot deployment began with equipping a Freightliner Cascadia heavy-duty semi-trailer truck (Figure 16) with two (2) Gen 3.3 breath sensors, a data acquisition system (DAS), a driver display (Figure 17). The initial truck was used as a platform truck for the DADSS engineering team to design, develop and test the sensors and instrumentation. The platform truck was extensively road tested by a Schneider operator, in collaboration with the DADSS program, prior to integrating the system into 7 additional trial vehicles. Schneider stipulated that the installation should not be distracting to the driver and should not include permanent modifications to the trim package or cab interior or interfere with the vehicle's safety features and equipment. In addition, efforts were made to minimize driver and supervisor actions (e.g., responding to false alarms).

Figure 17. Overhead Sensor Above Driver's Head and Dashboard Unit.

Engineers developed the operating software for the Schneider truck, with a goal to mitigate the likelihood of a false positive. A collaborative working group agreed upon a threshold of above 0.025% BrAC from both sensors to generate an initial on-screen notification to the driver while concurrently sending a message to key program personnel. Prior to moving the vehicle, the system prompts the driver to provide additional verification breaths to the sensors immediately. If the follow-on breaths are also above the threshold, the system would generate an alert via email and text message to the Schneider regulatory team as well as key program personnel monitoring notifications and alerts. The display also prompts the driver to contact their immediate supervisor or manager via on-screen messaging.

The on-road deployment of the platform truck allowed researchers to assess the value of the fleet sensor system and collect data for future technology improvements. In May 2022, the DADSS Team installed the sensors into seven (7) additional Schneider trucks and performed system updates to the platform vehicle in preparation for integration into the overall fleet pilot environment. After installation and testing of the systems, the team met with each volunteer driver to provide hands-on training in system operation and answer any questions about the technology or system operation. They were also provided with some basic troubleshooting techniques and briefed on contacting the team to resolve any problems, including an online fillable form to record any issues or concerns. Initially, a few concerns arose, including a few drivers who were too short to see the right front mirror over the dashboard unit. The resolution required disassembling the components of the all-in-one dashboard unit and segregating the components into alternate mounting locations (Figure 18). The engineering team also needed to redesign a new housing, inlet, and snorkel to reposition the dashboard sensor to an alternate area. The trucks now have three trucks with the initial system installation on the dash and three trucks with the revised alternative system for a total of six active Schneider trucks gathering project data.

Figure 18. Alternate DADSS Installation – System Display and Sensor.

As of October 1, 2022, test vehicles were operational for 6,987 hours and driven 61,143 miles, during which time sensors collected 25,171 breath samples.

CONCLUSIONS

Since its inception in 2008, the DADSS Program has made tremendous progress in the development of in-vehicle technologies that will prevent impaired drivers from driving their vehicles. The breath- and touch-based sensors have become increasingly refined, both in terms of hardware and software, as headway is made in meeting the high standards required for unobtrusive and reliable alcohol measurement. At the same time, additional research and development is paralleling the sensors' development to allow the characterization of sensor performance in the laboratory, on the road, and among human subjects.

Substantial progress has been made in the development of breath and touch-based calibration processes, materials and methodologies making the testing of multiple sensors at a time a reality. Moreover, instrumentation is being used that can enable the requisite testing across the range of specified environmental conditions.

As sensor development has progressed, research vehicles have been readied for on road testing with the latest versions of the breath sensors seamlessly integrated within the vehicle interiors. Vehicle instrumentation packages have been developed and installed and pilot testing trials are providing data on sensor performance under real-world driving conditions. The accumulated data from an extended program of on road driving trials under diverse conditions will determine whether the DADSS sensors are working as anticipated and allow the identification of areas for system improvement.

A comprehensive program of human subject research also is well underway, starting with the laboratory environment where better control of conditions can be exerted, and continuing in the vehicle where the sensors can be tested in the environment in which they will be used. This research has quantified alcohol absorption and elimination across a wide range of conditions that are anticipated to affect BAC and BrAC as well as some new scenarios specific to the breath and touch-based approaches. Testing showed that the data collected from the various generations of breath-based and touch-based prototypes were consistent, reproducible, and correlates very well with the "gold-standard" method of measuring alcohol in the body, which is accomplished by measuring blood via gas chromatography.

In summary, great progress has been made on a number of fronts to develop in-vehicle sensors that will seamlessly measure driver's blood and breath alcohol and prevent them from driving while impaired. Moreover, additional research is ongoing to continue the progress toward meeting the exacting performance specifications to ensure acceptance and longevity in the vehicle environment.

REFERENCES

Biondo, W., Zaouk A. K, Sundararajan, S. 2017. Driver Alcohol Detection System for Safety (DADSS) – Development of the Subsystem Performance Specifications. Paper Number 17-0301. Proceedings of the 25th International Technical Conference on the Enhanced Safety of Vehicles.

Ferguson, S. A. 2012. Alcohol-impaired driving in the United States: Contributors to the problem and effective countermeasures. Traffic Injury Prevention, 13, 427-441.

Ferguson, S.A., Traube, E, Zaouk, A, Strassburger, R. 2009. Driver Alcohol Detection System For Safety (DADSS) – A non-regulatory approach in the development and deployment of vehicle safety technology to reduce alcohol-impaired driving. Paper Number 09-0464. Proceedings of the 21st International Technical Conference on the Enhanced Safety of Vehicles.

Ferguson, S.A., Zaouk, A., Strohl, C. 2010. Driver Alcohol Detection System for Safety (DADSS). Background and Rationale for Technology Approaches. Society for Automotive Engineers Technical Paper, Paper No. 2010-01-1580, Warrendale, PA: Society of Automotive Engineers.

Ferguson, S.A.; Traube, E; Zaouk, A; Strassburger, R. 2011. Driver Alcohol Detection System For Safety (DADSS) – Phase I Prototype Testing And Finding. Paper Number 11-0230. Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles.

Fournier, R, Zaouk, A.K., Willis, M., Strassburger, R. 2019. Assessing system implementation readiness of the Driver Alcohol Detection System for Safety (DADSS) to reduce alcohol-impaired driving in a real-world driving pilot deployment project. Paper Number 19-0263. Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles.

Fratto, B.E. & Katz, E. Utilization of a fluidic infrastructure for the realization of enzyme-based Boolean logic operations. *Int. J. Parallel Emergent Distrib. Syst.* In Press. (DOI: 10.1080/17445760.2016.1213250.

Fratto, B.E. & Katz, E. Reversible Logic Gates Based on Enzyme-Biocatalyzed Reactions and Realized in Flow Cells: A modular Approach. *ChemPhysChem.,* 2015, 16, 1405 –1415 DOI: 10.1002/cphc.201500042

Hök, B., Pettersson, H., Andersson, G. 2006. Contactless measurement of breath alcohol. Proceedings of the Micro Structure Workshop. May, 2006, Västerås, Sweden.

Jones, A.W. 2008 Biochemical and Physiological Research on the Disposition and Fate of Ethanol in the Body. Garriott's Medicological Aspects of Alcohol (Fifth Edition), 47.

Katz, E. & Fratto, B.E. Enzyme-based reversible logic gates operated in flow cells. In: *Advances in Unconventional Computing,* A. Adamatzky (ED.), Springer, 2016, Volume 2, Chapter 2, pp 29-59.

Lindberg, L., Brauer, S, Wollmer, P., Goldberg, L., Jones, A.W., Olsson, S.G. 2007. Breath alcohol concentration determined with a new analyzer using free exhalation predicts almost precisely the arterial blood concentration. Forensic Science International. 168: 200-207.

Ljungblad, J., Hök, B., Allalou, A., Pettersson, H. 2017. Passive in-vehicle driver breath alcohol detection using advanced sensor signal acquisition and fusion. Traffic Injury Prevention, Vol 18, 31-36.

Lukas, S.E., Ryan, E., McNeil, J., Shepherd, J., Bingham, L.,Davis, K., Ozdemir, K., Dalal, N., Pirooz, K., Willis, M., Zaouk, A. K. 2019. Driver Alcohol Detection System for Safety (DADSS) – Human Testing of Two Passive Methods of Detecting Alcohol in Tissue and Breath Compared to Venous Blood. Paper Number 19-0268. Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles.

National Highway Traffic Safety Administration (NHTSA) DOT HS 813298 A Brief Statistical Summary May 2022, Early Estimates of Motor Vehicle Traffic Fatalities and Fatality Rate by Sub-Categories in 2021.

National Highway Traffic Safety (NHTSA) DOT HS 813 266 Overview of Motor Vehicle Crashes in 2020.

Privman, V., Fratto, B.E., Zavalov, O., Halámek, J., Katz, E. Enzymatic AND Logic Gate with Sigmoid Response Induced by Photochemically Controlled Oxidation of the Output. *J. Phys. Chem. B*, 2013, *117* (25), pp 7559–7568

Ridder, T.D., Hendee, S.P., Brown, C.D. 2005. Noninvasive alcohol testing using diffuse reflectance near-infrared spectroscopy. Applied Spectroscopy, 59, 181-189.

Verma, A., Fratto, B.E., Privman, V., Katz, E. Design of Flow Systems for Improved Networking and Reduced Noise in Biomolecular Signal Processing in Biocomputing and Biosensing Applications. *Sensors (MDPI)* 2016, 16, 1042; doi: 10.3390/s16071042

Ver Steeg B, Treese T, Adelante R, Kraintz A, Laaksonen B, Ridder T, Legge M, Koslowski N, Zeller S, Hildebrandt L, Koeth J, Cech L, Rumps D, Nagolu M, Cox D. 2017. Development of a solid state, non-invasive, human touch-based blood alcohol sensor. Paper Number 17-0036. Proceedings of the 25th International Technical Conference on the Enhanced Safety of Vehicles.

Willis, M., Zaouk, A.K., Bowers, K., Chaggaris, C., Spicer, R., Bahouth, G., Strassburger, R., Traube, E. Driver Alcohol Detection System for Safety (DADSS) – pilot Field Operational Tests (PFOT) vehicle instrumentation and integration of DADSS technology. Paper Number 19-0262 Proceedings of the 26th International Technical Conference on the Enhanced Safety of Vehicles.

Zaouk AK, Willis M, Traube E, & Strassburger R. 2015. Driver Alcohol Detection System for Safety (DADSS) – A Status Update. Paper Number 15-0276. Proceedings of the 24th International Technical Conference on the Enhanced Safety of Vehicles.

Zaouk A. K, Willis, M., Dalal, N., Traube, E., Strassburger, R. 2017. Driver Alcohol Detection System for Safety (DADSS) – A Status Update. Paper Number 17-0271. Proceedings of the 25th International Technical Conference on the Enhanced Safety of Vehicles.

Evaluation of simulated Level 2 hands-free driving in real traffic – an innovative method for an early SOTIF Human Factors assessment of ADAS under realistic driving conditions

Manuela Witt[a]*, Florian Raisch[a], Martin Götze[a] and Burak Gülsen[a]

[a]*Vehicle Safety, BMW AG, Munich, Germany, manuela.witt@bmw.de*

We declare we have no conflict of interest to disclose.

Evaluation of simulated Level 2 hands-free driving in real traffic – an innovative method for an early SOTIF Human Factors assessment of ADAS under realistic driving conditions

Objective: Recent activities in the development of assisted and automated driving involve vivid discussions about the necessity to evaluate the interaction between the driver and the system, especially while using high performant SAE Level 2 functions (see SAE, 2016). The assessment of safety in use of these systems is fundamental and includes several methods that can be applied to evaluate the controllability of the systems, e.g., simulation, driving simulator studies, and realistic driving studies on test tracks or in real traffic. In early development stages, it is barely possible to assess the functions in real traffic. However, some research questions need to be addressed early and can be answered the most appropriate by studies in real traffic. Therefore, a method to simulate new SAE Level 2 and even Level 3 systems in test vehicles has been developed.

Method: A new method to assess driver behavior and controllability of system limits in real traffic is presented: By using assisted driving functions of series vehicles, higher assisted functions can be simulated in the user interface and additional functional features can be implemented, such as automated lane changes that can be triggered by a trained safety driver sitting on the passenger seat. Thereby, it is possible to assess fundamental Human Factors aspects, such as mode confusion, overreliance or overtrust, under highly realistic study conditions or even assess controllability of lateral steering errors in real traffic. A realistic driving study to assess controllability of such system limits while driving SAE Level 2 hands-free is presented.

Results: Simulating new SAE Level 2 functions by using special test vehicles and trained safety drivers enables researchers to evaluate the driver's interaction with these functions under controlled and very realistic conditions. The results of such studies can help to identify risks and, thereby, define appropriate measures to address and minimize them. Moreover, hypotheses about driver behavior can be tested and validated to support a safety-oriented development process. The results of the presented study on controllability of sudden steering errors show that attentive drivers are able to control system-detected as well as system-undetected lane drifts while driving SAE Level 2 hands-free. Differences in reaction times were significantly correlated with if the steering error occurred and an urgent

warning was triggered or if the lane drift was undetected by the system and no warning was issued.

Conclusion: Evaluating driver behavior in real traffic while using SAE Level 2 systems is necessary to assess safety in use of these functions before introducing them into the market. Simulating new systems in series vehicles helps getting important insights into driver behavior while using such functions. System limits to be expected can be presented and controllability of the resulting situations can be assessed as well as driver reactions in terms of reaction times and quality of intervention.

Keywords: SAE Level 2 Hands-Free Driving, Safety in Use, Human Factors

INTRODUCTION

In recent years, SAE Level 2 driving functions, which equals Partially Automated Driving (PAD), is getting more and more attention in media and an increasing number of highly performant Advanced Driver Assistance Systems (ADAS) are flooding the market. While some OEMs already took it to the next level and have introduced Level 3 (Highly Automated Driving, HAD), still, there are major challenges which must be faced in the field of PAD. Technology is improving progressively and rapidly which makes it possible to provide drivers PAD in more and more complex situations and include features, such as hands-free driving. Yet, the driver is still accountable for the safe guidance of the vehicle and must stay attentive at any time as well as be prepared to intervene immediately, when the system suddenly fails, or system limits occur. Several studies have shown that automation can have a negative impact on the driver's behavior, such as loss of skill (Stanton & Marsden, 1996), loss of situational awareness (Endsley, 1999; Clark et al., 2017) or overreliance on the automation (Parasuraman & Riley, 1997), as the driver's role is shifted from an actively performing to a passive monitoring task (Parasuraman & Riley, 1997; Parasuraman et al., 1993). Studies on the controllability of system limits of PAD apart from studies in driving simulators but in real traffic are scarce. When PAD is active, the system supports the driver with active steering, hence, steering errors or system limits can occur that require steering input by the driver, i.e., because the car is following the wrong trajectory or is leaving the lane due to missing lane markings. There are few studies that investigated the driver's reaction to these situations while driving SAE Level 2 and none that have been carried out in real traffic. A most recent study by Schneider et al. (2022) assessed controllability of lateral drift failures while driving SAE Level 2 hands-free on the test track. Naujoks et al. (2014) investigated the driver's reaction to a sudden take over request while driving a Level 2 hands-free system in the driving simulator. Different lateral guidance system limits that were caused by either ending lane markings, poor lane markings in construction sites or due to high curvature occurred. Several driving simulator studies examined the driver's reaction behavior to sudden obstacles in the road such as broken-down vehicles or lost cargo (Gold et al., 2013; Strand et al., 2014; Naujoks et al., 2015). Other authors carried out studies to investigate the driver's reaction time to system limits compared to system malfunctions (i.e., vehicle following the wrong trajectory vs. vehicle leaving lane) (DeGuzman et al., 2020; Sieber et al., 2015), also in a driving simulator environment.

METHOD

Methods for the assessment of safety in use of ADAS

The necessity of developing new and innovative methods to assess rapidity and quality of the driver's intervention, when system limits occur, is manifest. Reaction times to such events and take over quality need to be assessed to identify measures that are necessary to increase controllability and establish functional limits, e.g., in terms of lateral and longitudinal dynamic parameters. These are crucial to increase the driver's time budget when system limits occur while the driver is using PAD to prevent accidents and fatalities. Adequacy of different methods to assess controllability of system limits depends on several factors. Crucial are the maturity of the respective system, the claim on the validity of the results which highly depends on the stage of development, the research question to be answered and the current state of knowledge in the respective field of research. Derived from these key aspects Figure 1 shows the variety of assessment methods. In early stages, literature reviews, online surveys and expert evaluations are used to support the early development process and address upcoming research questions. Most of the studies cited above are driving simulator studies which is a highly powerful tool to examine quality of human behavior in interaction with ADAS, PAD and HAD in a controlled environment. Certain isolated research questions are predestined to be conducted on the test track, e.g., parking maneuvers or highly dynamic maneuvers which for safety reasons cannot be tested in real traffic. Other research questions require a more realistic test environment, which includes driving in real traffic on the road. Realistic driving studies, especially in real traffic, necessitate a high functional maturity of the considered system as well as an elaborated safety concept. By means of personal, technical, and organizational measures, realistic driving studies can be designed in a safe way and risks can be minimized to an acceptable low level. The L3EC method (marked in bold letters in Figure 1) can be applied in an early development stage under realistic test conditions which increases validity of the results and minimizes limitations due to an artificial laboratory test environment or simulated situations instead of real situations on the road.

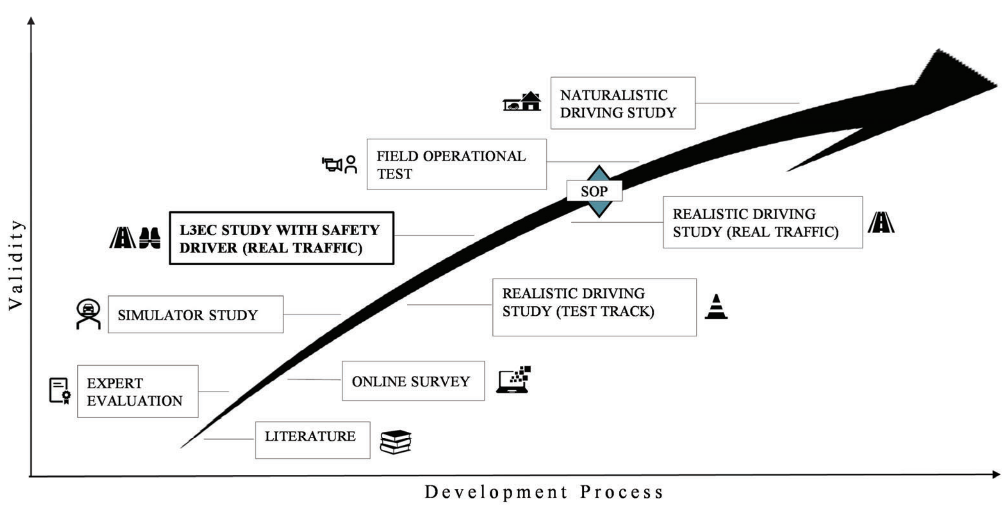


Figure 1. Methods for the assessment of safety in use of ADAS.

