Submitted by the expert from IEA

Informal document **GRVA-19-10**
19th GRVA, 25 June 2024
(For review during the
Troy meeting 20-24 May 2024)
Provisional agenda item 4 (i)

White Paper from HF-IRADS[1]
"Requirements for Safe in-Vehicle Interaction with Driving Automation Systems"

1. Purpose of document

This paper is intended to provide advice to GRVA as work on Automated Driving Systems (ADS) moves from the preparation of recommendations to the drafting of regulations. It can also be noted that the first version of a regulation on a Driver Control Assistance System has already been approved by WP.29. At this juncture, we wish to highlight the need to continue the work on user aspects that has previously been undertaken in the Functional Requirements for Automated Vehicles (FRAV) group and in the Integration Group combining the recommendations of FRAV with the output the group on Validation Methods for Automated Driving (VMAD). The paper addresses issues of user interaction both in vehicles designed to be driven manually at time and in those which cannot be driven manually, although it needs to be acknowledged that there are greater safety risks from failures in human-automation interaction in vehicles that offer multiple roles of operation.

2. The need for safe interaction with the user

Vehicle automation does not replace humans; rather it changes their roles. In the case of SAE Level 3 automation, the human becomes the fallback and a failure to respond appropriately could result in the vehicle being placed in a dangerous condition. In the case of the provision of Level 4 automation on a vehicle with multiple levels of automation, the manual driver becomes a passenger, but that user still needs to understand how responsibilities with automation enabled differ from those with Level 2 assistance enabled. Hence good interaction design is vital for safety. System providers need to ensure that systems are easy to use appropriately and hard to use inappropriately. ADS systems need to be designed to reduce user confusion, guard against human errors and promote quick human response. Thereby appropriate design for safety, i.e. human-centred automation will reduce the risks of problems arising and crashes occurring.

3. What is human-centred automation?

Based on Billings (1997), we can identify a set of core tenets of human-centred automation:

1. Users must be involved

2. Users must be informed

3. Humans must be able to monitor the automation

4. Automation must be predictable

---

[1] "Human Factors in International Regulations for Automated Driving Systems" (HF-IRADS) operates under the auspices of the International Ergonomics Association (IEA). It brings together human factors experts from across the world to support UNECE activities on the safety of automated driving systems.

5. Automation must monitor the human (input/state)

6. Intent must be dually communicated between automation and human

These tenets indicate the core role of interaction between system and user in determining vehicle safety.

4. Safety problems arising from poor interaction design

Problems in user interaction with driving automation systems have been identified as being at the root of real-world crashes. The U.S. National Transportation Board, following its investigation of a fatal crash between a vehicle operating with L2 support on 23 March 2018, recommended to the U.S. National Highway Traffic Safety Administration (NHTSA) to: "Evaluate Tesla Autopilot-equipped vehicles to determine if the system's operating limitations, the foreseeability of driver misuse, and the ability to operate the vehicles outside the intended operational design domain pose an unreasonable risk to safety; if safety defects are identified, use applicable enforcement authority to ensure that Tesla Inc. takes corrective action."[2]

More recently, the U.S. National Highway Traffic Safety Administration (NHTSA) has questioned whether a manufacturer's remedies to its Level 2 assistance system are sufficient "to address misuse, mode confusion, or usage in environments the system is not designed for."[3] NHTSA identified "at least 13 crashes involving one or more fatalities and many more involving serious injuries in which foreseeable driver misuse of the system played an apparent role." This has now been followed up with a formal recall query which includes a request to the manufacturer to explain and describe "what human factors considerations and principles were used when designing [a particular interaction] feature", and to "explain how the design was validated using human factors (including, but not limited to human participant evaluations)."[4]

The safety of Driving Automation Systems (DAS) is frequently described in terms of functional safety; however, the design of the interaction between vehicle systems and users in the vehicle is also safety-critical. Most obviously this applies to the person in the driving seat in vehicles with multiple levels of automation, but it also applies to passengers in single-level vehicles such as shuttles and robotaxis, where passengers need to understand, for example, how to exit the vehicle in case of emergency.

Some examples of foreseeable problems in interaction design that will lead to safety-critical situations are:

- Driving with a Level 2 hands-off system and thinking it is an ADS. This has already been shown to lead to inattention, and occupants being out of position, leading to crashes.
- Increased engagement in non-driving tasks after activating L2 assistance because of over-trust in system capabilities, which delays the driver's ability to promptly respond to emerging situations.
- Engaging automation in situations outside the operational design domain (here even a user being locked out of this possibility could lead to frustration and distraction if the interaction is not properly designed) where it is incapable of performing the DDT safely.

