A Novel Approach to Integrate Human-in-the-Loop Testing in the Development Chain of Automated Driving: The Example of Automated Lane Change [⋆]

Branko Rogic [*] **Demin Nalic** [**] **Arno Eichberger** [***]
Stefan Bernsteiner [****]

[*] *MAGNA Steyr Fahrzeugtechnik AG & CO KG, Department Electric/Elektronics and Driver Assistance, Austria, (e-mail: branko.rogic@magna.com).*
[**] *Automotive Engineering Department, Technical University Graz, Austria, (e-mail: demin.nalic@tugraz.at).*
[***] *Automotive Engineering Department, Technical University Graz, Austria, (e-mail: arno.eichberger@tugraz.at).*
[****] *MAGNA Steyr Fahrzeugtechnik AG & CO KG, Department Electric/Elektronics and Driver Assistance, Austria, (e-mail: stefan.bernsteiner@magna.com).*

Abstract: For market introduction of advanced driver assistant (ADAS) and automated driving (AD) systems on full vehicle level, testing and validation is one of the biggest challenges. The present study describes a novel approach that integrates a driving simulator in a virtual development process aiming to reduce time and effort for system development. The approach is demonstrated on a specific automated lane change assist (LCA) system. To this end, the LCA function and the corresponding human machine interface (HMI) are developed and implemented in the driving simulator. The core of the approach is a driving simulator-based testing method which proposes a novel two stage testing concept and involves multiple test drivers. The method provides better insight into the overall system performance and, moreover, detects potentials for improvements dedicated for the ADAS functionalities as well as for the design of the HMI system. Using this method, a driving simulator study with 20 volunteer drivers is conducted to evaluate the LCA system with respect to driver acceptance and user friendliness. The results of the study will be used for the parametrization and fine tuning of the LCA function as well as for the HMI improvement.

Keywords: Human Machine Systems, Transportation and Vehicle Systems, Advanced Driver Assistance Systems, Lane Change Assist

1. INTRODUCTION

Currently one of the great obstacles for a faster market introduction of automated driving functions of SAE level 3 and higher SAE (2019), is the unsolved problem of safety proof and/or guarantee of fail-safe operation Junietz et al. (2018b). Furthermore, drivers acceptance and trust in automated driving function is crucial for their use and market success. Drivers acceptance depends mostly on the obvious benefits for the user (such as usability and comfort) and the fail-safe characteristics of a system Brookhuis et al. (2019),Wang et al. (2001). The acceptance of an automated system will increase as soon as automated driving actions are performed in a way that is comprehensible for the driver. Therefore, the system should mimic human driver behaviour van Driel et al. (2007). Prior to this study an on-road naturalistic driving study

(NDS) with 20 drivers were conducted in order to gain data about human driver behaviour in the specific traffic situations of an overtaking manoeuvre. The recorded data were analysed and used for the initial parametrisation of the LCA function. Even though the naturalistic driving tests could be used for the initial parametrization, they were not sound enough in order to extract the final, statistically relevant parameters. The reason for this, beside the low number of drivers and performed manoeuvres, is the practical impossibility of providing the same traffic conditions for each repetition of the same manoeuvre on a public highway. However, the conducted tests provided some directives for the initial parametrisation and revealed two comfort-relevant parameters: lane change initialisation time and lane change duration. Current researchers studies of human drivers lane change behaviour (LCB) using NDS in Hetrick (1997), Henning (2010), Sporer et al. (1998),Salvucci et al. (2001), Salvucci (2006) and Olsen et al. (2002). In those studies the focus is mainly on the drivers LCB regarding the LC time, trajectory and

duration. According to those studies, the duration of LC manoeuvre is between 3.5 and 7.4 seconds. However, there are neither studies investigating LC manoeuvre in specific traffic situations under consideration of drivers comfort, nor studies investigating LC initialisation timing and LC duration in this context. The Acceleration and jerk are considered for the driving comfort Bellem et al. (2016) and used for the lane change algorithm developed for this work Samiee et al. (2016). Due to the fact LC initialisation timing and LC duration could not be derived from the naturalistic driving tests, the first objective will be to introduce a method to evaluate this parameters in the driving simulator using four standard traffic scenarios on a highway (Figure 3 on page 4). Along with the LCA function, a corresponding human machine interface is developed. Since drivers will still need to monitor and supervise the actions of the next generation automated driving functions SAE (2019), drivers acceptance of these systems will depend at great extent on the usability and user friendliness of the HMI. Consequently, if the acceptance was evaluated only in naturalistic driving tests, at later development stages, the rating of the test persons might be lower for a system with a basically good parametrized automated function but with an inconvenient HMI and vice versa. For this reason the second objective will be to evaluate the usability and user friendliness of the HMI. In order to cover both stated objectives, the developed method proposes a novel **two stage testing concept** with a driving simulator, whereby the first stage targets a function testing and the second stage an HMI testing. This should provide better insight in the overall system performance and detect potentials for improvements dedicated for the function as well as for the HMI system. The goal of the work is to investigate the potentials for applying virtual tools in the development process of automated driving systems. Some methods and challenges of virtual testing can be found in Broggi et al. (2013),Nidhi and Paddock (2016) and Stellet et al. (2015).

