
Driver Engagement

Adria Roig-Lafon and Francesco Deiana

Applus+ IDIADA, Spain

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 behaviour. 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 behaviour. 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 behaviour 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 (SAE, J3016-2014), 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. 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.

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 et al., 2013, 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., 2015 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 (UN Regulation No. 159) 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 (Euro NCAP 2030).

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. Brookhuis et al., 2009 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 et al., 2005 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 et al., 2004 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 (Hoc et al., 1994).

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 et al., 2008 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 and

formulated TTC as follows.

$$TTC = \frac{D}{\Delta V} \quad (1)$$

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 behaviour.

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 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 Vector Kit providing bus CAN data, RT and RT Range acquiring geolocation and vehicle dynamics, three video cameras recording at 30hz, and all components synchronized with CANape software on a CPU.

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 section “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 participants underwent a familiarization period of 10 minutes without any driving assistance systems, to acclimate themselves to the vehicle and track. This was followed by a 30-minute period of driving with the adaptive cruise control set at 60 km/h and all driving assistance systems enabled. 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 ADAC car dummy was placed in the middle of the lane without warning the participants. As depicted in [Figure 1], the obstacle was placed immediately after a bend in the road, as shown in Figure 1, 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 manoeuvre 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.

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: Non-professional drivers, 3/4 of participants without any experience with partial driving automation, 1/4 of participants with experience with partial driving automation, equal distribution of males and females (20/20), have a valid driver’s license. 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

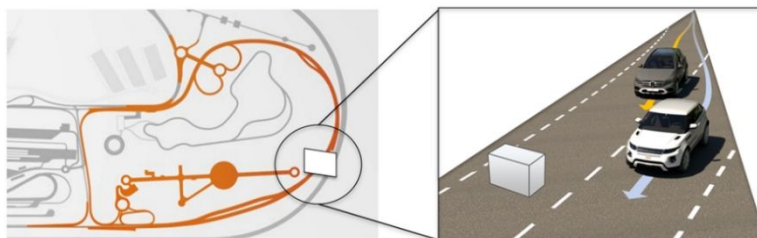


Figure 1: Highway loop B and manoeuvre description.

complete the DSQ questionnaire, (French et al., 1993) 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. The driving style participant distribution were: 4 participants with SPEED driving style, 3 participants with CALM driving style and 30 participants with FOCUS driving style.

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**. The Integrated Workload Scale (IWS) is a measure of a person's experienced peaks and troughs in workload in a specific scenario or interval. It consists of 9 points, asking the participant to rate their mental workload. The TASS (Trust in Automated System Survey) assesses trust between humans and machines and helps understand the impact of system characteristics on drivers. Participants 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 considered: Distance (D) at the start of swerve-avoidance and Relative Velocity (ΔV) at the start of swerve-avoidance between vehicle and stationary target.

To determine the threshold limit value, reference was made to the following criteria, defined by Saffarzadeh et al., 2013: “*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 2] and [Figure 3].

Points in [Figure 2], represent the average level of mental workload value for participants, differentiated between the two cars, while the vertical lines in each point represents the standard deviation. Same in [Figure 3] but as far as the trust level. As a static obstacle, the target has speed 0 km/h, whereas the

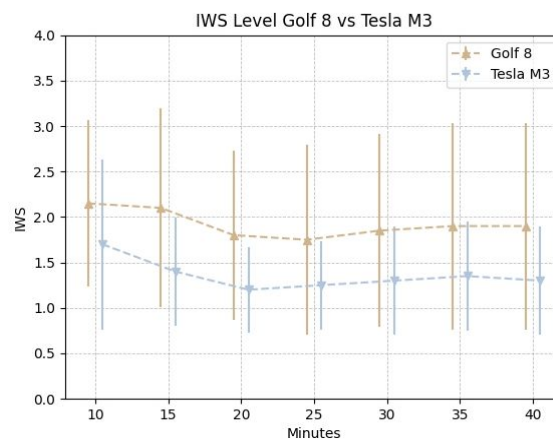


Figure 2: Workload level golf 8 vs tesla model 3.

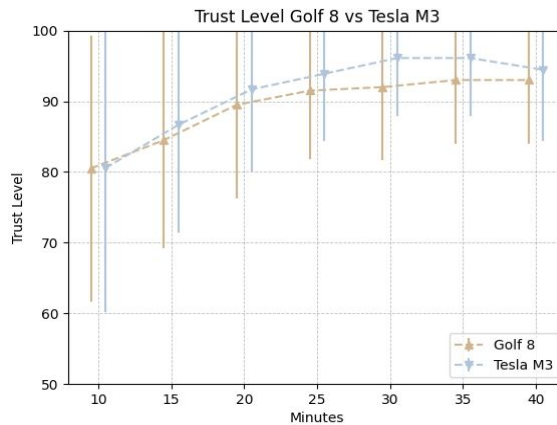


Figure 3: Trust level golf 8 vs tesla model 3.

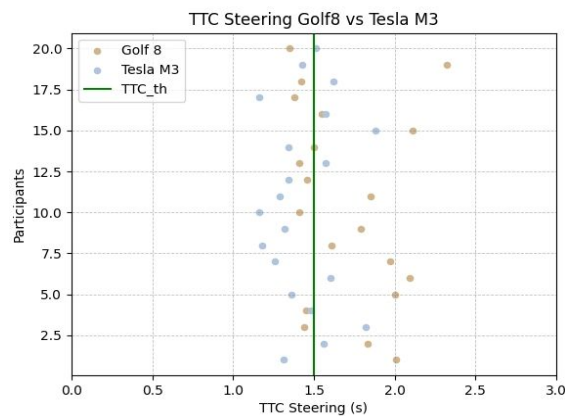


Figure 4: TTC steering golf 8 vs tesla model 3.

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 vehicles [Figure 4]: Participants below and above TTC threshold are distributed: Golf 8 with 45% below threshold and Tesla Model 3 with 60% below threshold.

CONCLUSION

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 2] and [Figure 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, it 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. 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.

REFERENCES

- 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.

- 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.
- Hoc, J., Cacciabue, P. C., Hollnagel, E. & Cacciabue, C. P. (1994). *Expertise and Technology: Cognition & Human-Computer Cooperation* (1). Psychology Press.
- 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.
- Latest Safety Ratings | Euro NCAP. (s.f.). <https://www.euroncap.com/en/ratings-rewards/latest-safety-ratings/>.
- 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.
- 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.
- 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.
- Peters, J., & Mattes, K. (2015). How driver engagement and inattention lead to accidents. *Journal of Safety Research*, 56, 79–87.
- 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.
- SAE, J3016 (2014). *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems*.
- 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.
- 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.
- 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.
- 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].
- Vision 2030 | Euro NCAP. (s. f.) <https://www.euroncap.com/en/press-media/press-releases/euro-ncap-vision-2030-a-safer-future-for-mobility/>.