---

[2] NTSB Safety Recommendation H-20-002, https://data.ntsb.gov/carol-main-public/sr-details/H-20-002.
[3] https://static.nhtsa.gov/odi/inv/2024/INOA-RQ24009-12046.pdf.
[4] https://static.nhtsa.gov/odi/inv/2024/INIM-RQ24009-12199.pdf

- Silent transitions to less capable automation models, such as from L3 to Adaptive Cruise Control (ACC) or Driver Control Assistance System (DCAS), where both the fallback user and automation are unable to safely perform the dynamic driving task.
- Patchy automation that lacks coherence, such as driving with L3 without lane change authority and subsequently switching back to Level 2 with lane change and an obligation on drivers to check their blind spots, which causes the driver to be confused about their role and the current capabilities of the system.
- In driving with an L3 system, the driver monitoring system DMS detects the common risk of the fallback user falling asleep but fails to effectively re-engage them when it encounters a situation it is not capable of handling safely.

5. How do we ensure safe interaction between vehicle and user?

Both intended use of systems and anticipated intentional and unintentional misuse must be considered. Therefore, there is a need to support user understanding for intended uses and perform a detailed safety analysis to address anticipatable misuse. Interaction design goes far beyond the design of visual screens, symbols and auditory and multimodal messages. It needs to address user understanding of changeable user roles, mode structure and interactions in, for example, enabling the Automated Driving System (ADS) and responding to requests to intervene.

System limitations must be clearly communicated to drivers, operation outside its Operational Design Domain (ODD) should be restricted, and robust mechanisms for monitoring and engaging drivers are necessary to mitigate foreseeable risks of complacency and inattentiveness.

6. Usability underlies safety

ISO 9241-11:2018 defines *usability* as "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". *Effectiveness* is defined in the same document as: "accuracy, completeness and lack of negative consequences with which users achieved specified goals". Here negative consequences include death and injury to the vehicle occupants and other road users. *Efficiency* is defined as "resources used in relation to the results achieved." There is a detailed discussion of the attributes of usability in Nielsen (2010).

The importance of the different usability dimensions differs depending on the context and the target users. In case of complex systems, such as automated vehicles in which there are safety critical tasks, it is essential to guarantee *effectiveness*, in terms of error prevention, recovery and tolerance and *efficiency*, in terms of reduced task execution time (interaction with a system should be as simple as possible) and learning time (where interaction should ideally be intuitive). Subjective aspects such as user satisfaction are also important.

6.1 Human-Centred Design contributes to usability and safety

Following a human-centred design (UCD) process is a core element in delivering usability and safety. The ISO 9241-210:2019 standard describes six principles to be followed to ensure the design is Human-Centred:

1. The design is based upon an explicit understanding of users, tasks and environments. This principle is about understanding users' "context of use".

2. Users are involved throughout the design and development process. This principle is about ensuring the users are involved in all design phases, avoiding running a focus group during the initial phase of the design and a survey at the end. Active user involvement is required, in which users are engaged in the design through dedicated studies in which they can use artefacts.

3. The design is driven and refined by user-centred evaluation. The standard recommends including usability testing throughout the design process, using different types of prototypes until the final design product/service is reached.

4. The process is iterative. The standard describes that the most appropriate design can be achieved only through iterative improvements.

5. The design addresses the whole user experience. The standard points out that, beyond usability, there are important emotional aspects and "include the kind of perceptual and emotional aspects typically associated with user experience."

6. The design team includes multidisciplinary skills and perspectives. It is of paramount importance to include different views and competences in the design team and to include a person with a background in human factors, ergonomics or cognitive engineering and experience in crafting human-technology interactions.

6.2 Iterative process of User-Centred Design

Moreover, HCD defines four activities as shown in Figure 1. These activities have to be **iterated** until the system satisfies the specified user and organisational requirements. At stage 4, designs need to be tested interactively, i.e. in a driving simulator or in real driving.

Figure 1: Iterative process of HCD. Adapted from ISO (2019)

The HCD process is the basis for developing user interface/interaction that is less vulnerable to user errors. In fact, human error analysis has an important role at each iterative step because, even if it is not possible to design a system in which users make no errors, it is possible to address error management aspects, through:

- Preventing errors through the design and understanding possible ones that may occur
- Minimising errors when full prevention is not achievable

4

- Ensuring easy detection and identification of the error of the user, if it still occurs
- Aiding error correction to quickly regain safe system behaviour
- Minimising the consequences of errors to avoid a catastrophic outcome

7. Role and mode awareness

The immediate role of the user, and any impending user role, should be obvious. However, there is currently a lack of recommendations on how role should be indicated to the user. A small indication on the dashboard is not sufficient. With roles go responsibilities and therefore we need recommendations on how to indicate responsibilities such as keeping one's eyes on the road in the case of DCAS, and the absence of a requirement to maintain eyes on the road in the case of an ADS being enabled.

Mode confusion is already evident in interaction with vehicle automation – leading to safety issues e.g. when people think lane keeping is switched on whereas it is not and they run off the road. The potential for such mode confusion will only increase with increasing varieties of driving systems capable of driving under various conditions and Operational Design Domains (ODDs).

Mode awareness can be enhanced by avoiding automated resumptions of previous modes and therefore transitioning to full manual driving from L2, L3 or L4. It can be degraded by very similar capabilities between L2 and L3, where the driver roles and responsibilities are vastly different.