2. METHOD

The driving simulator (DS) study is part of a the presented method for virtual integration and validation of advanced driver assistant systems (ADAS), see Figure 1 on page 2. A lane change assistance system is used as the system under development and test. The system is based on the LCA function presented in Samiee et al. (2016). This function is used for the trajectory planning and decision making regarding a lane change (LC) manoeuvre, based on a current traffic situation. The development process follows the steps shown on the left branch in Figure 1 on page 2. The proper functionality according to the functional concept is assured with defined test scenarios and test scenarios generated using the framework in Nalic et al. (2019). Some of this scenarios are described in Nilsson et al. (2017) and used in the development process for function modelling and implementation.

As mentioned before the initialisation time is the main metric we use in this work for the one part of the evaluation process. The initialisation time is defined here as time-to-collision TTC_{init} with the vehicle in front (target 1), in the moment when the LC manoeuvre is initiated, which is derived as follows
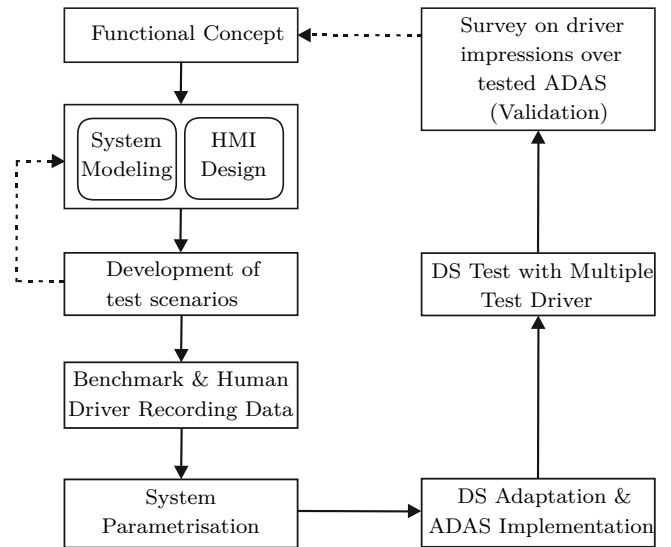

Fig. 1. Flowchart of the ADAS virtual development method.

$$TTC_{init} = \frac{x_{T_1,init} - x_{ego,init}}{v_{T_1,init} - v_{ego,init}} \qquad (1)$$

where $x_{T_1,init}$ describes the initial longitudinal position of the target 1, see Figure 3 on page 4, $x_{ego,init}$ is the initial longitudinal position of the ego vehicle, $v_{ego,init}$ describes the initial velocity of the ego vehicle and $v_{T_1,init}$ is the initial velocity of the target 1. Research works in the field of human driver lane change behaviour are shown in following studies Sporer et al. (1998),Salvucci et al. (2001), Salvucci (2006) and Olsen et al. (2002). The focus of these is mainly on the LC trajectory. According to those, the duration of a comfortable LC manoeuvre is between 3.5 and 7.4 seconds. However, there are neither studies investigating LC manoeuvre in specific traffic situations under consideration of drivers comfort, nor studies investigating LC initialisation timing and LC duration in this context. Therefore, and due to the fact that these two parameters could not be derived from the naturalistic driving tests, as explained above, the first objective of the presented method is to evaluate this parameters in the driving simulator using four standard traffic scenarios on a highway (Figure 3 on page 4). The results will be used then for final parametrisation of the LCA function. Based on the discussion above, following objectives for the driving simulator study are divided as following two testing stages:

- **The first test stage** investigates drivers preferred LC initiation timing and LC duration using four basic LC traffic scenarios shown in Figure 3. The advantage of scenario repeatability with a driving simulator will be used to overcome already discussed disadvantages of the conducted naturalistic driving tests. The findings from this test stage will be later used to define rules and parameters for the control strategy of the LCA function.
- **The second test stage** the LCA system is tested on a highway scenario with the microscopic traffic flow simulation Martin and Peter (2011). In this stage, the overall acceptance of an automated LCA system is evaluated in a "naturalistic" traffic scenario on a highway. The traffic flow test concept provides infor-

mation about the overall drivers impression regarding usability of the LCA system in three automation levels: manual, semi-automated and automated (see Section 2.2.2). Hereby the focus is on the HMI concept, with its logic, design and related drivers subjective workload, comfort and safety.

It should be stated here that the goal of this study was not to determine any definite characteristic of the system (e.g. system parameters, control rules etc.) but to validate the development method and demonstrate its potentials and the potentials for the application of a driving simulator for development and integration purposes. Therefore, the results examined using only 20 persons are not statistically significant.

2.1 Test setup

Participants A total of 20 drivers, 10 males and 10 females, participated in the study. Age range of all participants was from 20 to 61 years, with an exception of one female age 70 and one male age 80. Mean age was 42.1 and standard deviation was 16.6. They were chosen from a database of 284 people, which was built up in previous projects. Thus, most of them were already familiar with the driving simulator as they had already taken part in previous studies Koglbauer et al. (2017). For the selection of participants not only gender and age distribution criteria were set but also minimum annual mileage of 10,000 km, predominantly on highways, in order to target the end-user group of skilled drivers.