In the following section a study in the L3EC which has been carried out in real traffic with the main goal to assess controllability of system-detected and system-undetected steering errors, is presented.

Study set up and study design

By using the L3EC method, assisted driving functions of series vehicles, which are equipped with driving school pedals and additional mirrors and displays, higher assisted functions can be simulated in their behavior and interaction in regard to the user interface. Moreover, additional functional features can be implemented, such as automated lane changes that can be triggered by a trained safety driver sitting on the passenger seat. Therefore, additional switches are installed on the right side of the passenger seat so that the safety driver can give clearance of the function, issue warnings or even trigger error injections to simulate sudden steering errors. Additional cameras are installed to record traffic surrounding to all four sides of the vehicle (front view, rear view, right side, left side) as well as two more cameras facing the driver and one GoPro camera which is recording the driver's hands.

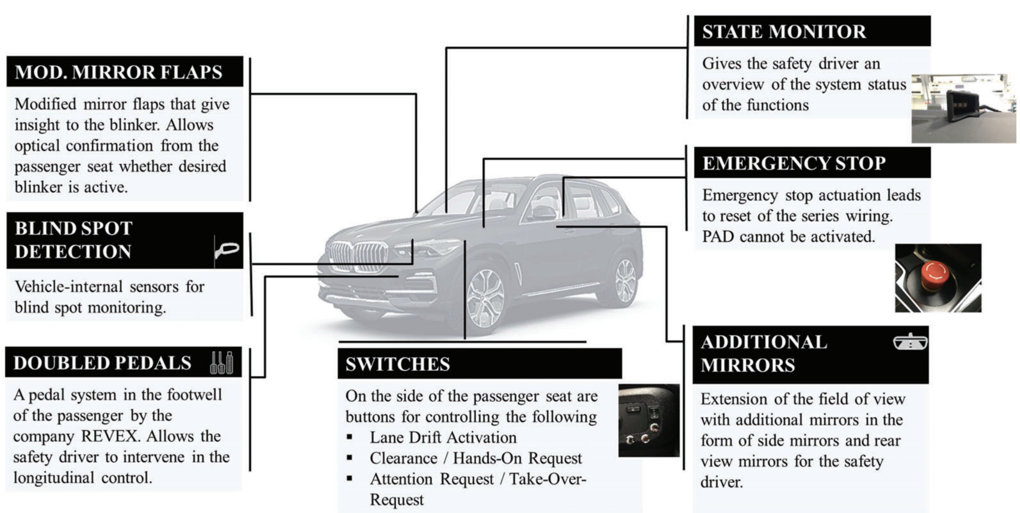


Figure 2. The Level 3 Experience Car.

The L3EC has been initially set up to facilitate a HAD (SAE Level 3) experience in real traffic under safe conditions and with real customers. With the introduction of Level 2 hands-free functions, a new experience of PAD is brought to the customer, allowing the driver to completely take his hands off the steering wheel. However, sudden steering errors can still occur, forcing the driver to put his hands back on the steering wheel and intervene immediately to take back control over the vehicle. By using the L3EC, a realistic driving study was carried out, in which the safety driver was able to trigger a sudden steering error that would result in the vehicle leaving lane, if the driver did not take over and steer back into the ego lane. Thereby, conditions under which the system limit would occur could be controlled completely by the trained safety driver who was also able to intervene by steering or using the pedals in case of emergency or if necessary. The safety drivers were specifically trained by BMW driving instructors to be able to control highly dynamic driving maneuvers from the passenger seat by using the pedals or grabbing the steering wheel from the passenger's side. The test drives have been executed in Phoenix, Arizona, with 8 Power Users of PAD and 12 drivers without any experience with PAD (= Non-Power Users). Drivers were declared as Power Users, if they were using PAD several times per month, and Non-Power Users were drivers with no experience with PAD at all. The participants have been recruited by a market investigation agency and the study has been conducted by an independent research institute. Every test drive consisted of one manual driving section, three PAD sections and two sudden system limits which were triggered by the safety in the first and the third PAD section. In between the different sections, the drivers had to drive off the highway to make a U-Turn and then drive back on the same highway. Figure 2 shows the study procedure which was split equally in the respective sections (~10 min. each section). The manual baseline was permuted, so that half of the drivers drove manually before, and half of the drivers drove manually after the PAD section to control possible sequence effects.

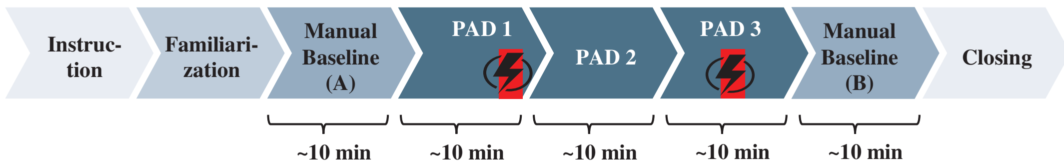


Figure 3. Study design and test procedure.

Drivers were told to obey traffic rules and activate and use the hands-free feature whenever it is available. Moreover, recommended speed was ~ 65 – 70 mph, if there was no other speed limit, and to keep up with the flow of traffic. Depending on time of day, traffic density and actual speed, the duration of each driving section varied between 7 and 10 minutes. The lane drifts were triggered at predefined spots to ensure reproducibility of the system limits for each participant. The first lane drift was combined with a take over request, whereas the second lane drift was undetected by the system and, therefore, was not issued with any warning. To minimize any risks for driver, passengers or other traffic participants, the safety driver made sure that there were no surrounding vehicles or lost cargo and debris on the shoulder, as the vehicle would always drift towards the shoulder on the right side of the road. PAD hands- was simulated by using the PAD hands-on lane keep assist function in the series vehicle. The capacitive hands-on detection was deactivated as soon as the simulated PAD hands-free was activated by the driver by pressing a button on the steering wheel. Driving PAD hands-on was still possible and clearance of the hands-free system was controlled and provided by the safety driver. By pressing a button on the right side of the passenger seat, the safety driver could give and take away availability of the hands-free function. The HMI was freely programmable and availability of the function as well as all other functional statuses were presented to the driver in the instrument cluster as well as by LED lights on a small display on the side of the safety driver. A warning cascade was integrated that would be triggered when the driver has been inattentive for a few seconds, and which escalated with urgent take-over requests when the driver did not react to the initial warning due to being inattentive. The warning cascade is shown in Figure 4.

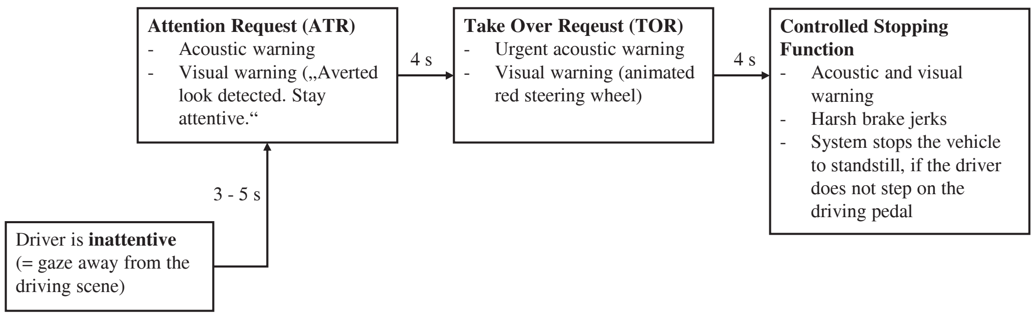


Figure 4. ATR warning cascade.

Reaction times, take over quality and necessity of safety driver intervention were evaluated to assess controllability of sudden steering errors while driving PAD hands-free. Differences in reaction time to lane drifts with an urgent take over request and lane drifts without a warning (= undetected system boundary) were evaluated. Lane drifts were only triggered when the driver was attentive and was not currently issued by an ATR or when other safety functions were currently active (= acute intervention of a safety system, e.g., emergency braking). Thereby, the impact of error variance on the test results, e.g., due to heterogeneous driver states while the system boundary occurred, was minimized and the results only represent controllability of these steering errors without consideration of other Human Factors aspects. Aberrant driver states, such as extreme fatigue, inattentiveness, or driver distraction, i.e., caused by driver engagement in secondary tasks, were not considered or tolerated during the test drive. Moreover, post interviews after the test drive were evaluated to get a deeper insight into how drivers perceived the system limits and get more information about the driver's mental model, e.g., concerning responsibility or necessity to permanently stay attentive and monitor the driving scene.

RESULTS

Reaction times

Reaction times were defined as time between the start of the lateral movement during the lane drift and (1) hands-on and (2) the first steering input by the driver. Data were evaluated by video analysis and time frames were clicked in 17 ms steps. In the following, results of the study are presented. Videos of two drivers were missing. Therefore, N was reduced to n = 18. Figure 5 shows reactions times to both lane drifts, with and without warning, as well as duration until hands-on solely and duration until first steering input of the driver. Difference between these values were minor, which means that drivers started steering immediately as soon as they put their hands back on the steering wheel.

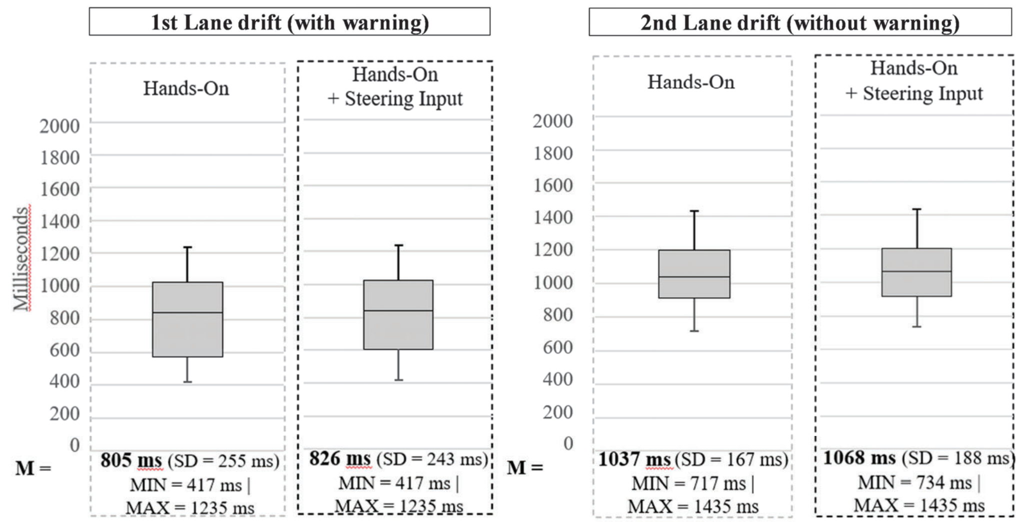


Figure 5. Reaction times to lane drift with and without TOR.

The drivers' reaction times varied between 417 and 1435 ms. There was a significant difference between reaction times to system-detected lane drifts with warning and system-undetected lane drifts without warning. Reaction times (hands-on + steering input) to the detected lane drift by the system with an issued warning were significantly shorter than reaction times to the undetected lane drift by the system without warning (U: 70; $p < 0.05$). There was no significant difference between reaction times with or without warning depending on the fact if the driver was an experienced Power User or an unexperienced Non-Power User. There was never a necessity for the safety driver to intervene during any of the lane drifts before the driver intervened. Both drifts, with and without warning, were controllable by all drivers. Some drivers reacted very strongly, which was one reason for some stabilizing safety driver interventions. Yet, none of the drivers started swerving or lost control over the vehicle and there were no crashes or near crashes, which indicates a sufficiently good take over quality. In the present study, a series Level 2 hands-on function was used to simulate a hands-free feature. The hands-on function allows cooperative steering and can be easily overridden by small steering inputs. In the series version of the simulated Level 2 hands-free system, additional technical measures are implemented to support stability and controllability of steering errors that require quick driver interventions, e.g., a slow ramp out of the steering torque. In the study, several drivers left the ego lane due to the lane drift and crossed the lane marking towards the shoulder with one wheel. There was enough space between the ego lane and the concrete wall or crash barrier next to the shoulder so that leaving the ego lane or crossing the lane marking was not dangerous. This might be an effect of subjective risk assessment by each driver. Figure 6 shows the road sections and impressions of the traffic situations, on and in which the drifts were triggered.

Figure 6. Road sections for the lane drifts

Post Interview

After the test drive, participants were asked questions about their experience with the system and the lane drift situations they were confronted with. There were striking errors in the mental model of some participants concerning accountability for the driving task, necessary monitoring behaviour and expectancies on warnings when system limits occur. Even though there were wrong beliefs about and false answers to the items on the topics mentioned above, the video evaluation of the drivers' behaviours in terms of monitoring behaviour or engagement in secondary tasks was unobtrusive. Moreover, all drivers – even if they stated in the post interview that they were not accountable for the driving task – reacted quickly to the lane drifts by grabbing the steering wheel and steering back into the ego lane. Figure 7 shows the answers of Power Users and Non-Power Users to the respective items on accountability, monitoring behaviour and expectancy towards the system on warnings in the event of a system limit.

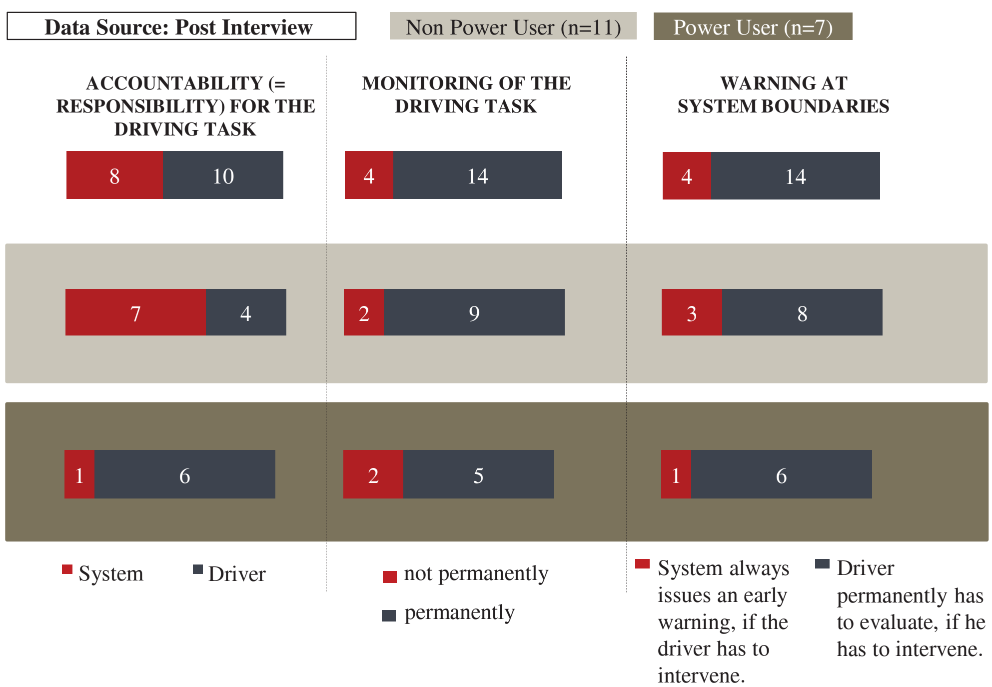


Figure 7. Post Interview on accountability, monitoring behaviour and system limits.

SUMMARY AND DISCUSSION

The L3EC is an innovative method to assess controllability of system limits in a realistic driving environment, which can and will occur to customers in their vehicles after market introduction of SAE Level 2 hands-free features. Knowing how drivers react in these situations is crucial during the development process of SAE Level 2 functions, most importantly, before their launch in the respective markets. Risks are minimized due to full control about the timing and surrounding conditions during the lane drift by the safety driver as well as due to additional safety measures which were integrated into the test vehicle, such as additional pedals and mirrors.

The goal of the presented study was the controllability assessment of PAD system boundaries while driving hands off under realistic traffic conditions. Results show that sudden system limits that require an immediate reaction and steering intervention are controllable by an attentive driver. Nevertheless, some drivers reported wrong beliefs and an incorrect mental model about responsibility, monitoring of the driving task and readiness for driver intervention to sudden events. Good performance of the system, being able to drive hands-free as well as external factors, such as low traffic density, were listed as reasons for why drivers thought they were not responsible for the driving task or didn't have to permanently monitor the driving task, when the system was active. The results indicate an impact of experience with PAD on the correctness of the driver's mental model. More Non-Power Users thought that the system was responsible instead of the driver, whereas only one Power User had a wrong belief about responsibility for the driving task. Therefore, it can be assumed that the driver's mental model can be corrected with growing experience with PAD and that there may have been an effect of first-time usage in the data of the present study. Reaction times to lane drifts were shorter, when an acoustic and visual urgent warning was triggered, when the system limit occurred, than when there was no warning. Take over quality was good, independently of system-detected or system-undetected situations, and drivers were able to control the vehicle and steer back into the ego lane without swerving or the necessity of an intervention by the safety driver. In the presented study, the lane drifts were only triggered in situations, in which the driver was attentive and when there was only low density of surrounding traffic and no object or other vehicle was in the corridor, into which the vehicle would drift after the steering error was injected. Therefore, there was no immediate danger and reaction times might have been shorter than how they were observed in the present study, if there was another object, vehicle, VRU (= Vulnerable Road User) or barrier next to the vehicle. Thus, a follow-up study has been carried out, with a similar study set up but with lane drifts, that were triggered by the safety driver, when the ego vehicle was close to another object. The study has shown that reaction times were significantly shorter, when there is an immediate risk of collision, than when there is enough space to evade until a concrete crash barrier next to the shoulder or another object would be reached. Results of this study will be published separately and will not further be discussed in the present paper. These results, along with the results of the present paper, show that drivers react instinctively, when there is an immediate threat, no matter, if they are convinced that they are responsible or if the system is responsible. False use of the system, which can e.g., lead to being less attentive to the driving scene or engaging in non-driving relevant tasks, can of course impact controllability as dangers may not be perceived in time. Long-term effects on the driver's behavior by the usage of these functions need to be observed in Field Operational Tests and Naturalistic Driving Studies.

REFERENCES

Clark, H., McLaughlin, A. C., & Feng, J. (2017, September). Situational awareness and time to takeover: exploring an alternative method to measure engagement with high-level automation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1452-1456). Sage CA: Los Angeles, CA: SAGE Publications.