Integrated design across automated features is needed because of the increased variety of systems, vehicles, ODDs. The potential for users transferring from one vehicle to another in the case of car rental, car sharing, and multi-vehicle households creates the case for commonality and a holistic approach to design.

8. Mental models and skill acquisition

A mental model is the user's concept of the functionality and operation of a system. It is built up over time. A user's mental model will often not conform to the designer's conceptual model. "The designer expects the user's model to be identical to the design model. But the designer does not talk directly with the user—all communication takes place through the system image. If the system image does not make the design model clear and consistent, then the user will end up with the wrong mental model." (Norman, 2002, p.16). UCD helps to enable conformity between the conceptual model of the designer and the user.

The initial mental model will be formed at the knowledge and rules level in the skills-rules-knowledge hierarchy and may have fundamental errors. Training can work to refine the mental model of a system and move the mental model into the skill category, where users react automatically in interaction with the system. However, requiring extensive training is impractical in the driving domain. So we need to establish a minimum form of instruction prior to first use that is sufficiently effective. This could be in the form of offline training or a short on-line tutorial. Such training should also be developed via a UCD process. The current DCAS regulation requires documentation but does not have any requirement to prove that this documentation is consulted or effective.

It is only by experience in using a system that users develop accurate mental models and acquire an accurate understanding of the procedures to use a system so that they can apply them at a skill level

(Kim et al., 2013). Then, for example, a user will respond instinctively to a request to intervene. Applying UCD to deliver enhanced usability and simplicity will hasten the process of reaching the skill level of interaction and help to ensure that the mental models of designers and users coincide.

9. Diversity of users

Depending on the type of vehicle, the user can be a passenger or a fall-back user. With both types of users, the diversity in skills, the variation in capabilities and physical impairments among users must be taken into account in the design of the interaction processes. To obtain the expected benefits of automation, the HMI should be iteratively designed to be fitted to the diversity of needs of these users and their expectations. Therefore, in the development phase, tests of interaction should be performed on the range of potential users and tested in a range of driving environments. Particular issues are that many potential users can have different characteristics, such as colour-blindness and many other users could suffer from chronic hearing loss, which means that they would not perceive acoustic warnings.

10. Recommendations to GRVA

As we move from preparing *recommendations* on system design to developing *regulations* on ADS validation and approval, we propose that:

1. Interaction safety must be set as an objective and integrated in the safety case for a DAS since it is mission-critical.

2. The use of Human-Centred Design process with the involvement of the expected range of users in the different stages of the development of driving automation systems must be verified.

3. There is a need to develop verifiable requirements on interaction, including on driver monitoring.

4. The Human-Centred Design process and established HMI principles be applied to create a high-level commonality of interaction design across vehicles and levels of automation. This will assist users in easy adaption to new vehicles and in switching from one vehicle to another in their daily use.

5. The consequent design recommendations could then be applied in feature development and be used in checklists at the verification or approval stage, which would substantially reduce development costs.

6. It will be Important to draw lessons from post-production In-Service Monitoring and Reporting to inform recalls and the refinement of driving automation system safety. Any substantive changes to interaction with driving automation systems, such as through over-the-air updates, require verification and notification of the changes to drivers.

7. Above all, we need an integrated approach that spans automation across the different SAE levels of automation in order to accommodate safe user interaction with these levels on a single vehicle.

References and further reading

Billings, C.E. (1997). Aviation Automation: The Search for a Human-Centered Approach. Mahwah, NJ: Lawrence Erlbaum Associates.

Collins, J. G. (1997). Prevalence of Selected Chronic Conditions: United States, 1990-1992. Vital and Health Statistics. Series 10, Data from the National Health Survey, 194: 1–89.

Coppin, R. S. and Peck, R. C. (1965). The Totally Deaf Driver in California. Highway Research Record, 79: 35–44.

ISO DIS 9241-11 - 2018: Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts (2018)

ISO 9241-210 - 2019: Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems (2019)

Kim, J.W, Ritter, F.E. and Koubek, R.J. (2013) An integrated theory for improved skill acquisition and retention in the three stages of learning. Theoretical Issues in Ergonomics Science, 14(1): 22-37 https://doi.org/10.1080/1464536X.2011.573008

Nielsen, J. (2010). What Is usability? In: C. Wilson (ed), User experience re-mastered: your guide to getting the right design, pp. 3-22. Burlington, MA: Morgan Kaufmann Publishers. https://doi.org/10.1016/B978-0-12-375114-0.00004-9

Norman D. and Draper S.W. (eds) (1986). User Centered System Design, Hillsdale, N.J.: Erlbaum.

Norman, Donald A. (2002). The Design of Everyday Things. New York: Basic Books.

Wegge, K.P. and Zimmermann, D. (2007). Accessibility, usability, safety, ergonomics: concepts, models, and differences. In: Stephanidis, C. (ed.) Universal Access in Human Computer Interaction: Coping with Diversity. UAHCI 2007. Lecture Notes in Computer Science, 4554: 294-301. https://doi.org/10.1007/978-3-540-73279-2_33