Driving Simulator The study was conducted using the fixed base full vehicle driving simulator of the Graz University of Technology. It uses an autostereoscopic visual system including nine LCD monitors, four 55″ monitors placed in front of the front wind shield, two 23″ monitors in each side window and one monitor covering the rear window. This setup of monitors covers the relevant field of view of the driver (over 180°) Eggeling et al. (2013). Haptic responses are provided by an active steering wheel, active brake pedal and passive throttle along with an automatic transmission gear stick. Integrated bass shakers in the drivers seat and vehicles compartment simulate the vibrations of the engine. Also engine, wind, traffic and wheel-rolling noise are simulated using a sound system. The simulation of the microscopic traffic flow is carried out by PTV Vissim software Leyn and Vortisch (2015) in co-simulation with the ego-vehicle dynamics. This is essential since the dynamic interaction between the ego-vehicle and the traffic is a key element for test and validation of the LCA system in the second testing stage of the presented method. The ego vehicle dynamics is simulated by the commercial vehicle dynamics software AVL-VSM, the output of this simulation (i.e. dynamic vehicle states) is then returned to Vissim, representing one of the simulated vehicles in the traffic flow simulation. In turn, the dynamic vehicle states of the individual vehicles are transferred from Vissim to the simulators co-simulation platform.

2.2 Two stage testing procedure

The essence of the proposed method is a two-stage testing procedure: the first stage for function performance testing,

Fig. 2. Driving simulator showing a lane change manoeuvre

	v_{T_1}	v_{T_2}	v_{T_3}	$d_{T_{2,3}}$
Scenario 1	100	-	-	-
Scenario 2	100	160	-	-
Scenario 3	100	160	-	-
Scenario 4	100	160	160	80

Table 1. Test scenarios setup parameters

named "fixed scenarios", and the second stage for HMI testing, named "traffic flow scenario".

First Testing Stage: Fixed Scenarios In order to guarantee a proper functionality, the system was previously tested in a systematic approach based on the variation of the test scenario parameters. Starting point for the fixed scenarios are four basic scenarios illustrated in Figure 3 on page 4. They represent four possible traffic situations on highway, in which overtaking manoeuvres can be executed. For each basic scenario, one fixed parameter combination was selected, out of a variety of combinations. These four scenarios are used for the driving simulator study as a fixed manoeuvers. The combination of parameters is set in a way that the resulting traffic situation allows a comfortable overtaking manoeuvre for most drivers, without having to change the constant velocity of 130 km/h for ego vehicle. Table 1 shows the setup parameters for each scenario with the corresponding target velocities v_{T_i} in $km\,h^{-1}$ with index $i \in \{1, 2, 3\}$ as target number and with the distance between Target 2 and 3 $d_{T_{2,3}}$ in m.

The main test scenario parameters are velocity and starting position of each vehicle. Further, a TTC in reference to Target 1 (Fig. 1) is the deciding factor for a driver in order to initiate an overtaking lane change manoeuvre, Zhou and Itoh (2016). Thus, it is a common metric criterion for different evaluation methods, see Junietz et al. (2018a), Satzoda and Trivedi (2016) and Rodemerk et al. (2012). By its definition, the same TTC value between two objects can be established by combining different velocities and distances 1. Since the goal here is to detect a preferred TTC (in reference to Target 1) for a comfortable lane change manoeuvre, one single combination of test parameters (velocities and starting positions) is chosen for each scenario. In this way the TTC depends only on the lane change initiation point during the scenario. At the beginning, the participants perform one initial drive in order to get used to the driving simulator and to get familiar with the scenario. After that every driver conducts each test scenario three times – once manually and twice automated with different function parametrisations. In the
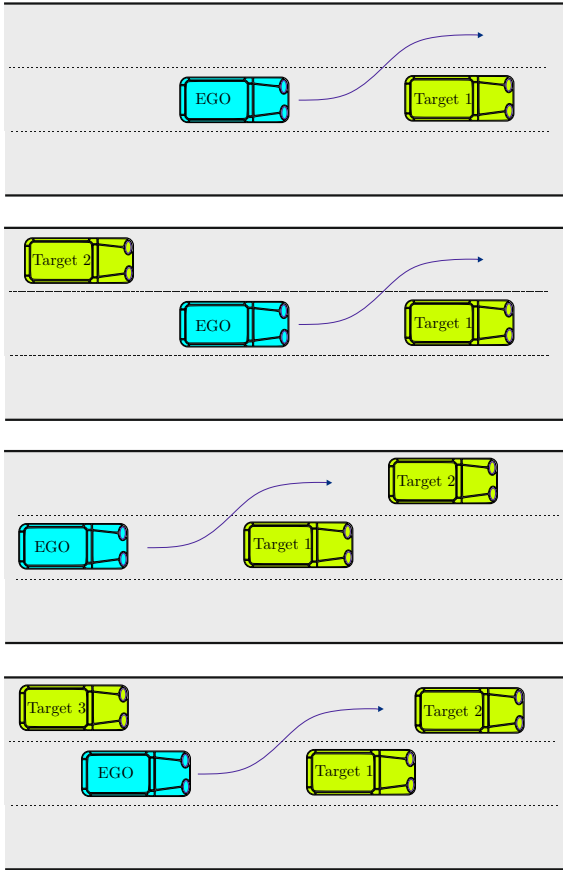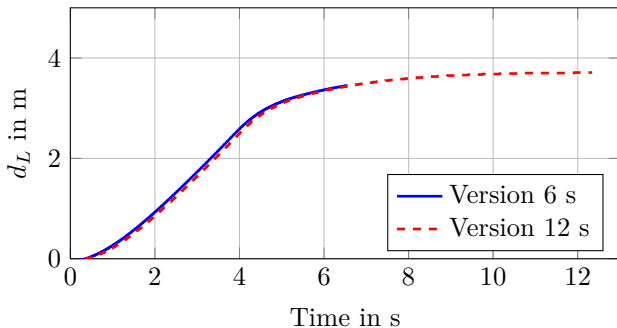
Fig. 3. LCA relevant traffic scenarios