DeGuzman, C. A., Hopkins, S. A., & Donmez, B. (2020). Driver Takeover Performance and Monitoring Behavior with Driving Automation at System-Limit versus System-Malfunction Failures. *Transportation Research Record,* 0361198120912228.

Endsley, M. R. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics, 42*(3), 462-492.

Gold, C., Damböck, D., Lorenz, L. & Bengler, K. (2013). "Take over!" How long does it take to get the driver back into the loop? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57, 1938-1942.

Naujoks, F., Mai, C. & Neukum, A. (2014). The effect of urgency of take-over requests during highly automated driving under distraction conditions. In T. Ahram, W. Karowski & T. Marek (Eds.), Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014 (pp. 2099-2106). Krakau: AHFE Conference.

Naujoks, F., Purucker, C., Neukum, A., Wolter, S., & Steiger, R. (2015). Controllability of Partially Automated Driving functions–Does it matter whether drivers are allowed to take their hands off the steering wheel?. *Transportation research part F: traffic psychology and behaviour, 35*, 185-198.

Parasuraman, R. & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 39*(2), 230-253.

Parasuraman, R., Mouloua, M., Molloy, R. & Hilburn, B. (1993). Adaptive function allocation reduces performance cost of static automation. In R. S. Jensen & D. Neumeister (Eds.), 7th International Symposium on Aviation Psychology (pp. 37-42).

Schneider, N., Ahrens, L., & Pruksch, A. (2022). *Controllability of lateral drift failures while driving with SAE Level 2 Advanced Driver Assistance Systems.* In: 14. Tagung Fahrerassistenz, München.

Sieber, M., Siedersberger, K. H., Siegel, A., & Färber, B. (2015, September). Automatic emergency steering with distracted drivers: Effects of intervention design. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems* (pp. 2040-2045). IEEE.

Society of Automotive Engineers. 2016. "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles." SAE International, Warrensdale, PA.

Stanton, N. A. & Marsden, P. (1996). From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. *Safety Science, 24*(1), 35-49.

Strand, N., Nilsson, J., Karlsson, I. M., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation research part F: traffic psychology and behaviour, 27*, 218-228.

PREVENTING DRIVER MISUSE WITH PROACTIVE ADAS

Robert Fuchs
Yuuta Sakai
Tsutomu Tamura
JTEKT Corporation, R&D Headquarters
Japan

Paper Number 23-0296

ABSTRACT

As a consequence of the fast adoption of driving automation systems, most vehicles available on the market are the result of a robot-centered development approach. A few decades ago, the major challenges faced by the engineers were to implement sensors and control capability enabling the vehicle to follow and remain within a lane. For safety and to ensure compliance with the evolving regulations, driver monitoring systems (hands-on detection, head and gaze cameras) and override or takeover strategies completed the necessary equipment. The human driver has been considered afterward the development of the robot-like vehicle.

Focusing on lateral control, the majority of level-2 vehicles use an override strategy, which segregates manual from automated steering operation. Sometimes, this causes confusion resulting in distrust and ultimately misuse. Consequently, the level of acceptance of ADAS functions remains under the expectation. Active interaction with the automation is proposed to leverage driver engagement, which is considered as one of the key indicators for assessing safety of ADAS functions. The concept of haptic shared control of the steering enables manual intervention over the automation without deactivation. Systematic and consistent reconsideration of level-2 ADAS functions becomes possible when haptic interaction is exploited. Two proactive ADAS functions: active lane centering assistance and assisted lane change are proposed to enhance driver engagement while reducing the risk of misuse. Furthermore, it raises the question of the relevance of the driver monitoring system.

STATE OF THE ART

As defined by the SAE standard J3016 [1], the responsibility of the *dynamic driving task* (DDT) is shared at partial or level-2 automation: the automation operates the sustained lateral and longitudinal vehicle motion control, while the driver is in charge of performing the *object and event detection and response* (OEDR) [2]. Two advanced driver assistance systems (ADAS) are used in combination to comply with the above definition. *Adaptive cruise control* (ACC) operates a vehicle speed control, regulating a preset speed and automatically slowing to maintain a preset following distance to the slower moving vehicle ahead in the same lane. *Lane centering assistance* (LCA) provides continuous lateral support to maintain the vehicle centered in the lane. Level-0 ADAS are additional safeguards that prevent imminent risk occurrence. *Automatic emergency braking* (AEB) for the longitudinal displacement and *lane keeping assistance* (LKA) for the lateral deviations complete the active safety envelop. Therefore, the vehicle is capable of following a lane at a preset velocity, however, it does not have the ability to detect the surrounding traffic condition and will not respond appropriately. Consequently, it is the driver who is the sole responsible and he/she is required to monitor both the automated driving system and the road ahead for any objects or events, responding appropriately when needed. During a partially automated drive, there can be no period of time that the driver can disengage from the OEDR task. *Driver monitoring systems* (DMS) are used to assess the driver status and, along with an alert escalation procedure, to prevent sustained noncompliance. For level-2 vehicles, the UNECE regulation 79 requires a means of detecting that the driver is holding the steering wheel [3]. Some, so-called level-2+ vehicles enable hands free driving by using camera for head or eyes gaze monitoring. In summary, a partially automated vehicle is equipped with level-2 ADAS for reducing the driver workload while ensuring engagement, level-0 safeguard functions to prevent traffic accident and DMS to assess the driver status.

SAFETY PARADOX

While increasing the level of automation is regarded as a measure to meet environmental, productivity and traffic safety requirements, the role of the driver is shifted to a monitoring task, increasing risks for human to lack operational understanding. The paradox of automation (not only in automotive) is that the more proficient and reliable the system evolves, the incentive for the human to maintain attention reduces. Overreliance or complacency is created when an unjustified trust on the system ability builds up over time. The consequent loss of situation awareness results in an *out of the loop* (OOL) phenomenon or disengagement. Statistics and research suggest that automation can lead to accidents because of the OOL phenomenon. As a consequence, driver engagement has become one of the most relevant indicators for assessing safety of partially automated vehicles [4]. Although monitoring systems and attention reminders increase engagement, they are reactionary to driver behavior and do not guarantee continuous engagement. Hence, if the vehicle cannot assure OEDR is being handled, the level-2 system is incomplete and should not operate [2]. The following examples point out some actual technological inconsistencies:

- The driver can take his/her hands off the steering wheel. Under the UNECE R79 regulation, a maximum of 15 seconds is permitted before a first warning is provided, while a few seconds are sufficient for the driver to become OOL.
- The driver can activate level-2 ADAS functions while not holding the steering wheel. This demonstrates that the driver status is not being used to permit activation of the ADAS.
- Automated lane change (ALC) is triggered under the confirmation that the driver is activating the indicator and is holding the steering wheel. The combination of these two conditions does not guarantee that the driver is performing the OEDR.

The focus placed on robotizing the car has relegated the driver to a peripheral role. To guarantee engagement and safety, reactionary solutions are currently being used: for disengagement, DMS and for ALC, blind spot warning and intervention. Consequently, these safety oncost seem justified for achieving compliance with the regulations and reaching top ranking in safety assessment.

TECHNICAL LIMITATIONS

ADAS control is often seen as the computation of a trajectory to be followed by the vehicle based on a set of exogenous sensors. Although it is one of the prerequisites for partial automation (and automated driving in general), it is the control of the power steering that renders the ADAS as a reaction torque to the driver. This study assumes that a trajectory corresponding to the centerline of the driven lane is available and focuses on the steering control only.

A single control structure is applied virtually to all available steering systems (Figure 1a). It takes the form of a combination of an assistance control loop for manual driving and of an angle control loop for automated driving (ADAS). Weights are used to adjust the control authority. Higher weight on the angle control results in stronger centering support that is effective on straight drive as well as curves but tends to reject manual input. Whereas, lower weight prioritize manual intervention at the expenses of lower angular tracking performance. This steering control for ADAS, called *blended control*, is characterized by a tradeoff between tracking performance and manual intervention. Additionally, an override strategy is used to manage this conflict by switching the weights depending on the torque applied by the driver. While level-2 ADAS aim at reducing the driver workload and ensuring engagement, override biased the concept of "driver assistance" as the notion of working together by giving the impression that the driver can be replaced by the automation. The discontinuous operation of ADAS with the override strategy is assumed to be one of the causes of driver misuse and disengagement occurrence.

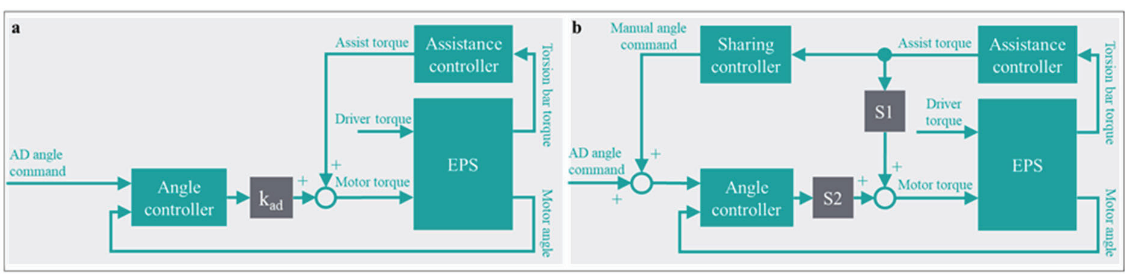


Figure 1. Two control structures for steering HSC. a. Blended control where the torque commands from the driver and the angle controller output are mixed. The gain K_{ad} is used to manage the conflict between driver and automation. b. Admittance control where the angle commands from the driver and the automation are mixed. The gains S1 and S2 are employed for setting the operating mode of the steering to manual, automated or shared mode.

HAPTIC SHARED CONTROL (HSC)

We argue that enhancing human-machine interaction with the steering system encourages driver engagement and is relevant in preventing misuse [5]. Continuous sharing of the steering control bonds the driver with the automation by mean of intuitive haptic communication. HSC of the steering enables manual intervention over the automation without deactivation. Two control configurations are available for HSC: blended control and admittance control (Figure 1). In some vehicles of the latest generation, driver engagement has been improved by lowering the angle controller weights of blended control enabling shared steering with the driver but with low trajectory tracking performance. Interestingly, this limited tracking performance is regarded as another contribution to better driver engagement because it refrains trust building on the capability of the automation. Nevertheless, dynamic adjustment of the weights of blended control as a function of the torque applied by the driver is technically challenging and consequently narrows the shared steering operation [6]. Rather than mixing the torque commands from the driver and the angle controller, admittance control mixes the angle commands. It features the advantage that the inner position control loop is purposefully made stiff so as to ensure high tracking performance in the absence of interaction. Conversely, the outer torque loop is naturally closed in the presence of interaction. The target position of the automation is corrected with an estimated manual displacement computed from a virtual admittance representing the steering system. Ideal lane centering function is achieved with admittance control based HSC because it enables driver intervention while ensuring high trajectory tracking performance [7]. Hence, admittance control is an alternative to blended control because it is not impaired by the tradeoff between tracking performance and manual intervention, and it enables continuous shared steering.

PROACTIVE ADAS

While conventional ADAS feature limited capability for manual intervention without deactivation and rely on reactionary DMS, proactive ADAS aim at enhancing driver interaction within an override free control framework. Rather than providing unconditional support unless misuse is detected, assistance is provided upon confirmation of engagement. This approach complies with the concept of assistance as working together and questions the relevance of DMS.

Active Lane Centering Assistance (aLCA)

Admittance control based HSC is appropriate for hands-free lane centering control as it enables driver intervention while providing high trajectory tracking performance. However, it is inappropriate for partially automated operation because of the potential risk increase of unjustified trust building on the capability of the automation. aLCA is based on the observation that a driver is tempted to take his/her hands off the steering wheel when sufficient safety margins are confirmed around the vehicle. It is assumed that this condition occurs primarily on straight drive and when the vehicle is well centered in the lane so that the driver is confident enough to let, intentionally or not, the automation take the drive. The condition relates to the lateral control of the vehicle as neither the driver nor the automation are needed to operate the vehicle. When the vehicle is centered in the lane, the driver does not need to apply torque and the output of the angle controller vanishes. aLCA consists in deactivating the centering support when the vehicle is well centered in the lane and when the driver applies no force. Conversely, it provides guidance only when the driver applies torque, similarly to power steering delivering torque assistance upon driver input. Hence, the driver is not tempted to take his/her hands off the steering wheel reducing the risks of misuse. While steering, the driver is continuously interacting with the ADAS providing intuitive haptic information in the form a torque directed towards the lane center (AD trajectory). In this way, the UNECE regulation 79 can be fulfilled without oncost related to DMS such as hands off detection (HOD). In the case of sustained hands-off driving with aLCA not providing the centering effort, steering based LKA remains to prevent lane departure like in manual operation. Figure 2 shows how the steering control switches between assisted and manual drive when using aLCA with the gain relation *S1+S2=1*. The torque threshold is small enough so that the switching is virtually seamless to the driver.

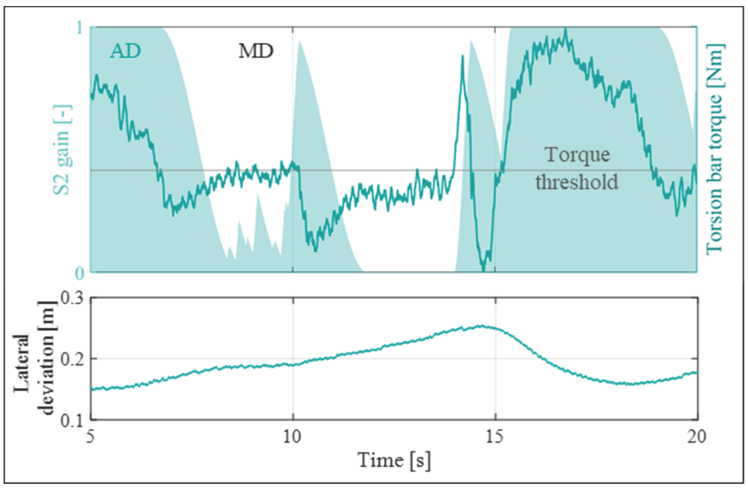


Figure 2: Operation of aLCA. Top figure: Relation between driver torque (torsion bar torque) and the control authority transfer between manual (MD) and centering assistance (AD) with the gain S2. Bottom figure: Vehicle lateral deviation from the planned trajectory. Between 7 and 13 s, the vehicle is relatively well centered, the driver activity is low and the centering support is inactive (MD). Because the driver hands are in contact with the steering wheel, engagement events are eventually occurring as observed at 10 s. From 13 s, the driver is actively steering trying to reduce the lateral deviation with the support of the centering function (AD).

Assisted Lane Change (aLC)

The aLC function requires the driver to manually initiate the lane change upon activation of the indicator. While steering, the driver encounters torque resistance at first. Overcoming this resistance triggers a trajectory shift to the adjacent lane. The driver is then guided towards the center of that next lane (Figure 3). Compared to the ALC function, driver engagement is enhanced because it is reasonable to assume that the driver will proceed to the OEDR during the initiation phase similarly to while changing lane manually. The need for safeguards like blind spot intervention becomes obsolete and safety oncost can be reduced.

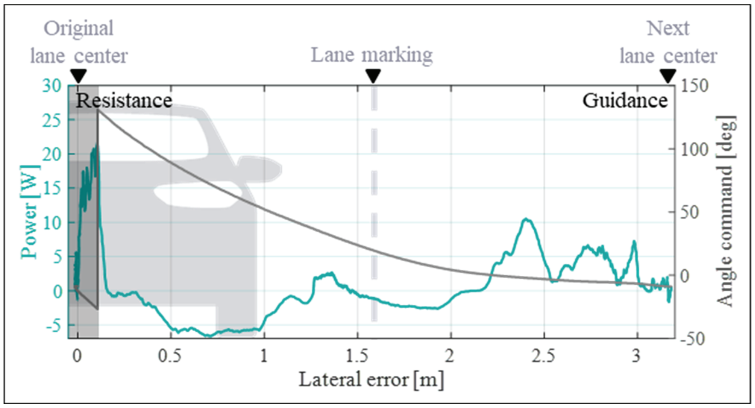


Figure 3: Steering system performance during a lane change maneuver while using the aLC function. Driver activity is required to deviate the vehicle away from the planned trajectory. When the lateral error increases above 0.2 m, the planned trajectory is shifted to the adjacent lane center (angle command trace). Guidance is then provided, which translates into negative driver power. As the vehicle gets closer to the center of the new lane, the driver proceeds to final adjustment of the lateral position (positive power from 2.2 m of lateral error).

ASSESSMENT OF DRIVER ENGAGEMENT

An evaluation made on a test vehicle across several individuals is presented to quantify the usefulness of the proposed proactive ADAS functions. The test vehicle is equipped with trajectory planning, tracking control and the steering shared control. The evaluation is made along a 1 km two-lane straight road portion. The nominal trajectory of the automation lies on the center of the right lane and the vehicle is controlled to track this nominal trajectory at 50 km/h using the Stanley path tracking model [8]. A high precision GNSS system (Global Navigation Satellite System) is used for the position feedback.

Six participants took part in the experiment with an average age of 44 (from 31 to 60) years old. All participants were experienced drivers and reported an average annual travel distance of 22,000 km. They were required to operate the steering wheel only and to keep their hands in contact with the steering wheel during the whole duration of the experiment. The drivers had to maintain a straight drive centered in the right lane for the assessment during lane centering and to perform a single lane change maneuver within a distance of about 100 m for the evaluation during lane change (Figure 4). Following accustomation with the different functions, each participant repeated the maneuvers three times in each operating mode of manual, conventional ADAS and proactive ADAS. The driver power, which is defined as the power input by the driver into the steering wheel, is employed for the assessment of engagement. Although there is no established way of measuring driver engagement, the choice of this indicator is motivated by its practical and non-intrusive aspects. The power input is a measure of how much involved is the driver in the operation of the lateral control of the vehicle. Positive power indicates that the driver is actively steering (driver active), while negative values represent situation where the human arms are driven by the automation through the steering wheel (driver passive). Inactivity is defined as when the power vanishes. Disengagement is not strictly captured with steering inactivity because zero power is obtained when either the torque or the velocity crosses zero, for example during a slalom maneuver. Furthermore, engagement cannot be assessed merely with the input power because it is a broader concept that is not limited to the physical contact of the driver with the steering wheel. Another acceptable level of engagement that satisfies the OEDR condition is obtained when the driver monitors the automation and the traffic (but this increases the risk of OOL as discussed in the Safety Paradox Section). Nevertheless, DMS are challenged in their intention of detecting the ill-defined concept of disengagement. For example, steering wheel touch sensor and camera for head or eyes gaze monitoring provide limited performance in assessing driver disengagement because neither hand in contact with the steering wheel nor the orientation of the driver sight are reliable indicators [9]. Hence because proactive ADAS are activated upon confirmation of physical activity, power is a pertinent and sufficient indicator to evaluate the driver engagement.

Technically, the driver power P_d is defined as $P_d = T_d\ \omega_{sw}$, where T_d is the driver torque input and ω_{sw} is the angular velocity of the steering wheel. Typically, power steering is equipped with a torque sensor in the form of a torsion bar. The measurement represents the torque difference at both extremities of that torsion bar, which is equivalent to the driver torque in steady state condition only. Estimation of the driver torque dynamics is obtained with an observer as detailed in [10]. While the input power captures driver activity, the lateral deviation away from the trajectory is another indicator, used in combination to assess how well the driver performs the task of controlling the vehicle position along the road. This deviation corresponds to the error between the planned trajectory and the actual position of the vehicle obtained from the GNSS.

Standard deviations of these indicators are used to quantify the driver performance. These were computed for each participant and averaged for the display in Figure 5 and 7.

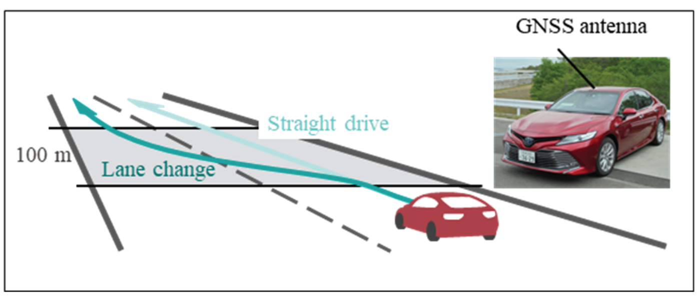


Figure 4. Test vehicle used for the assessment of the driver engagement. It is equipped with the Stanley path tracking model, the steering shared control and a high precision GNSS for the position feedback. Two driving scenarios are used for the assessment of the driver engagement during manual drive as well as when conventional and proactive ADAS are activated. The straight drive is used during lane centering. The driver is required to execute a lane change within a distance of 100 m for the assessment with lane change support.

Assessment during Lane Centering

Measurement examples of driver activity during manual drive, conventional LCA and the proposed proactive aLCA are shown in Figure 5. Technically, the driver can remain passive when using the LCA function because the vehicle is centered continuously. The remaining low activity observed is caused by the driver holding the steering wheel. Practically, it is only following a disengagement alert that the driver has to display activity temporary, such as touching or applying force on the steering wheel depending on the detection method. As shown in the top-right bar graph of Figure 5, less power is required for aLCA than during manual operation but higher than with LCA. This confirms that proactive aLCA reduces the driver workload compared to that measured during manual operation while ensuring a greater degree of engagement in comparison to that observed with conventional LCA. This assessment of engagement during lane centering correlates with the performance of tracking the lane center as displayed in the bottom-right bar graph of Figure 5. Hence, aLCA provides a compromise between LCA and manual operation. It reduces driver workload without impairing safety.

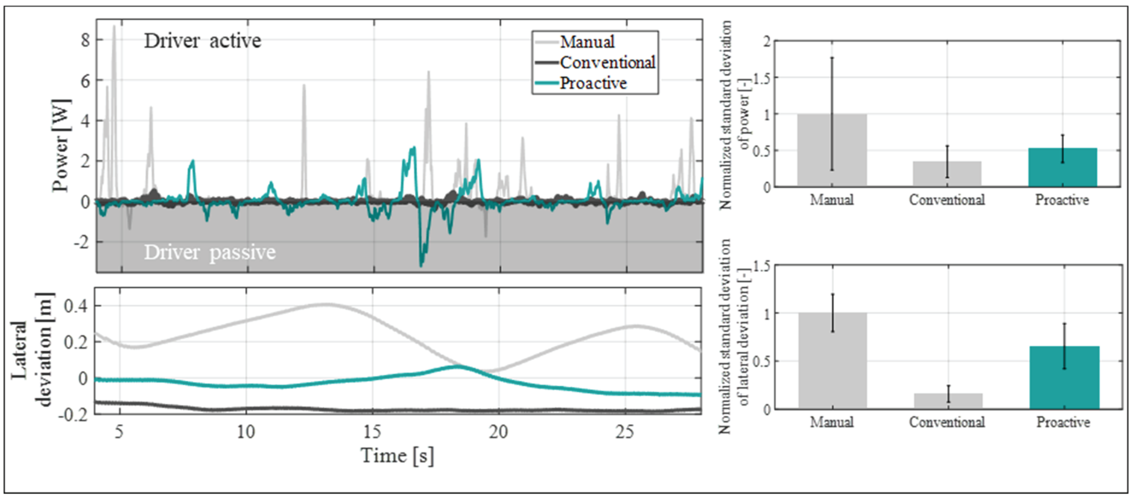


Figure 5. Assessment of driver engagement during lane centering. The time plots show a measurement example of one driver performance. The bar graphs give averaged standard deviations of the power and lateral error of the 6 participants including the one-sigma error. Virtually no driver activity is required when the conventional LCA function is activated. Less power is required while using the proactive aLCA function than during manual drive but more than with the LCA activated. A similar trend is observed for the accuracy of lane tracking. Greater engagement is observed when using the proactive aLCA function while still significantly reducing the workload.

Assessment during Lane Change

The data of three lane change maneuvers during manual drive as well as with conventional ALC and the proposed proactive aLC activated are depicted in Figure 6. Like most conventional ADAS, the ALC function does not require driver activity to execute the maneuver. Rather, the human is driven by the automation (negative power). However, driver activity is necessary to perform the lane change manually and with the aLC function. In this example, the aLC has been tuned so that the driver power required to initiate the lane change is even greater than that in manual operation. After this initial engagement, guidance is provided resulting in negative power until the vehicle approaches the center of the next lane where the driver proceeds to a final adjustment of the vehicle position (positive power, also observed with ALC). Figures 7 shows the standard deviation of the driver power. It has been split into two parts. The first part is composed of the data starting from the initiation of the indicator (assuming that the vehicle is centered in the initial lane) until a lateral deviation of 1 m is reached. The driver power of this portion of the maneuver is assumed to represents the initial engagement. The second part uses the data remaining until the vehicle attains the center of the new lane. Here, the driver power indicates the activity required to complete the maneuver. Considering the power for the manually executed lane change as baseline, the data of Figure 7 confirms that the lowest power required to the driver occurs while using the ALC function. In both, manual and with ALC, similar power levels are observed during the initial and guidance phases. Conversely, the aLC function features the compromise of ensuring initial driver engagement (highest initial power) while reducing the workload during the guidance phase (power lower than that in manual drive). The initial engagement level is tunable as shown with the markers in the left bar plot of proactive ADAS in Figure 7. Fine tuning of this initial power should be considered carefully. Indeed, a too high initial resistance might be misunderstood as a denial of the lane change, whereas a too low value would not guarantee engagement.

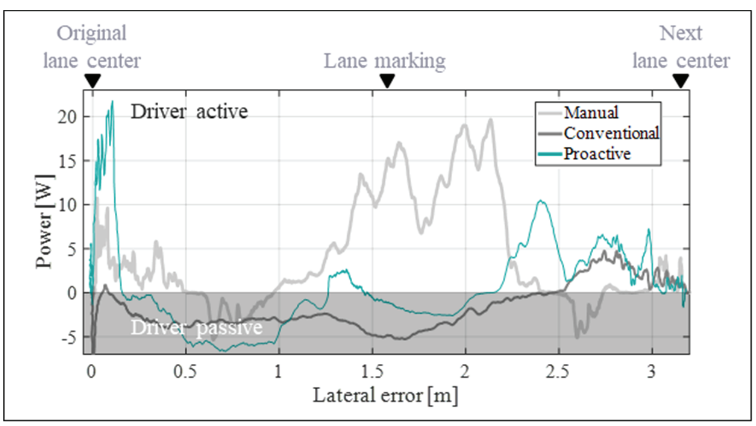


Figure 6. Measurement example of one participant over three lane changes: manual and with ALC and aLC activated. Driver activity is not required when the conventional ALC is activated resulting in negative power. The power required to initiate the lane change is highest with the proactive aLC (tuned as such) to force engagement. Then, as the AD trajectory shifts towards the next lane, guidance is provided and the driver becomes passive similarly to when using the ALC function.

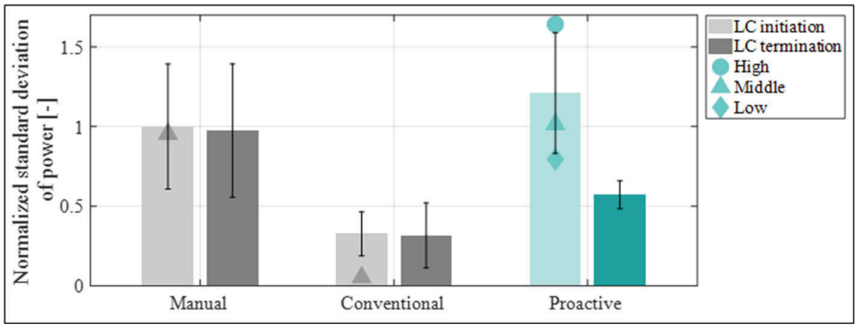


Figure 7. Standard deviations of the driver power over the lane change maneuvers (manual and with conventional ALC and proactive aLC activated). The experimental data is split into two parts: lane change initiation (from 0 to 1 m of lateral deviation) and termination (rest of the maneuver). Similar power levels at both initiation and termination phases are observed in manual operation and with conventional ALC respectively. However, they differ because virtually no driver activity is necessary when using the conventional ALC. The proactive aLC represents an alternative that guarantee initial engagement while reducing the workload during the termination phase. The markers indicate the standard deviation of one participant. Three different initial engagement levels of the proactive aLC function (low, middle and high) are displayed to demonstrate the tuning capability. The middle tuning has been used for the bar graphs.

CONCLUSION

Interestingly, most contributions on ADAS gravitates around a limited scope of human-machine interaction, which results in known safety challenges (OOL, disengagement, misuse, etc.). This paper is an attempt to demonstrate practically how enhanced human-machine interaction enables the development of ADAS that prioritize safety while still reducing significantly the driver workload. Override free and continuous ADAS are essential design requirements that enable intuitive haptic communication between driver and automation. Proactive ADAS, which provide assistance upon confirmation of driver engagement, enable the fulfillment of safety regulations (e.g. UNECE Regulation 79 for partial automation) without oncost related to DMS. In consequence, the relevance of DMS is questioned. The resulting cost reduction represents a significant contribution potential for the democratization of ADAS to all vehicle platforms so that the original road and traffic safety objective aimed by the automated driving technology can become reality.

REFERENCES

[1] 2021. “SAE Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles.” SAE Standard J3016, Rev. Apr. 2021.

[2] S. Joseph, U. Arun, F. Filip, S. Emily, and M. Neil. 2021. “Understanding misuse of partially automated vehicles - a discussion of NTSB’s findings of the 2018 Mountain View Tesla Crash.” Chartered Institute of Ergonomics & Human Factors, Tech. Rep., January.

[3] 2018. “Uniform provisions concerning the approval of vehicles with regard to steering equipment.” Economic Commission for Europe, E/ECE/TRANS/505/Rev.1/Add.78/Rev.4, Geneva, Switzerland, 18 October.

[4] 2023. “Assisted Driving, Highway Assist Systems, Test and Assessment Protocol.” EuroNCAP, Implementation 2023, Version 1.1.

[5] Mueller, A. S., Reagan, I. J., & Cicchino, J. B. 2021. “Addressing Driver Disengagement and Proper System Use: Human Factors Recommendations for Level 2 Driving Automation.” Design. Journal of Cognitive Engineering and Decision Making, 15(1), 3–27.

[6] Sakai, Y., T. Nakade, T. Tamura, and R. Fuchs. 2022. “Hi-Fi haptic shared control for automated steering.” in 15th International Symposium on Advanced Vehicle Control, September.

[7] Shoji, N., M. Yoshida, T. Nakade, R. Fuchs. 2019. “Haptic shared control of electric power steering – a key enabler for driver automation system cooperation.” In: Bertram T. (eds) Automatisiertes Fahren 2019. Proceedings. Springer Vieweg, Wiesbaden.

[8] Dominguez, S., A. Ali, G. Garcia, and P. Martinet. 2016. “Comparison of lateral controllers for autonomous vehicle: Experimental results,” in 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), 2016, pp. 1418–1423.

[9] 2022. “Effectiveness of driver monitoring systems.” AAA Report, February

[10] Moreillon, M., T. Tsutomu, and R. Fuchs. 2017. “Highly automated driving − detection of the driver’s hand on and off the steering wheel for ADAS and autonomous driving.” in 7th International Munich Chassis Symposium 2016, P. D. P. E. Pfeffer, Ed. Wiesbaden: Springer Fachmedien Wiesbaden, pp. 505–525.

EVALUATION OF INTERFACES FOR AUGMENTING A DRIVER'S ABILITY TO ANTICIPATE FRONT RISKS IN REAL TRAFFIC

Matti Krüger
Honda Research Institute Europe GmbH
Germany

Vanessa Krüger
Taisuke Mukai
Honda R&D Europe (Deutschland) GmbH
Germany

Paper Number 23-0307

ABSTRACT

Effective alerts are often subject to a tradeoff between relevance and utility. While it is easier to acknowledge the relevance of a warning about an imminent hazard than a more distant threat, the possibilities to act appropriately in response to notifications decrease with threat distance. To benefit from the advantages of early notifications without creating annoyance and ignorance, we introduce a variety of Human-Machine Interfaces that provide driver assistance by scaling stimulus saliency in accordance with the urgency of a front risk in traffic. Further, we report an initial investigation of the influence of the HMIs on measures of front collision risk and subjective driver experience after prolonged use in real traffic.

Three functional HMI prototypes were implemented in a roadworthy vehicle, equipped with additional hardware for front risk detection, stimulus presentation, assistance control, and data logging. Participants with advanced driving practice received these vehicles for 12 days in total for personal daily use, consisting of 3 guaranteed days of use for one of each HMI prototype and 3 days of driving without any added front risk notifications. Besides continuous logging of driving data and risk estimates, subjective data were acquired in the form of logbook entries and interviews.

Measures of driving safety were high across all conditions, indicating no occurrence of critical situations. No HMI specific safety effects on top of high baseline levels were observed. Subjective ratings show a trend for an increasing perceivability and usefulness of a sound-based HMI with extended system exposure. Participant feedback suggests that no such adaptations may have been necessary for the remaining vision-based HMIs because intuition could be gained quickly. Future HMI iterations should refine the balance between salience and subtlety to better align with actual safety levels while future investigations might benefit from longer individual exposures or an experimental control of safety levels.

INTRODUCTION

Driving creates a continuous demand for the driver's perceptual system. Traffic elements must be perceived and understood in time to allow for safe maneuver planning and collision avoidance. Yet, this requirement is not always being fulfilled so that around 90% of traffic accidents and critical situations are still attributed to human error [1–3], in particular to recognition failures [2]. To allow drivers to better recognize safety-relevant traffic elements, various driver assistance systems have been developed or proposed. Commonly available systems include forward collision warnings (FCW) [e.g., 4, 5], lane departure warnings [e.g., 6–9], blind spot warning systems [e.g., 10, 11], and parking assistance systems [e.g., 12, 13]. Some proposed systems, such as monitoring aids for laterally crossing traffic [e.g., 14–17], offer support in very specific situations that are known to be challenging. Furthermore, systems that monitor correlates of driver fatigue, workload, and attention [e.g., 18–21] are being introduced as means to account for variable mental states that may affect driving safety. Many of these assistance systems have a binary character: They are activated when a signaling threshold is surpassed and are otherwise silent. For example, a forward collision warning (FCW) may be triggered only once the time-to-collision (TTC) or another risk indicator [e.g., 22, 23] between the ego-vehicle and a front vehicle falls below a specific threshold. It then generates salient visual and auditory warnings to notify the driver about the predicted collision. One shortcoming of such notifications is the potential to disrupt ongoing attention processes [24] [see 25, p. 30], including attention on the driving task. Another issue is that warnings that are perceived as disturbing or unnecessary become likely to be ignored over time [see, e.g., 26, 27]. This creates a motivation to limit warnings to a minimum of cases for which the confidence about the necessity of the warning is high, i.e., situations in which it is certain that the driver appreciates the warning and situations in which a collision is imminent. The former case may be addressed by personalized assistance systems [28] and demand-based assistance triggered by an active user request [e.g., 17, 29, 30]. However, an automatic adaptivity carries the risk of being perceived as inconsistent while systems that rely on a user request are limited by the driver's ability to judge what situations require support.

Also a restriction to situations with an imminent collision risk comes with inherent challenges: From a driver's perspective, two issues of imminent threat conditions are that they leave only little time to react and are of little use for making foresighted maneuver plans. Binary warning systems are hence subject to a tradeoff between how much time they leave for driver reactions and their potential for disturbance [see 25, p. 31]. But this limitation of binary warnings does not necessarily persist for systems with ordinal or continuous properties [31, 32] and systems that provide non-critical warnings [33, 34]. For instance, so called *likelihood-alarm-systems* [31], which produce graded alerts that appear early and become more prominent with increasing event confidence, have demonstrated improvements in driver responses and system trust [31, 32]. Similarly, a proposed assistance that continuously encodes the temporal proximity between traffic participants in the strength of tactile stimuli has been rated as helpful and non-disturbing while improving driving safety [35–37]. As suggested by Krüger [25, pp. 54-56], systems with scalable output that is partially contingent on one's own actions may integrate well with existing sensory processing and possibly be utilized by drivers in a foresighted manner. Compared to binary signaling, utility may further be added through an inherent communication of dynamic aspects of a situation [see 25, p. 32], such as risk development over time instead of mere risk presence or "absence" information.

Accordingly, we developed and investigated assistance prototypes with a variety of human machine interfaces (HMI) that implement an early communication of front collision risks that is subtle at the beginning but can continuously become more salient when the risk increases. Here we introduce three of these HMIs, which use different means of encoding front risk. We aim to support drivers in integrating the conveyed information into their scene understanding to facilitate foresighted driving with a better avoidance of more dangerous situations. However, such an integration may take some practice and time to develop, especially when considering the initial stimulus subtlety at low risks. We hypothesize that extended experience of the HMI output in a wide range of situations will allow drivers to better pick up and use the encoded information. Specifically, we expect subjective measures of the experience with the HMIs to improve and the occurrence of situations that reach more critical levels to decrease over time with the introduced assistance. For this reason we decided to carry out evaluations in real traffic, where drivers are able to experience the HMIs for prolonged periods and in more diverse settings compared to a more controlled driving simulator environment. Following the introduction of three candidate HMIs, we describe a vehicle implementation of functional prototypes and report on a first user study about the effects of these HMIs on drivers in real traffic.