Fig. 4. Lateral displacement d_L of the two automated Lane Change variants

first run driver steers the ego vehicle manually. The driven trajectory is recorded, and the properties of interest (LC initiation timing and duration) are analyzed afterwards. In the second and the third run, driver was instructed to activate the automated LCA and to concentrate on his experience of being driven in an automated way through the same scenario. In the second run the LC function is set to 6 s for the LC initiation time, and 6 s for the LC duration time as well. In the third run the initiation time is set to 12 s and the duration time again to 6 s. Figure 4 shows the trajectory curves over time for both automated variants. The trajectory curves start with LC initialisation and end when the rear most point of target 1 is reached.

The first version, with the initiation time of 6 seconds, rests on the conducted naturalistic driving test and on

Table 2. Evaluation Questionnaire

Question	Scale Range	Scale Description
Evaluation questions with defined scales and criteria		
1. Overall impression over the lane change		
	[1,..,7]	1 - bad, ..,7 - perfect
2. Initialisation Time		
	[1,..,7]	1 - bad, ..,7 - perfect
3. Lane Change Duration		
	[1,..,7]	1 - bad, ..,7 - perfect
4. Overall impression over the lane change		
		1 - Much longer
		2 - A bit longer
		3 - Duration is perfect
		4 - A bit shorter
		5 - Much shorter
5. Initialisation Time		
		1 - Much longer
		2 - A bit longer
		3 - Duration is perfect
		4 - A bit shorter
		5 - Much shorter

the research in Salvucci et al. (2001). The second version, with the initiation time of 12 s, is mainly based on the naturalistic driving tests (see section 1. Introduction). this two-variants testing method offers a better basis for the participants subjective assessment, since they could directly compare different system performances. In order to benefit from this comparation effect, the 6s version was always driven as the first: the assumption is that, if the obviously less critical 12s version had been driven as the first, it would have resulted with rather poor rating for the 6s version afterwards. Therefore, participants first experience and rate the 6s version, uninfluenced by the other version, and only having in mind their own performance from the previous run. After their last run with 12s version, they will rate this one undoubtedly by comparing it with the 6s version. Thus, they would tend to be more critical against the much earlier start of the lane change. Such testing order should produce more balanced results and help participants to make a confident rate. The assessment is carried out using a short questionnaire with five subjective evaluations which are the same for each manoeuvre. Participants fill in the questionnaire immediately after each automated driving run. Unless stated otherwise, the evaluations are assessed on a scale of 1 – bad or inappropriate, to 7 – perfect. The questionnaire evaluations are explained in Tab. 2.

Second Testing Stage: Traffic Flow Scenario Unlike driving in predefined fixed scenarios, in this stage participants are driving in a straight three-way highway road with stochastic traffic generated by microscopic traffic simulation (see section Driving simulator). The version with 12 s was enhanced with a state-of-the-art Adaptive Cruise Control (ACC) and Lane Keeping (LKA) systems in order to provide a fully automated driving experience. In addition to the automated driving mode – a manual and a semi-automated LC modes, based on the same aforementioned decision making algorithm (Sammie et al. 2016), were also implemented. This automation level benchmark is intended to be a certain add-on for the study, demonstrating the methods potential. However, according to the participants reports, similar to the two variants testing model from the first stage, the participants were able to answer the questionnaire more precisely by comparing different automation modes afterwards. The HMI is realized using two tablet computers. The main one is

(a) Manual mode

(b) Semi-automated mode

(c) Automated mode

Fig. 5. Implemented HMI Modes which can be switched

positioned behind the steering wheel as a dashboard and represents the information part of the HMI Figure 5 on page 5. Its graphical design is principally dedicated to the LCA support. The blue squares represent the areas around the ego vehicle occupied by other traffic. In manual mode only the recommendation for possible LC will be displayed (green arrow in Figure 5 on page 5 (a)). In semi-automated mode, by activating the indicator, the driver makes a decision for automated LC execution. If the LC is not possible, the red square left to the ego vehicle (Figure 5 on page 5) indicates the unsafe situation and the system waits for a clear left lane in order to execute the LC. In the automated mode all actions are performed by the LCA system, still the driver is informed upon the oncoming LC and during the LC execution (green path arrow, Figure 5 on page 5). The second tablet computer represents the controlling unit of the HMI allowing for setting of ACC and LCA mode, see Figure 2 on page 3.

The traffic flow test concept gives a feedback about the overall driver impression regarding the usability of the

Value	age	Nr. of Participants	Init.	LC dur.
Manoeuvre 1				
μ	20 - 80	20	16.10	8.36
σ	20 - 80	20	9.03	4.06
μ	20 - 45	11	10.31	6.14
σ	20 - 45	11	2.52	1.81
μ	45 - 80	9	23.18	11.07
σ	45 - 80	9	9.14	4.47
Manoeuvre 2				
μ	20 - 80	20	13.00	5.75
σ	20 - 80	20	2.96	1.97
μ	20 - 45	11	11.68	5.63
σ	20 - 45	11	1.94	1.93
μ	45 - 80	9	14.60	5.88
σ	45 - 80	9	3.28	2.14
Manoeuvre 3				
μ	20 - 80	20	18.79	6.86
σ	20 - 80	20	7.69	2.27
μ	20 - 45	11	17.45	6.11
σ	20 - 45	11	8.58	2.21
μ	45 - 80	9	20.40	7.79
σ	45 - 80	9	6.55	2.11
Manoeuvre 4				
μ	20 - 80	20	12.26	6.41
σ	20 - 80	20	1.63	1.93
μ	20 - 45	11	12.00	6.40
σ	20 - 45	11	1.71	2.38
μ	45 - 80	9	12.57	6.43
σ	45 - 80	9	1.57	1.31

Table 3. Mean values and standard deviation of the manoeuvre initialisation and duration times

LCA systems. Hereby the implemented HMI concept, with its logic and design, is crucial. In this stage, the questionnaire focuses on the HMI assessment. The NASA Task Load Index (TLX) was used for a workload assessment Hart and Staveland (1988). The drivers completed the NASA-TLX for each mode (manual, semi-automated and automated). In addition, a questionnaire consisting of 12 items was used for evaluating the HMI usability and the drivers perceived safety and comfort when driving with the system.

3. RESULTS

The main objective of the research was to develop and validate the development and test methodology. Thus, a relatively low number of 20 drivers, regarding statistical significance, was included. Therefore results are presented descriptively.

3.1 Fixed Scenarios

Manual Driving The first two rows of Tab. 3 shows the mean value μ and corresponding standard deviation σ of the LC times for all 20 participants of four manoeuvres regarding LC initiation timing and LC duration. It is noticeable that the drivers tend to initiate overtaking quite early with mean initialisation timings ranging from 12 to 19 s for the different manoeuvres. If the LC manoeuvre is not prevented by the rear vehicle in the faster left lane (target 2), like in manoeuvres 1 and 3, the LC is initiated especially early.

The large standard deviation values from 2-9 s for the different scenarios show variations in driver behaviour.
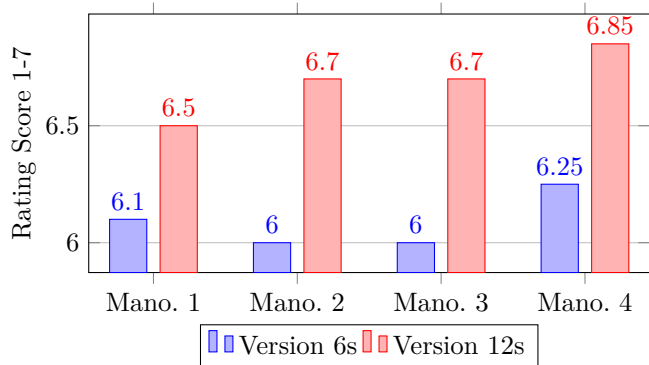
Fig. 6. General impression with the rating score

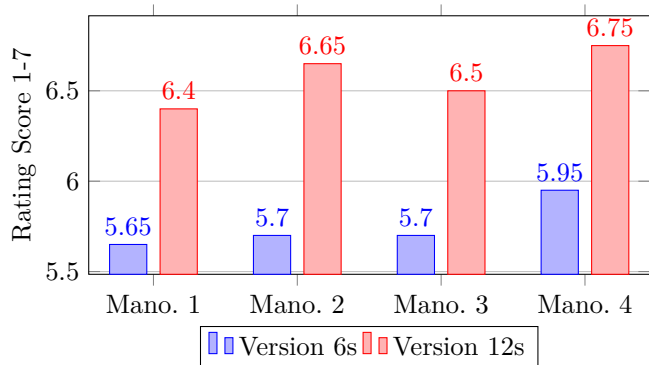Fig. 8. LC Initialisation timing with the rating score

Fig. 7. LC Duration timing with the rating score

There is a noticeable effect of age on the two variables. Mean and standard deviation value for manoeuvre 1 are lower for the age group 20-45 years than for the age group 45-80 years, see Tab. 3. Manoeuvre 3 is yet an exception because of extremely deviating values of two participants with very long initiation times (24.6 s and 40.6 s).