METHODS

Estimating Front Collision Risk

A continuous mapping from front collision risk to HMI activation requires an estimate of the front collision risk. For classical forward collision warnings the collision risk is typically estimated through the time-to-collision (TTC). The TTC is a very direct but also erratic predictor of collisions. It falls as another vehicle is approached but quickly jumps to infinity once a distance is stable or increasing. More indirect collision risks, such as those produced by tailgating are not adequately captured. Another temporal measure known as time headway (THW) describes the time at which the current position of a front vehicle will be reached. This makes it continuous in the presence of a front vehicle and sensitive to tailgating but insensitive to the velocity of the front vehicle.

To account for both direct and indirect front collision risk factors while ensuring risk continuity, we have chosen to estimate the front collision risk by a measure known as time to closest point of approach (TCPA) [38–40] within specified bounds. The TCPA interpolates between TTC and THW by a term a_L that expresses the potential deceleration of the leading vehicle at any time. For two vehicles that follow each other on the same trajectory, the TCPA may be obtained as follows.

$$t_L = -\frac{v_L}{a_L}. \tag{1}$$

$$TTC_a = \frac{-\Delta v - \sqrt{\Delta v^2 - 2a_L d}}{a_L}. \tag{2}$$

$$TCPA = \begin{cases} TTC_a, & \text{if } TTC_a < t_L \\ \frac{d - \frac{v_L^2}{2a_L}}{v_F}, & \text{otherwise.} \end{cases} \tag{3}$$

Here t_L stands for the potential stop time of the leading vehicle with velocity v_L, d is the spatial distance between both vehicles, v_F the velocity of the following vehicle, and Δv is the difference between v_L and v_F. Normalizing up to a threshold θ yields a simple collision risk estimate r that can be used to scale HMI output continuously. For all HMIs we have selected an activation threshold of $\theta = 7s$.

$$r = max\left(\frac{\theta - TCPA}{\theta}, 0\right). \tag{4}$$

HMI Candidates

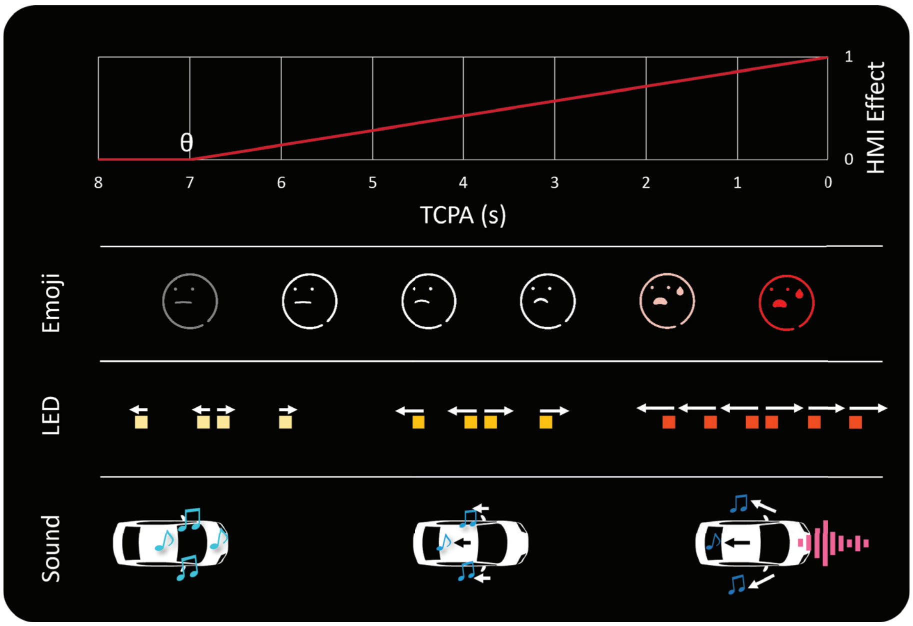


***Figure 1* - Illustration of the relationship between TCPA and the three described HMIs. Emoji: The face becomes more concerned with increasing risk and changes color at a high risk. LED: apparent movement of light dots from the risk center towards the periphery changes in frequency, speed, and color with increasing risk. Sound: Modulation of infotainment sound source location and presence increases with risk.**

We developed three HMI prototypes that utilize the described TCPA-based risk estimate to continuously adjust aspects of their output. Each HMI targets unique properties of human sensory processing to explore specific potential in driver assistance and allow for redundant or complementary future utilization[1].

Emoji HMI: In the mobility domain only little time can be spent on the perception and processing of individual areas of interest because information can become outdated in proportion to the velocity of oneself and other traffic participants. Accordingly, an HMI to convey existence and magnitude of a collision risk should facilitate fast and intuitive perception and understanding to minimize interference with the driving task. HMI output in conditions of low risk should further be subtle and be easy to ignore.

For most people, vision is the primary sensory modality in various tasks, especially in those related to mobility [42, 43]. For some types of visual stimuli there are highly specified processing networks that favor and speed up processing and recognition of such stimuli. One prominent form of preferred stimulus are human faces. Already as infants humans start to preferably seek out and focus on human faces [44, 45]. The specialization on faces [46, 47] even leads to commonly experienced forms of pareidolia [48], i.e., a perception of faces on objects and patterns when there are none. Faces are also a primary identifier of individuals and effectively convey most emotions within 200 ms [49], and even in a stylized form such as through cartoons [50, 51].

The first HMI prototype aims to exploit the speed and intuition of emotion processing through faces. It consists of a transparent emoji with a size of $2°$ of visual angle that appears in the head-up display and displays emotions

[1]In addition to the HMIs specified here, an investigation of a different set of HMIs for encoding the same underlying risk in a driving simulator environment is reported by Matsuoka et al. [41]

that are mapped to a range of risk values. Starting at a low risk threshold, the face gradually becomes more concerned and eventually transitions into expressions of fear and finally panic at high risk. For high risk values the color of the face further changes from white to red. The transitions are reversed when the risk falls and briefly even change into a smile once a safe TCPA is reached. By being positioned in the head-up display, the face is almost aligned with the movement direction so that it requires only short saccades to come into focus most of the time. Nevertheless, while a transition between emotions may be picked up peripherally, a reliable recognition of the current emotion depends on an actual fixation. Due to this property the information provided by this HMI must be actively sampled by a driver on demand. This makes it unobtrusive but also suggests that it introduces competition with other regions of interest that must be monitored during driving.

LED HMI: The second HMI aims to avoid a need for active fixations and competition with existing regions of interest. To achieve this it targets the visual periphery rather than the center. The periphery of the retina and hence the visual field is not capable of processing shapes in detail. However, it is as sensitive to moving stimuli as the retinal center [52, 53] and also able to resolve the direction of movement [54]. This capability allows us to detect moving objects in the environment and also perceive self-movement through optical flow. Optical flow during movement is characterized by a radial expansion of visual elements from the direction of motion to the periphery. Similarly, approaching objects generate a radial expansion from a stable center. The second HMI aims to make use of this property to augment the visibility of approaching risks. It is realized through an LED strip that spans across the whole bottom side of the windshield. When an object in the front is considered to be a risk, the LED strip creates repeatedly appearing dots that move outwards from the direction of the risk, thereby augmenting the optical flow associated with the risk on one dimension. With increasing risk the speed and spatial frequency of the movement pattern increases. Additionally, a gradual color change of the movement pattern from white to yellow to red is applied according to the risk level. The LED movement pattern has the theoretical advantages of being recognizable by peripheral vision alone and of guiding the user towards the direction of risk by tracing the origin of visual expansion. But in contrast to the emoji HMI, individual levels of flow have no direct absolute correspondence to emotions.

Sound HMI: The first two HMIs both relied on the addition of visual stimuli to convey risk information. Interference with the visual demands of the driving tasks was designed to be minimal either by reliance on the fast human visual processing of faces and emotions or by avoiding a need for foveal vision. Nevertheless, HMIs that add visual stimuli in a visual task are at the risk of eventually becoming distracting in situations of high sensory load. The third HMI therefore focuses on the potential to make a driver more aware of risks by reducing the efficacy of distracting stimuli. One typical source of stimuli that are not directly relevant to driving safety is a vehicle's infotainment system. In the event of an approaching collision risk the driver's attention should be focused on the driving task as much as possible. To facilitate that focus, this HMI reduces the perceived presence and salience of audible infotainment elements, such as the radio sound, in proportion to the measured risk level. Using a combination of auditory filters and variations in speaker balance and volume, an effect is created that leads to infotainment sound appearing to retreat to a more distant location opposite to the direction of the approaching risk. This modulation is intended to reduce infotainment salience but not take away the driver's autonomy to voluntarily select stimuli to attend to. A reduction in risk reverses the modulation and implicitly rewards the driver with infotainment sounds of improved presence and clarity.

The described form of assistance relies on the presence and activation of an infotainment system. A second component of this HMI therefore acts independent of the availability of infotainment sounds. This second component consists of the generation of an artificial ambient "noise" sound that becomes more prominent with increasing risk. This noise sound contains oscillating low frequency components for a loose association with objects in motion as well as slightly dissonant components that are intended to create a feeling of uneasiness. It is set to appear from the direction of the risk to draw focus to that direction. At low risks the noise is faint with a low and narrow frequency spectrum. With increasing risk it becomes more prominent and localizable through added volume and a gradual activation of higher frequency components that facilitate identification of interaural intensity and spectral differences [55–57]. A reduction in risk reverses the prominence of the noise component and eventually makes it disappear. In contrast to the infotainment modulation, the onset of this noise is set to a higher risk (5 seconds TCPA instead of 7 seconds). This, coupled with its directionality, makes it more a complementary escalation step in HMI output than a redundant stimulus. Conceptually the infotainment modulation clears the space that is then occupied by the noise and the risk it represents.

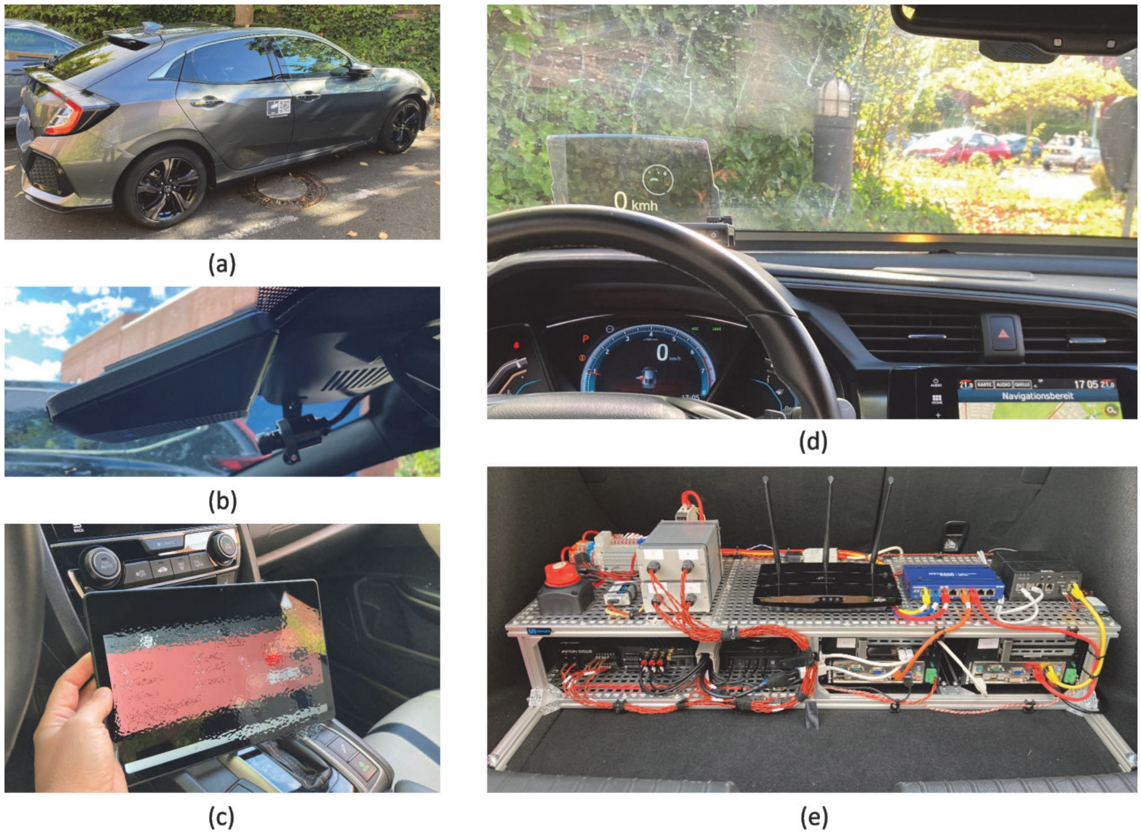


Figure 2 - **Overview of the vehicle setup for functional prototype realization. (a) 2017 Honda Civic, (b) aftermarket computer vision system and scene camera, (c) user interface for HMI parameter control, (d) head-up display with concerned emoji on top of the dashboard and LED strip with intermediate risk light pattern below the windshield, (e) hardware for computation, network, power-, and sound control installed in the car trunk.**

Vehicle Setup

Functional prototypes for the three HMIs have been implemented in a 2017 Honda Civic with automatic transmission. Figure 2 shows an overview of the installed components. For sensing of external risk-relevant measures, such as object presence, distance, and relative velocity, an aftermarket open protocol computer vision system (Mobileye Global Inc.) was mounted to the top of the windshield next to a front facing camera for optional scene recording. Measures from the ego-vehicle are accessed through a CAN read-only interface (Vector Inc.). Data integration, risk calculation, HMI control, and data logging are carried out on two PCs mounted in the trunk of the car together with a Wi-Fi router, sound card, amplifier, and voltage distributor. The SOLID software platform (Usaneers GmbH) is used as a backend for data integration, risk calculation, logging, and HMI control. Access to HMI toggles and parameters is realized through a custom Android app (Usaneers GmbH) installed on a tablet that connects to an on-vehicle Wi-Fi network.

For the emoji HMI an aftermarket head-up display (Maxwin) was installed on top of the dashboard. For the LED HMI an RGB LED strip was installed below the windshield, ranging across the full width of the windshield base. To realize the sound HMI, sound signals from the vehicle's amplifier are relayed to a sound card via a Hi-Lo converter. The signals are then modulated according to the current risk and output via the vehicle's original 7.1 speaker system. This allows us to modulate any sound generated through the vehicle's infotainment system, including the radio sound. Technical elements and cables are hidden within the car trunk and behind components of the interior to maintain the look and feel of a regular vehicle and reduce the risk for interference of the setup with the driving task and safety.

User study

We used the described vehicle setup in a first user study to investigate the development of effects of each HMI on driving experience and safety-relevant driving behavior and to find potential issues that could inform future system refinements. Specifically, the study should investigate the following research questions for each HMI:

1. Does driving safety with respect to front collision risks improve with HMI use compared to a classical forward collision warning?
2. Does driving safety, in terms of front collision risk, increase with HMI exposure over time?
3. Does the subjective HMI experience improve with HMI exposure over time?

We wanted drivers to put their main focus on the driving task rather than on the HMIs. Further requirements were an extended exposure to each HMI for multiple days and a natural variability in driving situations. To that end, we opted for a study setup in which participants would use the vehicle as a temporary substitute for their personal vehicle for multiple consecutive days. By using the vehicle in daily driving, such as for commuting, we hoped to facilitate the participants' natural driving behavior.

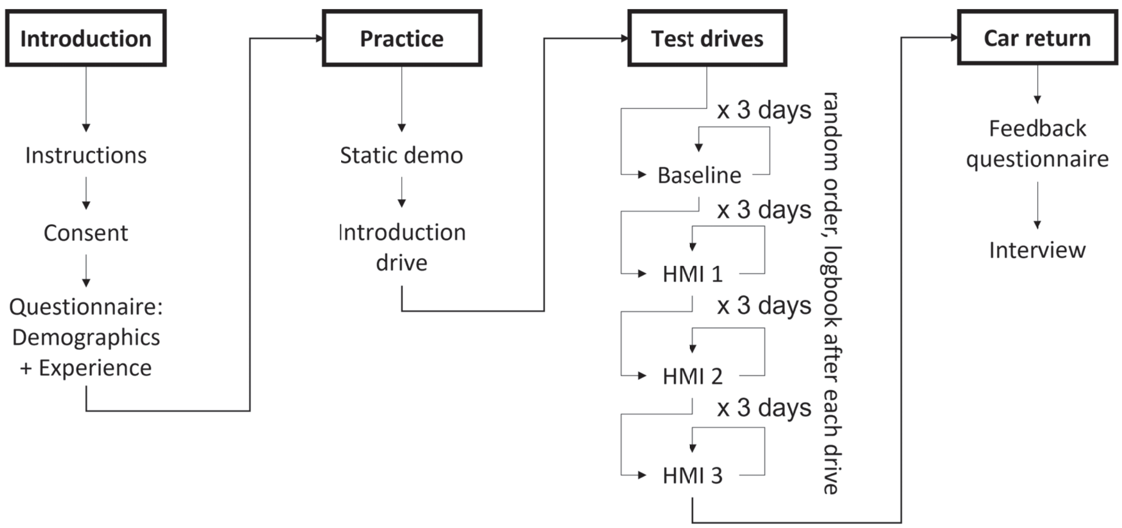


***Figure 3* - Study procedure for each participant.**

Procedure: Before the start of the study, each participant received an introduction that explained the procedure and conditions of the study. After giving written consent about their understanding of the study and the permission to use their collected data, participants completed a short questionnaire about their demographic data and driving experience. Next, a 30 minute introductory session was conducted, consisting of a demonstration of each HMI in a standing vehicle, followed by an accompanied introductory drive with each HMI.

For the subsequent testing period, participants were asked to drive for at least three days with each HMI. As a baseline condition, participants further drove for three days without any of the new HMIs. Instead, only a classical late forward collision warning (FCW) was available, consisting of a salient visible and auditory alert at a higher risk level. Participants were instructed to fill out a logbook after each drive, which contained fields to indicate their experience in the interaction with the car and the HMI. At the end of the last drive, participants were further asked to comment on their overall experience with the HMIs and general comments and remarks. In addition to these subjective evaluations, video data of the front scene and driving data required for risk calculation, such as front and ego vehicle location, relative velocity, acceleration, and braking, were recorded during each drive. The described study procedure has been evaluated and approved by the Honda R&D Bioethics Committee.