Automated Driving   Because the automated version with 6 s was always driven first, the drivers had the option to rate the first version again after driving the second version with 12 s. This comparison option provides certain accuracy in the rating process (see Section 2.2.1). The general impression was good for both versions. However, the scores tend to be higher for the version with 12 s compared to the one with 6 s, see Figure 3.1.2 on page 6.

Rating scores of LC initialisation timing and LC duration provide additional understanding of the drivers preferences during the LC manoeuvre in the four different scenarios, see Figure 7 on page 6 and Figure 8 on page 6. Although the LC duration was the same for both versions, participants still rated the version with 12 s better in this category, see Figure 7 on page 6 . The higher score difference for manoeuvre 3 and 4 can be explained by a higher risk perception caused by an oncoming rear vehicle (target 3) resulting in the driver's wish to reach the target lane sooner. In general, drivers preferred an early LC initiation under all tested conditions, see Figure 8 on page 6.

A closer analysis of the data reveals similar age differences for the preferred initialisation timing as in the manually driven sessions, see Figure 9 on page 6 and Figure 10 on page 6.
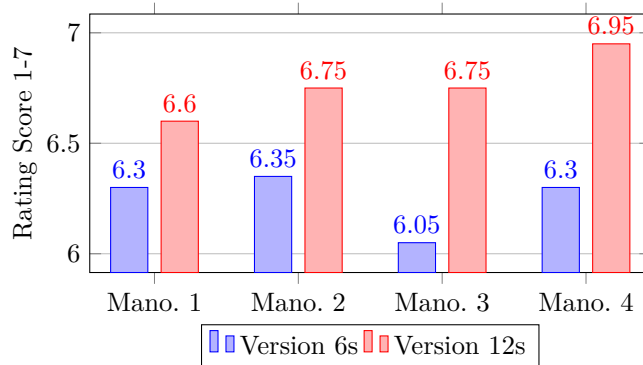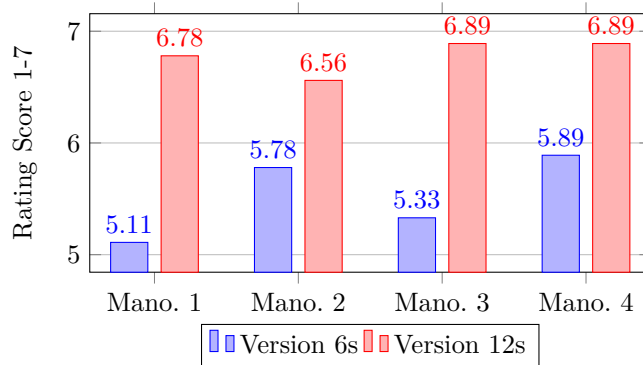

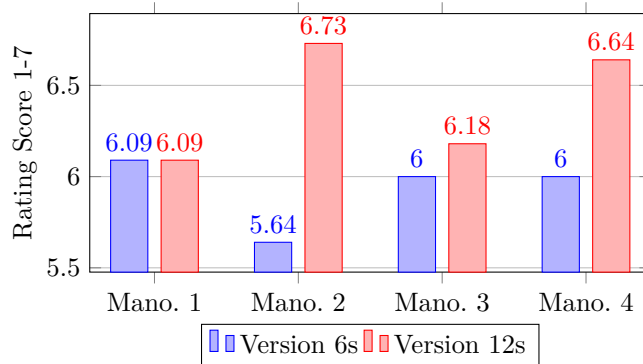Fig. 9. Initialisation timing for different the ages over 45 years

Fig. 10. Initialisation timing for different the ages under 45 years

Remarkable are the results for manoeuvres 1 and 3 where the ratings of young drivers are comparable for both versions, as opposed to older drivers who tend to favour the version with 12 s.

3.2 Traffic Flow Scenario

A general overview of the mean workload ratings shows an overall decrease of mental, physical and temporal demand, effort and frustration level with the increase of the automation level, pointing to advantages of automation as a means to reduce drivers workload, see Figure 11 on page 7.

On the one hand, the results may appear obvious due to the expectation that automation reduces the workload
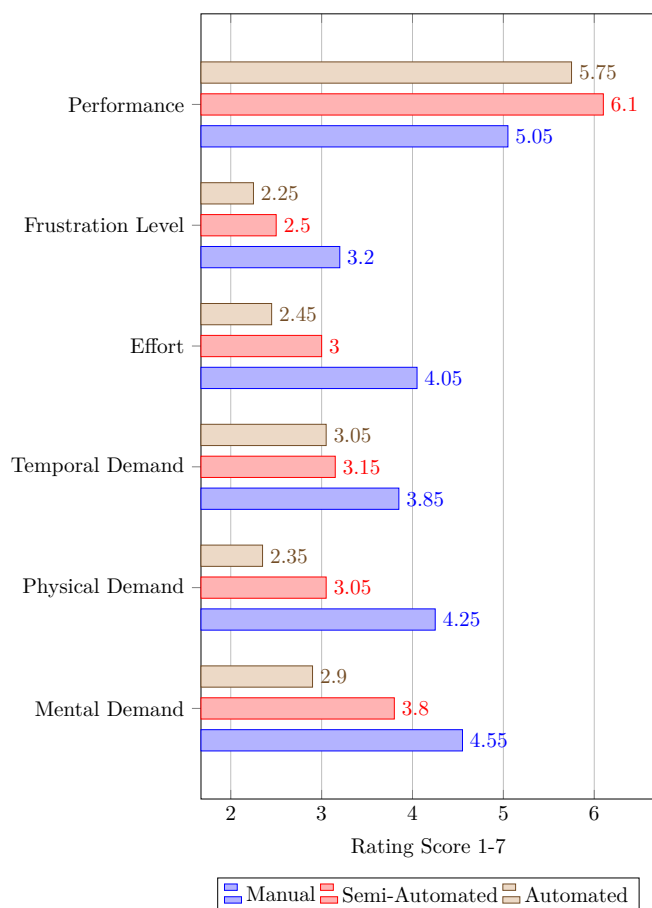
Fig. 11. NASA task load index with the rating scores

Description	
Question Q_1	The familiarisation was easy
Question Q_2	Actions are easy to understand
Question Q_3	Increases the feeling of safety
Question Q_4	Enhances the driving comfort
Question Q_5	Eases the LC manoeuvre

Table 4. Question description for the NASA task load

per se. On the other hand, it is the task of the HMI to transmit information in a comprehensive way, so that the driver can also perceive this workload reduction. The steady decrease of rating scores for mental demand from manual- to automated mode, along with the noticeable lower physical demand and effort of automated modes compared to manual mode, prove the usability of the HMI. In contrast, a badly designed HMI, causing confusions and thus frequent driver interventions in automated modes, would obviously be rated worse than the manual driving mode, especially regarding mental and physical demand, effort and frustration.

The NASA-TLX category "performance" reflects drivers satisfaction concerning their experience of driving with different modes, see Figure 12 on page 7. Here, slightly higher ratings of the semi-automated over the automated mode could be explained with the drivers wish not to be completely excluded from the driving task. It also shows the potential for improvement of the automated mode in order to satisfy driving preferences of the drivers.
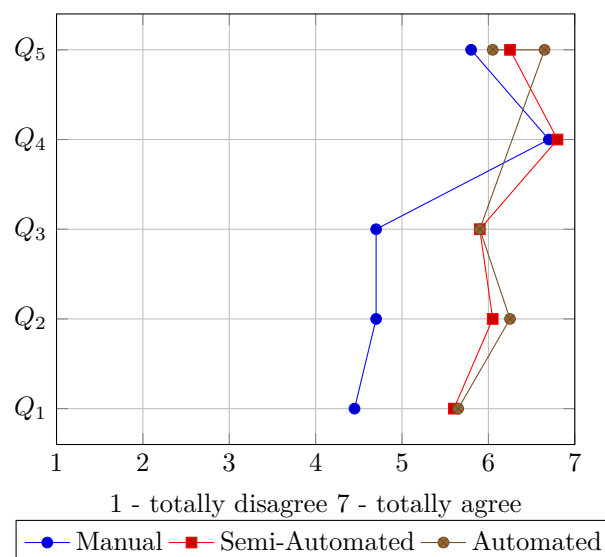

Fig. 12. Custom questionnaire with the rating scores see 4

Out of the 12 additional criteria, regarding usability, safety and comfort, a short semantic difference is presented in Figure 12 on page 7. The five most significant criteria represent the usability profile for single modes. The scale of 7 points ranges from 1 – "totally disagree", to 7 – "totally agree". Good rating scores can be stated in each five criteria for both - semi-automated and automated modes. Obviously the aspects regarding driving easiness, comfort and safety are much more related to the automated modes than to the manual one, where the driver still has to rely predominantly on his own observations and driving skills. Such a rating score picture would not be possible with a counter-intuitive system and confusing HMI for automated modes. All three modes come along with a very good assessment regarding the two criteria - HMI understandability and familiarisation.

4. CONCLUSION

A two-step test approach for development and validation of SAE level 3+ automated driving functions was presented and applied on an automated lane change using a driving simulator. The first step included a study with predefined "fixed" scenarios, specific to lane change and focused on system lay-out. The second step included a study with stochastic traffic and concentrated on HMI test and validation. The testing strategy of the fixed manoeuvres provided a possibility for comparison of the participants LC driving performances. It also assured accuracy in their ratings of the automated system because they could compare different parametrizations of the system. Based on the findings, some recommendations regarding the lane change duration and initiation timing for the execution of the lane change manoeuvre were made. The lane change duration time of 6 s was an acceptable value that should be reduced in situations when a faster vehicle is approaching from behind. The lane change initiation time of 12 s time-to-collision to the vehicle to be overtaken received high acceptance scores. However, the reduction of this time might even increase the overall acceptance for some manoeuvres. The results also show that the customisation of these parameters could also lead to higher

acceptance in a broader user scale (e.g. age dependency). Furthermore, more precise dependencies to the vehicles in the target lane for the lane change execution could be detected by additional tests based on the same method. The experiences of using an automated driving function in a naturalistic traffic flow simulation gave drivers a proper idea about the system and the usability of the Human-Machine Interface. By comparing different automation modes, the participants were able to answer the questions and rate the systems more precisely. The study will be repeated with a larger sample of volunteers and modified parameters based on this first study, prior to implementing the lane change function in a prototype vehicle that ought to be tested on public roads.