Participants: Six participants (aged between 30 and 59 years, one woman) took part in the study. All participants are experienced drivers with an annual mileage of at least 15000 km and certified advanced qualifications for handling vehicles in critical situations. This added qualification reduced the risk of being unable to safely handle the vehicle in cases of system failure at the expense of reducing generalizability to a more diverse population. All participants had a valid driver's license and normal or corrected-to-normal vision.

Dependent Measures: Because of the real traffic setting of the study the participants were not subjected to any specific predetermined traffic situations. The occurrence of front collision risks was out of experimental control and depended on the increasing chance for a natural occurrence of HMI-relevant conditions with prolonged exposure. This design choice makes safety evaluations through, e.g., activation-frequency measures uninformative. Instead we looked at measures that can be obtained within periods of HMI-relevant conditions, i.e., situations with a TCPA below the 7 s HMI activation threshold.

As a first measure of driving safety within those periods we chose the minimum TCPA. For each HMI-relevant period this minimum TCPA expresses how dangerous the corresponding situation became at most. Besides this measure of risk magnitude we were interested in how well drivers would prevent such situations from developing into situations with medium or higher risk. We defined medium risk as the midpoint between HMI onset and a collision, i.e., at a TCPA of 3.5 s. The inverse of the ratio between medium risk situations and all HMI-relevant situations indicates the success in risk prevention. To assess the subjective driving experience we analyzed interview responses and 7-point logbook ratings on HMI usefulness, annoyance, difficulty, and naturalness.

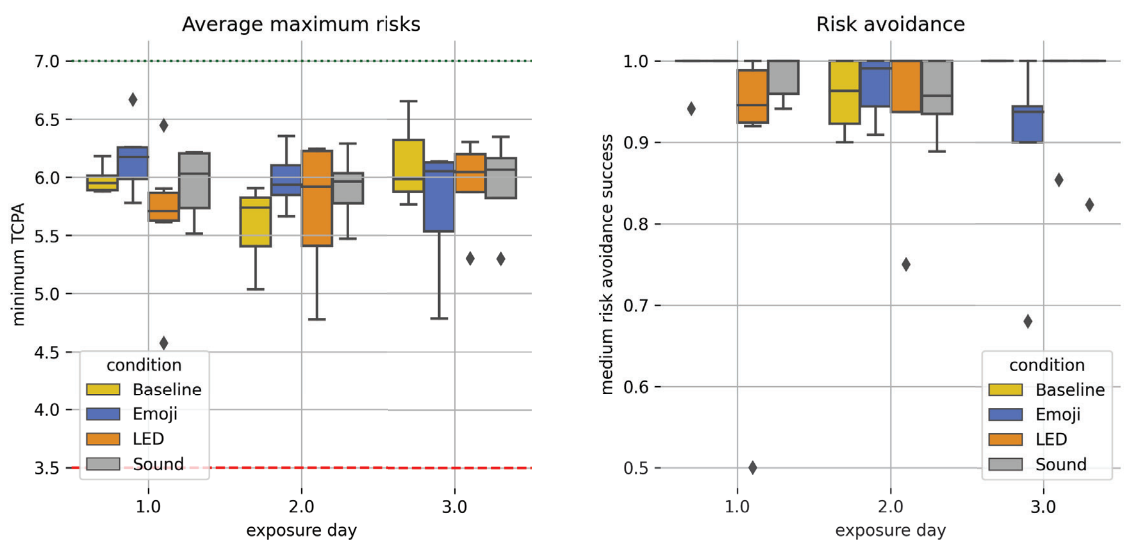
RESULTS

Safety

Table 1.
Occurrence of HMI-relevant conditions during free driving

condition	count	mean	std	min	50%	max
Baseline	6.0	112.00	65.22	30.0	129.0	194.0
Emoji	6.0	142.50	60.55	57.0	142.5	243.0
LED	6.0	150.16	46.01	88.0	163.5	201.0
Sound	5.0	199.20	58.99	131.0	187.0	284.0

For each condition, participants entered HMI activation ranges more than 100 times in total on average and 30 times at minimum (see Table 1 for further details). This suggests that the requirement for a regular natural occurrence of HMI exposure could be met and that a statistical analysis of the targeted safety effects is possible.



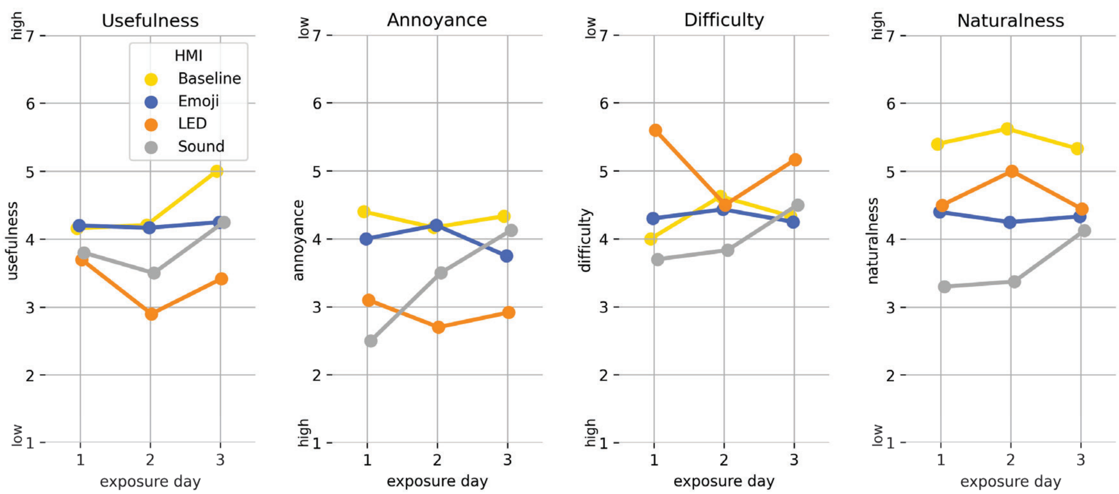
(a) *(b)*

Figure 4 - **Safety measures within periods of HMI-relevant conditions. 4a: Distribution of maximum risks quantified inversely (high value → low risk) as minimum TCPA for each condition and day of HMI exposure. 4b: Distribution of ratios for the avoidance of medium risks (TCPA ≤ 3.5 s).**

For all HMIs and exposure days the minimum TCPA values within relevant periods were on average between 5.6 and 6.1 s (min and max ranges per HMI - Baseline: 5.03-6.65; Emoji: 4.78-6.66; LED: 4.57-6.44; Sound: 5.29-6.35). This means that the highest risks were, on average, still close to the 7 s TCPA threshold, i.e., in a range that can still be considered as rather safe. A two-way ANOVA with the factors condition (HMI) and day of exposure indicated no main effect for condition ($F(3,49) = 0.76$, $p = 0.52$), day of exposure ($F(2,49) = 0.65$, $p = 0.52$), or their interaction $F(6,49) = 1.19$, $p = 0.32$. Participants typically drove safe, irrespective of whether one of the newly introduced HMIs or only a classical FCW was available. This also means that there was almost no HMI activation at all in the baseline condition because the FCW onset threshold lies much lower.

The high level of driving safety is also reflected in the measure of risk avoidance. Risk avoidance, defined as a minimum TCPA above 3.5 s, was high for all HMIs and exposure days with an average success rate between 0.88 and 1.0 (min and max ranges per HMI - Baseline: 0.9-1.0; Emoji: 0.68-1.0; LED: 0.5-1.0; Sound: 0.82-1.0). A two-way ANOVA with the factors condition (HMI) and day of exposure indicated no main effect for condition ($F(3,49) = 0.98$, $p = 0.4$), day of exposure ($F(2,49) = 0.07$, $p = 0.93$), or their interaction $F(6,49) = 1.3$, $p = 0.27$. The data therefore provide no answers to the first two research questions concerning differences in safety compared to a classical FCW or changes in safety with increasing exposure.

Subjective Experience



(a) *(b)* *(c)* *(d)*

Figure 5 **- Averaged subjective ratings on perceived usefulness, annoyance, difficulty, and naturalness for each HMI and day of exposure.**

Logbook Ratings: The subjective experience, quantified in terms of perceived usefulness, annoyance, difficulty, and naturalness, was assessed through logbook ratings, which participants were asked to provide after each drive. This request was met in 76% of all cases, yielding an average of 1.2 ratings per participant, condition, and exposure day. Due to this small sample size no inferential statistics on the ratings are provided. Figure 5 shows the average experience ratings for each HMI and day of exposure.

Usefulness of all HMIs including Baseline (FCW) was initially rated neutrally (Figure 5a). While ratings were constant over days of exposure for the Emoji HMI, the perceived usefulness for LED and Sound HMI dropped on day two but returned to neutral or slightly positive levels on the third day. Only the Baseline (FCW) ratings increased from neutral to positive on day three.

The LED and Sound HMIs were initially perceived as slightly annoying (Figure 5b). But while LED ratings remained low, the Sound HMI appears to have lost its perceived annoyance over time. Baseline and Emoji were rated neutrally for all days.

Difficulty ratings (Figure 5c) indicate whether participants found it difficult to notice and interpret HMI output. All except the LED HMI were assigned neutral difficulty ratings on the first exposure day while the LED HMI was easy to perceive on the first and the third but not the second day. Over time the sound HMI appears to have become easier to perceive whereas ratings stayed similar for Emoji and Baseline conditions.

The temporal upward trend of the Sound HMI also repeats for naturalness ratings (Figure 5d). It was perceived as slightly strange on the first two days but shifted towards more neutral ratings on day three. Emoji and LED HMI were perceived as more natural from day one. Baseline drives were perceived as the most natural. In these drives HMI output could only appear in situations of high risk, making them mostly identical to unassisted driving (see Section 4.1).

In summary, the subjective HMI experience only appears to vary consistently for the Sound HMI in line with our third research question. This variation is characterized by an initially slight rejection and perception difficulty that develops into a more positive and accessible perception with added exposure. LED HMI ratings are less consistent, showing no clear recovery from negative ratings despite low difficulty and high naturalness ratings. The Emoji HMI was consistently rated slightly above neutral and not difficult or unnatural to perceive. Baseline drives were mostly rated neutrally but also as the most natural, in line with the absence or rarity of HMI output.

Interview Feedback: Final interviews with participants provided some general insights on the common approach of the newly introduced HMIs. The different levels of expression that scaled with risk were regarded as a desirable property and were reported as helping in understanding the front risk reference. However, it was not universally understood whether situations were already serious at HMI onset, especially when initial stimuli were strongly perceived. An overly strong onset was further regarded as potentially distracting.

For the Emoji HMI it was appreciated that information could be accessed through fixations on demand without being obtrusive. However, the requirement for a fixation and the low salience were also seen as weaknesses. Some participants requested the Emoji to disappear more quickly after risk reduction while others would have preferred them to stay longer, suggesting potential in customization. One participant reported that Emoji appearance at HMI onset was more noticeable than Emoji changes at increasing risk.

The LED HMI was described as very recognizable and intuitive. It could capture a drivers attention also peripherally, e.g., while looking at the navigation screen, suggesting a better compatibility with the largely vision-guided driving task. However, it was also interpreted as urgent soon after onset and as drawing attention too strongly.

For the Sound HMI it was reported that the reference to a developing risk became easy and intuitive over time. However, a temporal impairment to listen to the radio was partially seen as a weakness. Furthermore, perceivability was reported to vary depending on the radio program and driving speed.

CONCLUSIONS

Here we introduced three human-machine interfaces (HMI) for a driver assistance designed to inform a driver about the presence and estimated urgency associated with a detected front collision risk. The HMIs continuously scale aspects of their output (one auditory, two visual) with the risk level to realize a risk communication that starts informing a driver much earlier than classical warning systems to facilitate foresighted driving. Functional prototypes of the interfaces were realized in a vehicle for further investigation in real traffic. We carried out a user study to investigate effects of the HMIs on driving safety and subjective user experience, as well as the potential role of the amount of HMI exposure.

Driving safety, in terms of front collision risk, was high across all conditions, including largely unassisted driving and irrespective of the amount of exposure. A safety-level benefit on top of high baseline levels was not observed. If such effects should exist, more data would likely be required to detect significant differences within low risk ranges and to detect any exposure-related safety effects. Alternatively, more controlled experimental settings that artificially provoke situations with higher risks or situations that are less predictable could facilitate an understanding of HMI effects on safety and driver behavior in general.

On the subjective experience level exposure-related benefits on ratings for usefulness, annoyance, difficulty, and naturalness were observed for a sound-based HMI. No such effects were observed for the two visual HMIs. However, a potentially higher initial intuition for these HMIs, which are exploiting specific pre-existing aspects of human visual processing such as optical flow and emotion recognition, might account for a reduced need for initial adaptations.

Overall the need for both sufficient stimulus salience to appropriately capture a driver's attention and initial subtlety to avoid causing driver distractions remains a challenge. Participant feedback suggests that this may be partially addressed by parameter adjustments and personalization. Investigations of future HMI refinements should further adjust the experimental design to promote a detection of potential effects on driving safety.

REFERENCES

[1] John R Treat et al. *Tri-level study of the causes of traffic accidents: final report. Executive summary.* Tech. rep. Indiana University, Bloomington, Institute for Research in Public Safety, 1979.

[2] Santokh Singh. *Critical reasons for crashes investigated in the national motor vehicle crash causation survey*. Tech. rep. Stats. Report No. DOT HS 812 115. US National Highway Traffic Safety Administration, 2015.

[3] Statistisches Bundesamt. "Verkehrsunfälle". In: *Fachserie 8 Reihe 7* 8.7 (2020).

[4] Robert Graham. "Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application". In: *Ergonomics* 42.9 (1999), pp. 1233–1248. DOI: 10.1080/001401399185108.

[5] B. Stein, Gideon P., Dan Rosenbaum, and Amiad Gurman. "Method and system for forward collision warning". EN. EP2993654. EP2993654. 2016.

[6] Sumio Motoyama, Takashi Ohta, Takeshi Watanabe, and Yoshihito Ito. "Development of lane departure warning system". In: *7TH World Congress on Intelligent Systems*. 2000.

[7] Shinnosuke Ishida and Jens E Gayko. "Development, evaluation and introduction of a lane keeping assistance system". In: *IEEE Intelligent Vehicles Symposium, 2004*. IEEE. 2004, pp. 943–944. DOI: 10.1109/IVS.2004.1336512.

[8] Frank Beruscha, Klaus Augsburg, and Dietrich Manstetten. "Haptic warning signals at the steering wheel: A literature survey regarding lane departure warning systems". In: (2011).

[9] Mathieu Deroo, Jean-Michel Hoc, and Franck Mars. "Effect of strength and direction of haptic cueing on steering control during near lane departure". In: *Transportation research part F: traffic psychology and behaviour* 16 (2013), pp. 92–103. DOI: 10.1016/j.trf.2012.08.015.

[10] Christopher Hinninger, Makoto Murata, Masahito Shingyoji, Fumie Nakamura, and Yasuhiko Fujita. *Vehicle blind spot detection and indicator system*. US Patent 8,310,353. Nov. 2012.

[11] Toshihiro Hashimoto et al. "Benefits of intuitive auditory cues for blind spot in supporting personalization". In: *National Highway Traffic Safety Administration (NHTSA)* ().

[12] Nobuhiko Yasui and Atsushi Iisaka. *Parking assistance system*. US Patent 6,483,429. Nov. 2002.

[13] Chunxiang Wang, Hengrun Zhang, Ming Yang, Xudong Wang, Lei Ye, and Chunzhao Guo. "Automatic parking based on a bird's eye view vision system". In: *Advances in Mechanical Engineering* 6 (2014), p. 847406.

[14] Cuong Tran, Karlin Bark, and Victor Ng-Thow-Hing. "A left-turn driving aid using projected oncoming vehicle paths with augmented reality". In: *Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications*. 2013, pp. 300–307. DOI: 10.1145/2516540.2516581.

[15] Alexandria M Noble, Thomas A Dingus, and Zachary R Doerzaph. "Influence of in-vehicle adaptive stop display on driving behavior and safety". In: *IEEE transactions on intelligent transportation systems* 17.10 (2016), pp. 2767–2776. DOI: 10.1109/tits.2016.2523438.

[16] Martin Heckmann, Dennis Orth, and Dorothea Kolossa. ""Gap after the next two vehicles": A Spatio-temporally Situated Dialog for a Cooperative Driving Assistant". In: *Speech Communication; 13th ITG-Symposium*. VDE. 2018, pp. 1–5.

[17] Martin Heckmann, Heiko Wersing, Dennis Orth, Dorothea Kolossa, Nadja Schömig, and Mark Dunn. "Development of a cooperative on-demand intersection assistant". In: *International journal of automotive engineering* 10.2 (2019), pp. 175–183. DOI: 10.20485/jsaeijae.10.2_175.

[18] Oskar Palinko and Andrew L Kun. "Exploring the effects of visual cognitive load and illumination on pupil diameter in driving simulators". In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. 2012, pp. 413–416. DOI: 10.1145/2168556.2168650.

[19] Gulbadan Sikander and Shahzad Anwar. "Driver Fatigue Detection Systems: A Review". In: *IEEE Transactions on Intelligent Transportation Systems* 20.6 (2019), pp. 2339–2352. DOI: 10.1109/TITS.2018.2868499.

[20] Honda Motor Company. *Honda Driver Attention Monitor*. [Online; accessed 06-2021]. 2021.

[21] Honda Motor Company. *Honda Unveils Next-generation Technologies of Honda SENSING 360 and Honda SENSING Elite*. [Online; accessed 12-2022]. Dec. 2022.

[22] Julian Eggert and Tim Puphal. “Continuous risk measures for ADAS and AD”. In: *Future Active Safety Technology Symposium*. 2017.

[23] Julian Eggert and Tim Puphal. “Continuous Risk Measures for Driving Support”. In: *International Journal of Automotive Engineering* 9.3 (2018), pp. 130–137. DOI: 10.20485/jsaeijae.9.3_130.

[24] Neil Lerner et al. *Driver-vehicle interfaces for advanced crash warning systems: Research on evaluation methods and warning signals*. Tech. rep. NHTSA, 2015.

[25] Matti Krüger. “An enactive approach to perceptual augmentation in mobility”. PhD thesis. Ludwig-Maximilians-Universität München, 2022. DOI: 10.5282/edoc.30558.

[26] Shlomo Breznitz. *Cry wolf: The psychology of false alarms*. Psychology Press, 1984.