REFERENCES

Bellem, H., Schönenberg, T., Krems, J.F., and Schrauf, M. (2016). Objective metrics of comfort: Developing a driving style for highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41, 45 – 54.

Broggi, A., Buzzoni, M., Debattisti, S., Grisleri, P., Laghi, M.C., Medici, P., and Versari, P. (2013). Extensive tests of autonomous driving technologies. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1403–1415.

Brookhuis, K., de Waard, D., and Janssen, W. (2019). Behavioural impacts of advanced driver assistance systems–an overview. *European Journal of Transport and Infrastructure Research*, 1, 245–253.

Eggeling, E., Fellner, D.W., Halm, A., and Ullrich, T. (2013). Optimization of an autostereoscopic display for a driving simulator. In *GRAPP/IVAPP*.

Hart, S.G. and Staveland, L.E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (eds.), *Human Mental Workload*, volume 52 of *Advances in Psychology*, 139 – 183. North-Holland.

Henning, M. (2010). Preparation for lane change manoeuvres: Behavioural indicators and underlying cognitive processes.

Hetrick, S. (1997). Examination of driver lane change behavior and the potential effectiveness of warning onset rules for lane change or "side" crash avoidance systems.

Junietz, P., Bonakdar, F., Klamann, B., and Winner, H. (2018a). Criticality metric for the safety validation of automated driving using model predictive trajectory optimization. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 60–65.

Junietz, P., Wachenfeld, W., Klonecki, K., and Winner, H. (2018b). Evaluation of different approaches to address safety validation of automated driving. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 491–496.

Koglbauer, I., Holzinger, J., Eichberger, A., and Cornelia, L. (2017). Drivers' interaction with adaptive cruise control on dry and snowy roads with various tire-road grip potentials. *Journal of Advanced Transportation*, 2017, 1–10.

Leyn, U. and Vortisch, P. (2015). Calibrating vissim for the german highway capacity manual. *Transportation Research Record*, 2483(1), 74–79.

Martin, F. and Peter, V. (2011). *Microscopic traffic flow simulator VISSIM*, 63–93.

Nalic, D., Eichberger, A., Hanzl, G., Fellendorf, M., and Rogic, B. (2019). Development of a co-simulation framework for systematic generation of scenarios for testing and validation of automated driving systems. *IEEE Transactions on Intelligent Transportation Systems*.

Nidhi, K. and Paddock, S. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice*, 94, 182–193.

Nilsson, J., Brännström, M., Coelingh, E., and Fredriksson, J. (2017). Lane change maneuvers for automated vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 18(5), 1087–1096.

Olsen, E., Lee, S., Wierwille, W., and Goodman, M. (2002). Analysis of distribution, frequency, and duration of naturalistic lane changes. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 46, 1789–1793.

Rodemerk, C., Habenicht, S., Weitzel, A., Winner, H., and Schmitt, T. (2012). Development of a general criticality criterion for the risk estimation of driving situations and its application to a maneuver-based lane change assistance system. In *2012 IEEE Intelligent Vehicles Symposium*, 264–269.

SAE (2019). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*. URL https://doi.org/10.4271/J3016 201806.

Salvucci, D.D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors*, 48(2), 362–380.

Salvucci, D., Liu, A., and Boer, E. (2001). Control and monitoring during lane changes. *Vision in Vehicles*, 9.

Samiee, S., Azadi, S., Kazemi, R., and Eichberger, A. (2016). Towards a decision making algorithm for automatic lane change maneuver considering traffic dynamics. *Promet - Traffic and Transportation*, 28.

Satzoda, R.K. and Trivedi, M.M. (2016). Safe maneuverability zones metrics for data reduction in naturalistic driving studies. In *2016 IEEE Intelligent Vehicles Symposium (IV)*, 1015–1021.

Sporer, A., Prell, G., Buck, J., and Schauble, S. (1998). Realsimulation von spurwechselvorgängen im straßenverkehr. *Verkehrsunfall und Fahrzeugtechnik*, 36, 67–76.

Stellet, J.E., Zofka, M.R., Schumacher, J., Schamm, T., Niewels, F., and Zöllner, J.M. (2015). Testing of advanced driver assistance towards automated driving: A survey and taxonomy on existing approaches and open questions. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 1455–1462.

van Driel, C., Hoedemaeker, M., and Arem, B. (2007). Impacts of a congestion assistant on driving behaviour and acceptance using a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(2), 139–152.

Wang, F., Ma, N., and Inooka, H. (2001). A driver assistant system for improvement of passenger ride comfort through modification of driving behaviour. *International Conference on Advanced Driver Assistance Systems*, 38–42.