[27] David J Getty, John A Swets, Ronald M Pickett, and David Gonthier. “System operator response to warnings of danger: A laboratory investigation of the effects of the predictive value of a warning on human response time.” In: *Journal of experimental psychology: applied* 1.1 (1995), p. 19. DOI: 10.1037/1076-898x.1.1.19.

[28] Martina Hasenjäger, Martin Heckmann, and Heiko Wersing. “A survey of personalization for advanced driver assistance systems”. In: *IEEE Transactions on Intelligent Vehicles* 5.2 (2019), pp. 335–344.

[29] Matti Krüger and Christiane Wiebel-Herboth. “Gaze-guided communication for assistance in mobility”. U.S. pat. 10543854B2. Jan. 2020.

[30] Chao Wang, Matti Krüger, and Christiane B. Wiebel-Herboth. ““Watch out!”: Prediction-Level Intervention for Automated Driving”. In: *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI ’20. Virtual Event, DC, USA: Association for Computing Machinery, 2020, pp. 169–180. ISBN: 9781450380652. DOI: 10.1145/3409120.3410652.

[31] Robert D Sorkin, Barry H Kantowitz, and Susan C Kantowitz. “Likelihood alarm displays”. In: *Human Factors* 30.4 (1988), pp. 445–459. DOI: 10.1177/001872088803000406.

[32] John D Lee, Joshua D Hoffman, and Elizabeth Hayes. “Collision warning design to mitigate driver distraction”. In: *SIGCHI Conference on Human factors in Computing Systems*. ACM. 2004, pp. 65–72. DOI: 10.1145/985692.985701.

[33] Jens Hoffmann and Jens E Gayko. “Fahrerwarnelemente”. In: *Handbuch Fahrerassistenzsysteme*. Springer, 2012, pp. 343–354. ISBN: 978-3-8348-8619-4. DOI: 10.1007/978-3-8348-8619-4_25.

[34] Frederik Naujoks, Andrea Kiesel, and Alexandra Neukum. “Cooperative warning systems: The impact of false and unnecessary alarms on drivers’ compliance”. In: *Accident Analysis & Prevention* 97 (2016), pp. 162–175. DOI: 10.1016/j.aap.2016.09.009.

[35] Matti Krüger, Christiane B. Wiebel-Herboth, and Heiko Wersing. “Approach for Enhancing the Perception and Prediction of Traffic Dynamics with a Tactile Interface”. In: *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI ’18*. ACM. New York, New York, USA: ACM Press, 2018, pp. 164–169. ISBN: 9781450359474. DOI: 10.1145/3239092.3265961.

[36] Matti Krüger, Christiane B. Wiebel-Herboth, and Heiko Wersing. “The Lateral Line: Augmenting Spatiotemporal Perception with a Tactile Interface”. In: *Proceedings of the Augmented Humans International Conference*. AHs ’20. ACM. Kaiserslautern, Germany: ACM Press, 2020. ISBN: 9781450376037. DOI: 10.1145/3384657.3384775.

[37] Matti Krüger, Tom Driessen, Christiane B. Wiebel-Herboth, Joost CF de Winter, and Heiko Wersing. “Feeling Uncertain - Effects of a Vibrotactile Belt that Communicates Vehicle Sensor Uncertainty”. In: *Information* 11.7 (2020). ISSN: 2078-2489. DOI: 10.3390/info11070353.

[38] Kearon J. “Computer programs for collision avoidance and traffic keeping”. In: *Conference on Mathematical Aspects on Marine Traffic*. Academic Press, 1977.

[39] H. Imazu and T Koyama. “The Determination of Collision Avoidance Action”. In: *Journal of Japan Institute of Navigation* 70 (1984), pp. 31–37.

[40] Hoong Chor Chin and Ashim Kumar Debnath. “Modeling perceived collision risk in port water navigation”. In: *Safety Science* 47.10 (2009), pp. 1410–1416. ISSN: 0925-7535. DOI: 10.1016/j.ssci.2009.04.004.

[41] Takahiro Matsuoka, Tsuyoshi Nojiri, Vanessa Krüger, and Matti Krüger. “Novel Interfaces that Enhance a Driver’s Ability to Perceive Forward Collision Risks”. In: *27th International Technical Conference on The Enhanced Safety of Vehicles (ESV)*. 2023.

[42] Francis B Colavita. “Human sensory dominance”. In: *Perception & Psychophysics* 16.2 (1974), pp. 409–412.

[43] Cynthia Owsley and Gerald McGwin Jr. “Vision and driving”. In: *Vision research* 50.23 (2010), pp. 2348–2361. DOI: 10.1016/j.visres.2010.05.021.

[44] Charles A Nelson. “The development and neural bases of face recognition”. In: *Infant and Child Development: An International Journal of Research and Practice* 10.1-2 (2001), pp. 3–18. DOI: 10.1002/icd.239.

[45] Mark H Johnson. “Subcortical face processing”. In: *Nature Reviews Neuroscience* 6.10 (2005), pp. 766–774. DOI: 10.1038/nrn1766.

[46] Romina Palermo and Gillian Rhodes. “Are you always on my mind? A review of how face perception and attention interact”. In: *Neuropsychologia* 45.1 (2007), pp. 75–92. DOI: 10.1016/j.neuropsychologia.2006.04.025.

[47] Alumit Ishai. “Let’s face it: it’sa cortical network”. In: *Neuroimage* 40.2 (2008), pp. 415–419. DOI: 10.1016/j.neuroimage.2007.10.040.

[48] Nouchine Hadjikhani, Kestutis Kveraga, Paulami Naik, and Seppo P Ahlfors. “Early (N170) activation of face-specific cortex by face-like objects”. In: *Neuroreport* 20.4 (2009), p. 403. DOI: 10.1097/wnr.0b013e328325a8e1.

[49] Shichuan Du and Aleix M. Martinez. “Wait, are you sad or angry? Large exposure time differences required for the categorization of facial expressions of emotion”. In: *Journal of Vision* 13.4 (Mar. 2013), pp. 13–13. ISSN: 1534-7362. DOI: 10.1167/13.4.13.

[50] Lia N Kendall, Quentin Raffaelli, Alan Kingstone, and Rebecca M Todd. “Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic”. In: *Cognitive research: principles and implications* 1.1 (2016), pp. 1–14. DOI: 10.1186/s41235-016-0021-8.

[51] Janet Wessler and Jochim Hansen. “Facial mimicry is independent of stimulus format: Evidence for facial mimicry of stick figures and photographs”. In: *Acta Psychologica* 213 (2021), p. 103249. ISSN: 0001-6918. DOI: 10.1016/j.actpsy.2020.103249.

[52] Joseph S. Lappin, Duje Tadin, Jeffrey B. Nyquist, and Anne L. Corn. “Spatial and temporal limits of motion perception across variations in speed, eccentricity, and low vision”. In: *Journal of Vision* 9.1 (Jan. 2009), pp. 30–30. ISSN: 1534-7362. DOI: 10.1167/9.1.30.

[53] Suzanne P McKee and Ken Nakayama. “The detection of motion in the peripheral visual field”. In: *Vision research* 24.1 (1984), pp. 25–32. DOI: 10.1016/0042-6989(84)90140-8.

[54] Manfred Fahle and Christian Wehrhahn. “Motion perception in the peripheral visual field”. In: *Graefe’s archive for clinical and experimental ophthalmology* 229.5 (1991), pp. 430–436. DOI: 10.1007/bf00166305.

[55] John C Middlebrooks. “Narrow-band sound localization related to external ear acoustics”. In: *The Journal of the Acoustical Society of America* 92.5 (1992), pp. 2607–2624. DOI: 10.1121/1.404400.

[56] Simon Carlile. *Virtual Auditory Space: Generation and Applications*. Springer, 1996. DOI: 10.1007/978-3-662-22594-3.

[57] Craig Jin, Anna Corderoy, Simon Carlile, and André Van Schaik. “Spectral cues in human sound localization”. In: *Advances in Neural Information Processing Systems* 12 (1999).

DRIVER ENGAGEMENT

Adrià, Roig
Francesco, Deiana
James, Jackson
Clara, Cabutí
Cristina, Periago

IDIADA Automotive Technology SA
Spain

Paper Number 23-0315

ABSTRACT

One of the great challenges around the advent of driver assistance systems is to ensure that drivers understand the true capability of technology, such that they can behave accordingly for safe vehicle operation. This understanding can be influenced by a range of factors including vehicle instructions, user interface and warnings, and system control behavior. Validation accounting for these important aspects is therefore central to understanding and comparing safety performance for real world use for overall system design implementations.
This paper presents a test methodology specified for implementation on an automotive proving ground facility capturing pre-use information, and driver-vehicle interaction during assisted driving regarding user interface and system control behavior. Data collection was defined around the quantification of driver engagement with the driving task using subjective measures to assess progressive effects of system use and objective metrics considering driver behavior and capability to respond to an emergency scenario.
In a pilot assessment, a between-subjects test was conducted using two vehicles with differing assisted driving concepts. A sample of naïve drivers (n=39) was recruited and, following a customer focused description of system functionality, was instructed to drive on a test track in continuous highway driving scenario with longitudinal and lateral driver assistance features active. Subsequently, a critical 'cut-out' event was presented requiring a driver response to avoid an in-lane obstacle.
Results indicate variability in how drivers interact with the system during 'normal driving' with subjective measures demonstrating differences in metrics associated with engagement. Likewise, objective measures for driver reaction to the critical event signify differing levels of driver vigilance associated with perceived functionality of individual systems.
Outcomes from this experimental test mark a step in the development of test methods for global assistance system assessment and provide a platform for further progression and refinement of tests. This has implications system design verification with highly replicability whilst accounting for use by representative drivers, alongside possible applications in consumer and regulatory testing with representative drivers.

INTRODUCTION

Background

Systems automating parts, or all, of the driving task have the potential to provide significant benefits in safety and comfort. Such systems have been classified according to the degree automation they provide by SAE International [16], with different degrees of automation corresponding with 6 distinct levels, with these being defined by the responsibilities between driver and system for safe vehicle control [Table 1]. Levels 0, 4, and 5 classify a human driver or a system as being solely responsible for all aspects of the driving task, whilst levels 1, 2, and 3 each involve differing degrees of shared overall responsibility according to execution of the driving task, monitoring of the environment, and the fallback in case of failure. Crucially for all systems classified under these levels the system is able to perform at least part of the dynamic driving task, however drivers must maintain involvement, such that they are able to manage situations where the system is unable due its functional limitations.

In recent years automated driving systems have become more and more prevalent on consumer vehicles with systems classified as level 2 – the main focus of this paper – as being at the forefront. As use has increased there have been growing concerns around the safe use of the technology, particularly around misuse of systems. This is characterised by lowered levels of driver attention observed during use and in accident reports, which indicate drivers exhibiting behaviours inappropriate for the assistance functionality limitations of a level 2 system where they are responsible for vehicle control at all times. Driver condition, and associated behaviours during use can be closely related to the information received by users around system marketing, instructions, and interactions during use influences a drivers condition and vigilance over the system.

In response, one of the central themes of development as level 2 driving assistance systems have advanced has been to ensure that drivers are able to effectively cooperate with systems to maintain safe vehicle control. With the novelty of these types of consumer technology in vehicles a key challenge has been to ensure that a driver perceives and understands the true capability of the system.

Table 1.
SAE Levels of Driving Automation

	SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?	You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You are not driving when these automated driving features are engaged – even if you are seated in "the driver seat"		
	You must constantly supervise these support features, you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
	These are driver support features			**These are automated driving features**		
What do these features do?	These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to driver	These features can drive the vehicle under limited conditions and will not operate- unless all required conditions are met		This feature can drive the vehicle under all conditions

Driver Engagement

Due to its importance in safe operation a key challenge in implementation of assisted driving systems is to ensure that the driver is appropriately engaged in the driving task, corresponding to the level of assistance provided by the system. In practice this involves implementation of the vehicle and its systems in such a way that appropriate driver engagement is prompted, providing clear and timely communication about the vehicle's status and capabilities, and establishing clear expectations for the driver's role in the driving process, relative to the level of assistance or automation provided when active.

The concept of driver engagement has become a central theme given its recognized importance in the development of ADAS systems geared towards the improvement of driver performance, comfort, and roadway safety. Driver engagement has been studied extensively in the field of human factors and transportation psychology, and different definitions can be found in literature, within which certain recurring elements can be recognized: complexity and difficulty of the driving task, driver's level of motivation and involvement in the task, and the presence of distractions or other competing demands on the driver's attention.

Stanton [1] defines driver engagement as a *"measure of the degree to which drivers are actively participating in the driving task and is characterized by the driver's ability to recognize and react to potential hazards in a timely and appropriate manner"*. Peters et. Al [2] provides a more specific definition that emphasizes the importance of the cognitive factor and level of involvement of the driver: "*Driver engagement is defined as the extent to which the driver is mentally and physically engaged in the driving task. It is a combination of the driver's attention, focus, and involvement in the driving activity"*. One challenge in developing ADAS systems is to ensure that the driver is appropriately engaged in the driving process, regardless of the level of automation.

This may involve designing the vehicle and its systems in a way that promotes driver engagement, providing clear and timely communication about the vehicle's status and capabilities, and establishing clear expectations for the driver's role in the driving process. Current regulations on assisted driving consider performance and assessment requirements related to driver engagement, as it ensures that drivers maintain the correct level of control and vigilance in relation to the system's behaviour.

Reflecting its importance, regulation and consumer testing programmes for assisted driving have begun to include performance and assessment requirements related to driver condition, reflecting the importance of drivers maintaining the correct level of control and vigilance in relation to the system's functionality such as European Union Vehicle Regulation R159 [15] Additionally, driver engagement has been identified as an area for further exploration, particularly in the context of quantifying driver condition and associated risk in Euro NCAP Roadmap for 2030 [14].

Evaluation Methods

There are several methods that can be used to evaluate driver engagement in advanced driver assistance systems (ADAS) and autonomous vehicles. These methods include self-report measures, which rely on drivers to report their level of engagement while driving; behavioural measures, which involve observing and measuring driver behaviour during the driving task; performance measures, which involve measuring the driver's performance on tasks related to driving; and physiological measures, which involve measuring the driver's physiological response to the driving task. In the literature, it is possible to find various evaluation experiments that have been conducted using driving simulators to study the development and design of human-machine interaction. However, it is more difficult to find studies that have been conducted in real-world environments or on test tracks.

It is important to recognize that no single method is likely to be sufficient for evaluating driver engagement in advanced driver assistance systems (ADAS) and autonomous vehicles. Instead, a combination of methods may be needed to fully understand the driver's level of engagement. To address this issue, there is a need for a comprehensive testing procedure that can validate the assessment of driver engagement in a controlled environment. This procedure should consider both subjective data obtained directly from the driver and objective data obtained from the vehicle, in order to provide a complete picture of the driver's level of engagement. This approach can help to ensure that ADAS and autonomous vehicle systems are designed and developed in a way that effectively supports driver engagement and safety.

Subjective metrics

Subjective metrics are measures or evaluation criteria that rely on the opinions or personal perceptions of individuals. In the context of evaluating driver engagement some of the most relevant subjective metrics are mental workload (MWL) and trust.

Mental Workload is a widely studied concept in human factors and ergonomics and has been defined and measured in various ways. There is not a universally accepted definition of mental workload, although some are more widely accepted than others. Hoedemaker (2002) [4] defines Mental workload as the *"amount of mental effort or cognitive resources that an individual must expend in order to perform a task or set of tasks"* Pickup (2005) [5] proposed a multi-dimensional conceptualization of mental workload that is based on the core psychometric properties of load, demand, effort, and effects. This has been widely cited and used in research on mental workload and has been adopted as the basis for the IWS scale, a recognized tool for measuring mental workload. For what concerns trust, the definition proposed by Lee and See (2004) [6] is widely accepted in the literature on human factors and has been extensively cited and validated. According to this definition, "*trust is an attitude that will help an individual achieve their goals in a situation characterized by uncertainty and vulnerability, which has been shown to play a role in influencing operators' strategies toward the use of automation*" This definition provides a valuable framework for understanding and assessing trust in automated systems. This concept has been measured using scales such as mental workload, trust has been evaluated by using scales in automated systems scale. Trust is known to play an important role in how drivers use ADAS and their level of disengagement from driving tasks. [7]

Objective metrics

Objective metrics, on the other hand, are measures or evaluation criteria that are based on observable and quantifiable data, rather than subjective opinions or perceptions. In the context of evaluating driver engagement, objective metrics might include things like the driver's speed, acceleration, braking, or steering behaviour, as well as metrics related to vehicle performance, such as fuel efficiency or emissions.

In this study, we are using Time To Collision (TTC) as an objective metric to evaluate the performance of advanced driver assistance systems (ADAS) and autonomous vehicles. Ozbay [8] defines Time to Collision (TTC) as the *"time it would take a following vehicle to collide with a leading one if the vehicles do not change their current movement characteristics. It can also be explained as the time needed to avoid a collision by applying certain countermeasures"*. By tracking TTC, it is possible to determine how well a vehicle can avoid collisions and maintain a safe distance from other objects in its environment.

For specific TTC calculation, former studies generally used the relative distance D (m) between the two vehicles divided by their relative speed ΔV (m/s) and formulated TTC as follows [Equation (1)]:

$$TTC = \frac{D}{\Delta V}$$

Equation (1)

This paper describes a test methodology that has been developed for evaluating the performance of advanced driver assistance systems (ADAS) on an automotive proving ground facility. The methodology involves collecting data on the driver's pre-use characteristics, as well as their interactions with the vehicle's user interface and control systems during assisted driving. The data collection is focused on quantifying the driver's engagement with the driving task, using both subjective measures to assess the effects of system use over time and objective metrics to assess the driver's behaviour and ability to respond to emergency situations. The aim of this test methodology is to provide a comprehensive and reliable means of evaluating the performance of ADAS and autonomous vehicle systems in terms of their impact on driver engagement and safety.

OBJECTIVES

The main objective of the study is to test a methodology for implementation on an automotive proving ground facility capturing pre-use information, and driver-vehicle interaction during assisted driving regarding user interface and system control behavior.
Methodology aims to identify collect subjective and objective data to contribute to the development of systems that are as well-suited as possible to human characteristics and variability.

METHOD AND PROCEDURE

Summary

The methodology developed in this study aims to compare level 2 (L2) advanced driver assistance systems (ADAS) with different characteristics, to identify differences in terms of driver engagement. To accomplish this, two vehicles with different characteristics were selected based on their EuroNCAP [3] safety assist ratings, with the aim of comparing them. The selected vehicles were a Volkswagen Golf 8 (2020) as a medium L2 vehicle and a Tesla Model 3 (2020) as an advanced L2 vehicle. In order to enable comparison between the two vehicles, they were both instrumented in the same way, with the same sensors and measurement devices. Specifically, the vehicles were instrumented with:

a. **Vector Kit:** Vector provides all de data from test vehicle bus CAN and the other CAN signals;
b. **RT & RT Range:** RT is the device used to acquire precise geolocation, velocity, acceleration and lateral acceleration to determine the steering moment;
c. **Camera set-up:** Test included *three* video cameras (forward-facing, rear facing-dashboard/environmental). Video data were recorded at 30hz.

all measurement and data collection components were connected to a CPU that was equipped with CANape software, used for data collection and synchronization.

Table 2.
Vehicles models and features description

Vehicle	Engine type	AD System	EURONCAP SAFETY ASSIST (%)	L2 level definition	Features
Golf 8	Gasoline	Travel Assist (IQ Drive)	82%	Medium L2	ACC (Autonomous Cruise Control) AEB (Autonomous Emergency Braking) LCA (Lane Centring Assist)
Tesla Model 3	Electric	Autopilot	98%	Advanced L2	

Test Specifications

Participants were asked to drive for 40 minutes on a simple, motorway-like track. They were accompanied during the test by a professional co-driver in order to ensure safety throughout the test. The same co-driver was also in charge of administering the mental workload and confidence level scales, explained in [During the test].
participants were required to follow a lead vehicle, a Seat Leon (2018), for the entire duration of the test. The test consisted of three stages:

- The first 10 minutes were conducted without any driving assistance systems, in order to allow the participants to become familiar with the vehicle and the track.

- The next 30 minutes of driving were conducted with driving assistance systems enabled, including the adaptive cruise control (ACC) set to a speed of 60 km/h.

- Each participant completed a total of 12 laps around the track.

Scenario

This study was conducted on the IDIADA Highway loop B, a 2.7 km long highway scenario with consistent lane markings and a standard lane width of 3.75 meters. In order to minimize risk, the track was used in semi-exclusive mode, which means that the drivers were unfamiliar with the track and had not previously driven on it. This was also done in order to more closely simulate real-world driving conditions and to ensure that the drivers were fully engaged and attentive while driving.

On the penultimate lap, an obstacle [Figure II. ADAC car dummy 2D] was placed in the middle of the lane without warning the participants. As depicted in [Figure I. Highway loop B and manoeuvre description] The obstacle was placed immediately after a bend in the road, as shown in Figure I, so that it was not visible from a distance and was hidden by the lead vehicle. In order to simulate a real-world emergency situation, the lead vehicle had to perform a cut-out maneuver 15 meters before the obstacle. The participants were then required to react to the unexpected event with the driving assistance systems enabled, depending on their level of attention and engagement with the system at that moment.

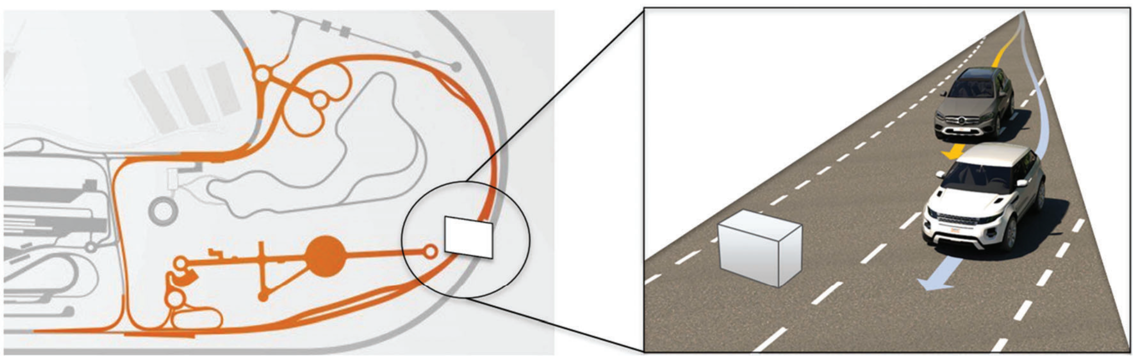


Figure I. Highway loop B and manoeuvre description

Figure II. ADAC car dummy 2D

Participants

In this study, 39 naïve participants were recruited from a specialized agency based on predefined criteria that were determined based on literature review and project specifications. [Table 3.]:

- Non-professional drivers.
- ¾ of participants without any experience with partial driving automation
- ¼ of participants with experience with partial driving automation
- Equal distribution of males and females (20/20)
- Have a valid driver's license

- Age between 21 and 58, reflecting the average age of drivers in Spain.

Table 3.
Participants demographic data sample

Age range	Gender	Mean	Range
20-29	4M/4F	25	21-28
30-39	6M/4F	34	30-39
40-49	2M/6F	43	40-48
50-70	8M/6F	58	50-68

Before the test, all participants were contacted via email and provided with information about the procedure, criteria, and general objectives of the study. They were also asked to complete the DSQ questionnaire, French et al., 1993 [9] in order to identify their driving style. The questionnaire included items related to various aspects of driving style, such as speed, calmness, social resistance, focus, planning, and deviance. The aim of this questionnaire was to provide a more complete understanding of the participants' driving habits and behaviours, and to allow the researchers to better interpret the data collected during the test.

Table 4.
Driving style participants distribution

Driving Style	#Participants
SPEED	4
CALMNESS	3
SOCIAL RESISTANCE	0
FOCUS	30
PLANNING	0
DEVIANCE	0

Pre – test procedure

Upon arrival, the participants were given a short briefing by a Human Factors expert, which aimed to explain to participants the procedure of the study, vehicle characteristics and functionality (including the instrumentations and assisted functions), safety and data protection measures and to let them fill the consent form.
Afterwards, they were given a short *ad-hoc* questionnaire created with respect to their emotional state at that moment. the questionnaire consists of 16 items with a 4-point Likert-type scale (*1 not at all, 2 a little,3 moderately, 4 a lot*). Upon arriving at the testing facility, participants were asked to sit in the driver's seat of the vehicle and adjust their driving position to their liking. They were also given the opportunity to ask any questions they had about the vehicle or ADAS systems. Once the participants indicated that they were ready to start, the test began. This process was designed to ensure that participants were comfortable and familiar with the vehicle and its systems before starting the test, and to allow them to ask any questions or raise any concerns they might have.

During the test

During the test, the two scales of IWS (implemented workload scale) and TASS (trust in automated system survey) were administered by the co-driver to the participant at a **regular interval of 5 minutes**.

- **Integrated Workload Scale (IWS) (readapted)**, developed and tested for signallers by Pickup et. Al, 2015 [5], is a valuable measure of individually experienced peaks and troughs in workload over an interval or within a particular set of scenarios. The scale consists of 9 progressive points, asking the naïve driver to define his level of mental workload in that given situation.

- **TASS (Trust in Automated System Survey)**, developed by Jian et. Al [10] provides a model for assessing trust between humans and machines based on empirical data and help understand how the system characteristics might affect drivers. In reference to earlier studies that assessed automation trust with single-item trust ratings [11] and perceived risk in ACC [12] the experimenter asked the participants to report their automation trust on a scale from 0% to 100%.

As mentioned above, some objective data were also collected during the test such as video recording, TTC response, vehicle trajectory (GPS). During the test, the co-pilot also had the responsibility of monitoring participant driving behaviour and intervene in case of dangerous situations.

For the objective data, we have analysed the TTC (Time To Collision) in the moment that participant starts to turn the steering wheel. To calculate TTC [Equation (1)] the following parameters were considerated:

- Distance (D): The exact moment when the car starts to swerve to avoid the target
- Relative velocity (ΔV): The speed of the target, which will always be 0, and the speed of the vehicle at the moment when the swerve-avoidance begins

To determine the threshold limit value, reference was made to the following criteria, defined by [13]: *"Various TTC thresholds can be defined to adapt to different road users and contexts different road users and contexts. Early research suggested critical TTC thresholds of 1 to 1.5 seconds and considered values up to 5 seconds to enable collision avoidance systems on highways"*. Considering these researchers set a collision avoidance threshold of 1.5 seconds.

Post-test

After completing the driving test, the participants were subjected to a short, semi-structured interview about their experience. The interview was designed to collect additional subjective data, impressions, and suggestions. It was also useful for gathering opinions on perception of system reliability following the critical event (the obstacle avoidance test), as well as feedback on how to improve the methodology. The data from the interviews were not included in this analysis, but they may be used in future studies and to corroborate data from other sources

RESULTS

The first round of testing for the validation of the methodology was successfully completed with a total of 39 naive participants, as previously mentioned.
Collected data has been cross-checked between the different car models, the Golf 8 and the Tesla Model 3, in order to underline differences between the level of engagement with L2 medium and advanced systems.
For the subjective data the IWS level and Trust Level with positive results in [Figure III. Workload level (IWS)

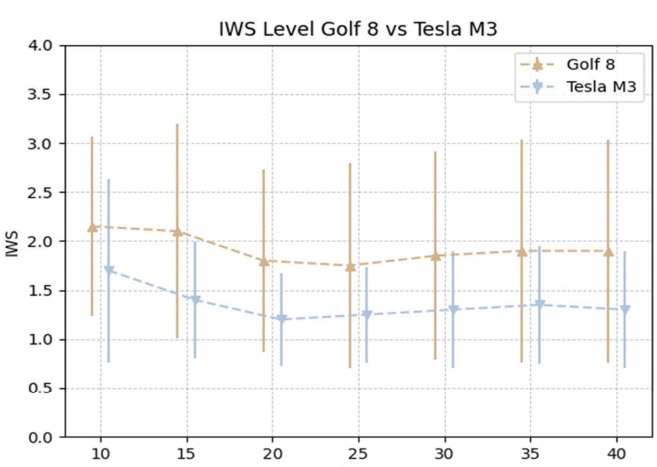


Figure III. Workload level (IWS) Golf 8 vs Tesla Model 3

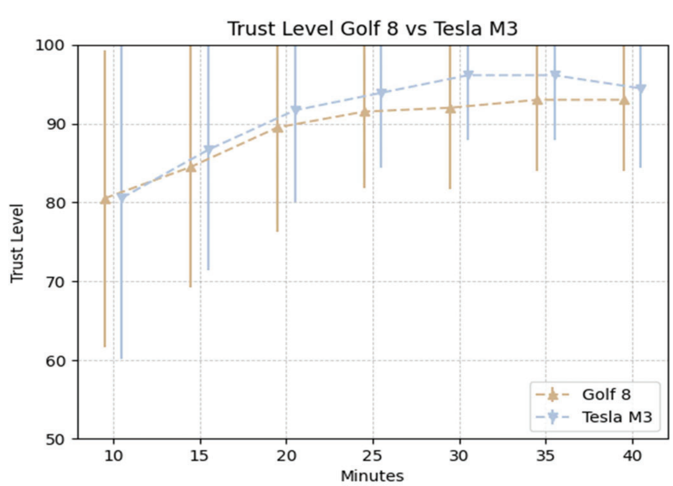


Figure III. Trust Level Golf 8 vs Tesla Model 3

Golf 8 vs Tesla Model 3] and [Figure IV. Trust Level Golf 8 vs Tesla Model 3].
Points in [Figure III. Workload level (IWS) Golf 8 vs Tesla Model 3], represent the average level of mental workload value for participants, differentiated between the two cars, while the verticals lines in each point represents the standard deviation. Same in [Figure IV. Trust Level Golf 8 vs Tesla Model 3] but as far as the trust level. As a static obstacle, the target has speed 0 km/h, whereas the speed of the test car was around 60 km/h, depending on user case. With the data acquired from the different participants, by applying the [Equation (1)] we have identified the differences in TTC between the two vehicle [**Error! Reference source not found.**]:

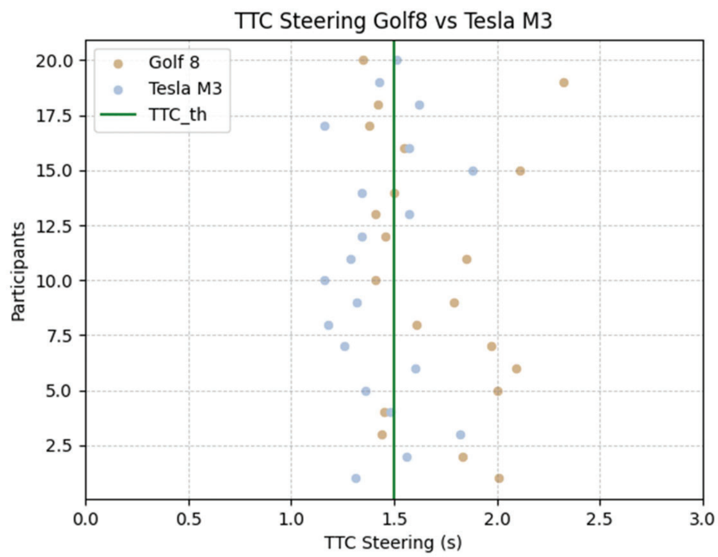


Figure IV. TTC Steering Golf 8 vs Tesla Model 3

Participants below and above TTC threshold are distributed in Table 5.:

Table 5.
Percentage TTC threshold distribution

	TTC < TTC threshold	TTC > TTC threshold
Golf 8	45%	55%
Tesla Model 3	60%	40%

CONCLUSIONS

The analysis of subjective and objective metrics related to the mental workload perceived by drivers and the level of trust revealed interesting differences between the two types of L2 systems. For what concerns subjective data, as represented in [Figure III. Workload level (IWS) Golf 8 vs Tesla Model 3] and [Figure IV. Trust Level Golf 8 vs Tesla Model 3], perceived mental workload and trust levels are low in both systems. None of the participants reported a mental workload level above 5, and the minimum average trust level for both systems was 80%. However, is possible to underline interesting differences between the two systems. The vehicle equipped with a medium L2 system has higher mental workload value (max. mean value 2,10 – min. mean value 1,75) in the whole test than the vehicle equipped with an L2 advanced system (max. mean value 1,75 – min. mean value 1,2).

According to the results of the study, the participants perceived a higher level of trust in the vehicle equipped with the advanced L2 System. More specifically, L2 advance vehicle has a max. mean value of 96,1% while the other vehicle has a max. mean value of 93%. Minimum mean value is almost the same in both vehicles (80,6% for L2 advanced and 80,5 for L2 medium).

The time-to-collision (TTC) values for the two vehicles showed significant differences. In the advanced L2 vehicle, 60% of participants had a TTC value below the threshold of 1.5 seconds. In contrast, only 45% of participants in the L2 medium vehicle had a TTC below the threshold [Table 5.]. These results suggest that the type of L2 system used (advanced or medium) may influence a driver's reaction time and ability to take control of the vehicle to redirect the maneuver.

According to the subjective and objective data, it appears that the perceived level of trust in the advanced L2 vehicle is inversely related to the mental workload experienced by the participants. As trust increased, mental workload decreased. This is reflected in the lower time-to-collision (TTC) values observed in participants with high trust levels; 60% of participants with high trust had TTC values below 1.5 seconds when using the advanced L2 vehicle

Next steps

This study represents the initial phase of a larger project that aims to develop a methodology for assessing the level of driver engagement in different advanced driver assistance systems (ADAS). The methodology uses both subjective and objective data in order to provide a more complete understanding of driver behavior and performance. In subsequent phases of the project, the research protocol will be refined and improved to create a more solid and comprehensive database. Although the methodology is still being developed, it has already been useful in identifying differences between the two systems tested in this study, such as the effectiveness of the test scenario and the performance of the vehicles on the proving ground. In future implementations, it is intended to also analyze the Time to Collision (TTC) based on the braking time, which was not possible in this study due to the automatic braking system of the vehicles. To enable this analysis, an additional camera will be placed above the brake pedal. Additionally, the data collected through the methodology will be further analyzed and compared in different clusters, such as by participants' age, gender, and driving style, in order to identify any patterns or trends that may be relevant to the development of ADAS. Also, it is intended to extend the study to other vehicles and automation systems such as L3.

BIBLIOGRAPHY

[1] Stanton, N. A., Salmon, P. M., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). A review of driver engagement and distraction. Ergonomics, 56(9), 1453-1469).

[2] Peters, J., & Mattes, K. (2015). How driver engagement and inattention lead to accidents. Journal of Safety Research, 56, 79-87).

[3] Latest Safety Ratings | Euro NCAP. (s. f.). https://www.euroncap.com/en/ratings-rewards/latest-safety-ratings/

[4] Brookhuis, K. A., van Driel, C. J., Hof, T., van Arem, B., & Hoedemaeker, M. (2009). Driving with a congestion assistant; mental workload and acceptance. Applied Ergonomics, 40(6), 1019–1025.

[5] Pickup, L., Wilson, J. R., Norris, B. J., Mitchell, L., & Morrisroe, G. (2005). The Integrated Workload Scale (IWS): A new self-report tool to assess railway signaller workload. Applied Ergonomics, 36(6), 681–693.

[6] Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. Human Factors: The Journal of the Human Factors and Ergonomics Society, 46(1), 50–80.

[7] Hoc, J., Cacciabue, P. C., Hollnagel, E. & Cacciabue, C. P. (1994). Expertise and Technology: Cognition & Human-Computer Cooperation (1.). Psychology Press.

[8] Ozbay, K., Yang, H., Bartin, B., & Mudigonda, S. (2008). Derivation and Validation of New Simulation-Based Surrogate Safety Measure. Transportation Research Record: Journal of the Transportation Research Board, 2083(1), 105–113.

[9] French, D. J., West, r. J., Elander, j., & wilding, j. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. Ergonomics, 36(6), 627–644.

[10] Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. International Journal of Cognitive Ergonomics, 4(1), 53–71.

[11] Madhani, K., Khasawneh, M. T., Kaewkuekool, S., Gramopadhye, A. K., & Melloy, B. J. (2002b). Measurement of Human Trust in a Hybrid Inspection for Varying Error Patterns. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46(3), 418–422.

[12] Saffarian, M., de Winter, J. C. F., & Happee, R. (2012). Automated Driving: Human-Factors Issues and Design Solutions. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56(1), 2296–2300.

[13] Saffarzadeh, M., Nadimi, N., Naseralavi, S., & Mamdoohi, A. R. (2013). A general formulation for time-to-collision safety indicator. Proceedings of the Institution of Civil Engineers - Transport, 166(5), 294–304.

[14] Vision 2030 | Euro NCAP. (s. f.) https://www.euroncap.com/en/press-media/press-releases/euro-ncap-vision-2030-a-safer-future-for-mobility/

[15] UN Regulation No 159 – uniform provisions concerning the approval of motor vehicles with regard to the Moving Off Information System for the Detection of Pedestrians and Cyclists [2021/829]

[16] SAE, J3016 (2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems.