Advanced usability testing framework of in-vehicle digital user interfaces

Róbert Junas[*]
Supervised by: Miroslav Laco[†]

Faculty of Informatics and Information Technologies
Slovak University of Technology
Bratislava,Slovakia

Abstract

As digital products become more prevalent within the automotive industry, user experience and usability have become critical factors in developing in-vehicle systems, particularly infotainment interfaces. Testing of these interfaces poses a significant challenge in the conduct and evaluation. The testing needs to evaluate not only the interface's usability but also help to mitigate the driver's distraction factor caused by these systems. In our research work, a novel testing framework for use in simulator-based environments was developed by leveraging human-centered design principles, eye-tracking technology, and novel metrics.

Keywords: automotive, user experience, usability, usability testing, eye-tracking, mobile eye-tracking, distraction, infotainment

1 Introduction

In recent years, automobile manufacturers have been using increasingly complex systems e.g. infotainments systems also known as *In-Vehicle Information Systems* (IVIS) [21]. IVIS allows drivers to use touch screens to browse music, use messaging and maps directly on the vehicle's dashboard [22]. While these systems can enhance the driver's experience, they may also cause unwanted distractions and create potential risks to the driver as well as their surroundings. According to NHTSA [18], between the years 2017 and 2021, around 8% (or 3000) of all fatal crashes could be attributed to distracted driving, and approximately 14% of all crashes were caused by distractions. The most observed types of distractions are cell phones, IVIS, and navigation systems, and others [6]. IVIS research has shown that many features should be disabled, but on the contrary, a trend emerged of expansion of functions and features. One of these features is Apple's CarPlay[1] and Google's Android Auto[2] [21].

Usability and usability testing of automobiles requires the development and design of interfaces for Human-Machine Interactions. For this process, a classic user-centered process can be used [1]. It is important to keep in mind that controlling the vehicle is **the primary task** and usage of the designed system is always **a secondary task** [1, 17]. Validation of the created interfaces in the automotive industry can be targeted at the driver or front passenger [20] and comes in forms of reports of accidents, surveys, field studies, laboratory experiments, task simulations [1]. Systems could be evaluated on the basis of usability of the system (task completion etc.), but a more important measurement is the distraction of the driver.

2 Related Work

Studies can be categorized into the use of simulators or field studies to measure the usability of vehicle systems. Field studies are conducted with a working car in real traffic or on a racing track [16, 5, 12]. Drivers often wear eye-tracking glasses and the car interior is modified to fit the experiments. On the other hand, studies done in the field have a problem with testing dangerous driving situations (like blind spot overtaking, pedestrian crossing, and so on) [23]. Therefore we are more interested in usability testing in simulators that do not possess this limitation, but used simulators vary greatly in complexity and immersion. The simplest simulator consisted of one LCD screen rendering the whole cockpit of the vehicle and steering wheel attached to a table [14, 15]. Other simulators consist of one or multiple screens rendering the virtual environment, while drivers have access to a car seat, steering wheels, and other controls [7, 3, 13, 23]. Similar simulators can be created with mixed reality, which does not show significant differences between real driving and simulators [2]. In extreme cases, the simulator consists of a real vehicle in front of which the virtual environment is projected [19, 11].

In the study by Naujoks F. et al. [17] a methodological framework for the evaluation of human-machine interactions (HMI) with automated driving systems (ADS) was proposed. This framework is based upon specific use cases. In the initial step of the evaluation framework, it

---

[*] rob.junas@gmail.com

[†] miroslav.laco@stuba.sk

[1] Apple CarPlay - `https://www.apple.com/ios/carplay/`

[2] Google Android Auto - `https://www.android.com/auto/`

is recommended to do a heuristic evaluation built on the provided checklist, which consists of, but is not limited to, prevention of unintentional activation/deactivation and continuous display of the system's state. The next step is empirical evaluation whose purpose is to investigate if use cases are falsified. The framework specifies that testing should be conducted with no less than 20 participants from diverse age groups with comparable knowledge about ADS. Participants should also be informed about the procedure and capabilities of ADS and have enough time to interact with the ADS. The recommended test environment is a driving simulator that provides realistic controls and detection. Another evaluation methodology was proposed by Crescenti R. et al. [5] which uses eye-tracking technology to evaluate infotainment systems with delimits. To showcase the protocol, a test was conducted in the field on a track. The protocol consists of two phases. First is training, in which participants get instructed on how to perform different tasks on vehicle systems to ensure a standard level of knowledge. The second phase is the test where participants have to complete 14 tasks 3 times. The result evaluation consists of the total task time, different gaze, and fixation metrics on the infotainment system. They validated their solution with 10 vehicles and found similar distractions for each vehicle.

Infotainment systems were also studied by Nagy V. [16]. They focused on investigating the driver's behavior, cognitive load, and level of distraction of the infotainment systems compared to physical buttons. To measure them, They used psychological tests, questionnaires, eye-tracking, and pupillometry. The study was conducted on a racing track using Volkswagen e-Golf and participants were given tasks to do while driving at different speeds. Results showed that using an infotainment system was slower and was more cognitively demanding. On the other hand, the faster the vehicle was moving, the task completion was done faster. Similarly, eye tracking was used in a study by Jung S. [10] which measured the effect of the button styles (shapes, colors, etc.) on the usability of an infotainment system complemented with cognitive load metrics. The study found that the style of the buttons does not have an effect on glances on the infotainment system.

A simulator framework was introduced in a study by Bolder A. [2] for usability testing in mixed reality. The virtual environment was created with Unity3D and consists of a mixed reality headset, steering wheel, car seat, pedals, and shift stick. The validation of the simulator was conducted on a simple task regarding infotainment systems and compared to the same task done in a real car. The result showed that the usability of the system and time to completion didn't vary significantly, but the readability of the text was the main issue reported by the participants. On the other hand, the gaze of the participant was not measured. Also, Hildebrandt M. et al. [9] created a collection of tools for managing complex data in behavioral research for roads and aviation simulators called Synopticon. Synopticon consists of four modules: one for mapping eye and hand tracking data into a 3D environment model in real-time. The second module synchronizes multiple video/audio recordings from an unlimited number of cameras from multiple computers. Then there are modules for generating and presenting the task to users and modules for analysis of data using machine learning.

3 Contribution

The main goals of this paper is to propose a formalized methodology and framework for usability and user experience testing in the automotive industry building upon statistically relevant quantitative and corresponding quantitative outputs relying on the state-of-the-art and novel proposed usability metrics resulting in a usability testing framework for a car simulator focused on infotainment testing while performing primary driving task. This framework consists of the establishment and interpretation of metrics, collection, and analysis of data in the context of vehicle usage.

To validate the proposed framework, we proposed a set of research questions:

Q1: We are able to propose a measure of distraction for the effectiveness of the design of infotainment systems.

Effective design of the infotainment system should minimize time not concentrating on driving, or in other words, minimize the amount of needed interactions with IVIS.

Q2: Objective measures are more effective in the formalized evaluation of the usability testing of complex driving tasks in comparison with self-reported subjective measures.

Objective measures such as gaze, interactions, and time to complete tasks might be better at identifying the complexity of tasks and the usability of vehicle systems than subjective measures (thinking aloud, post-testing interviews, and questionnaires) which might be prone to change in cognitive processes during tasks or misremembering experiences afterwards.

Q3: The information gain from visual scene converges with an increasing amount of participants.

Visual attention-supported metrics are proposed to verify and standardize the well-known usability testing methodology of a few participants revealing most of the information gain without the need for vast testing on a statistically relevant sample.

4 Research proposal

In this section, we introduce a novel approach to measuring the driver's distraction caused by infotainment systems as well as introduce the software solutions needed to measure distraction and answer our research questions.

4.1 Method

This study focuses on task-based usability testing. The tasks are given by a moderator while the participant continuously drives inside a simulator.

To facilitate this testing, we propose a novel methodological as well as a simulation and analysis framework. The simulator was chosen because it allows to test different driving scenarios without endangering the driver and others. The simulator framework must allow the user to drive freely within the simulator while their gaze, as research suggested, is captured with eye-tracking technology. Furthermore, it is important to define the areas of interest (AoIs) within the vehicle simulator and the way to measure the gaze within the AoIs while wearing eye-tracking glasses that allow free head movement. In our study, we monitored gaze on the road and the IVIS system.

Furthermore, the gaze data should also be visualized with heatmaps and scan paths, to help researchers better aid the analysis on top of defined metrics (more in section 4.3). To create visualization and calculate metrics, the start and end of tasks are set during testing whenever researchers consider the tasks have begun/ended.

Moreover, a physical environment with IVIS module has been created (diagram of the setup is in figure 1). This environment is in the form of a simulation of a real car interior, in which participants conduct the testing scenarios. Also, multiple driving scenarios (driving tasks) which include the use of the IVIS need to be created to measure as many possible use cases of the IVIS module and their distraction.

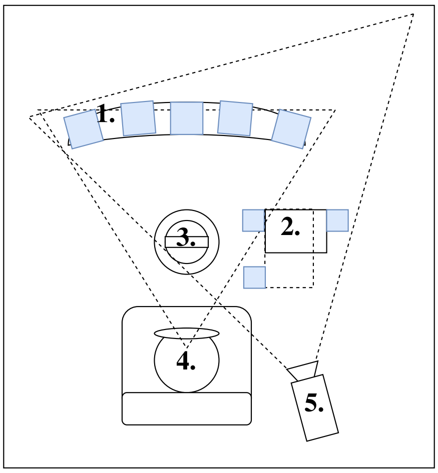


Figure 1: Diagram of the top view from of the proposed simulator with following components: 1. main screen with a driving simulator, 2. tablet screen representing IVIS (either vertically or horizontally), 3. steering wheel, 4. participant wearing an eye tracking device, 5. static camera viewing the whole setup. Additionally, blue squares represent the theoretical position and size of used markers. The markers on the main screen (1.) should be present on the top and bottom of the screen. Dotted lines represent views from both eye-tracking and static cameras.

4.2 Framework

The testing environment is a simulation of the driver's cockpit created by *the Faculty of Architecture and Design, Slovak University of Technology in Bratislava*, with collaboration from *Škoda Auto*, and has a modular interior structure that is able to house interchangeable elements of Škoda's mass-produced cars (shown in figure 2). The driver's cockpit consists of a driver seat, a steering wheel and pedals, ultra-wide monitor, an Android/iOS tablet, and a static ultra-wide angle web-cam to capture the whole scene.

On the main screen, a simulation of driving is shown with which participants interact with a steering wheel and gas/brake pedals with automatic transmission. On the second screen (or simulator of IVIS) a prototype of the tested application is shown. Similarly, participants can and must interact with the second display. During testing, participants have to wear a mobile eye-tracking device to collect gaze data and effectively measure distractions. To better visualize and understand the collected eye-tracking/interaction data, all screens and camera views (eye-tracking camera and static camera) are recorded.

Figure 2: The implementation of the proposed setup created at the Faculty of Architecture and Design. The markers on are placed by us according to our proposal and pilot testing.

.

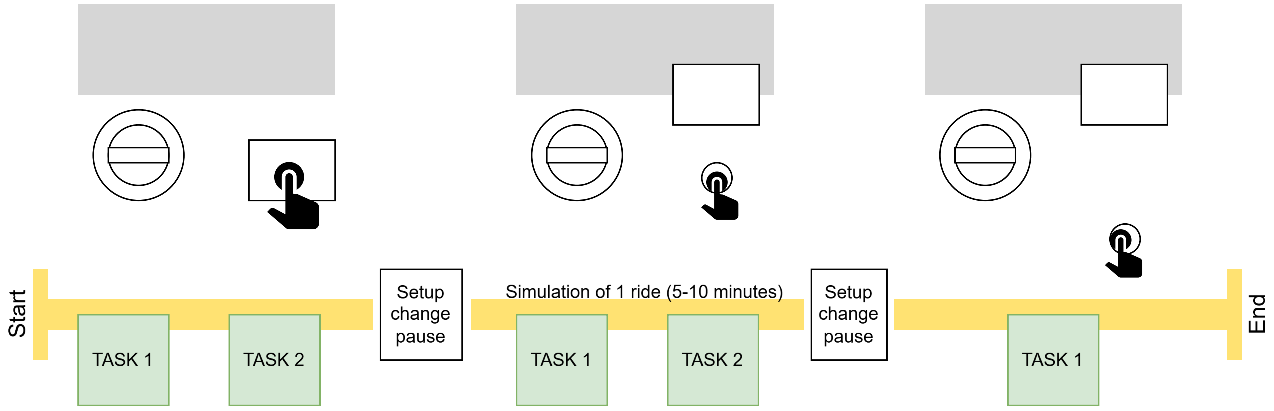


Figure 3: Experiment task progress. The figure shows the different IVIS setups used in the experiment in the order they were conducted: touch-based, parallel knob, perpendicular knob.

4.2.1 Enabling remapping of data

Because of the nature of mobile eye-tracking technology, we need to be able to map participant gaze into the screen recordings and another static view camera. To map gaze the data from eye-tracking camera into other views we propose to use the AruCo marker mapping methodology because reviewed studies [16, 5] build upon them and they are generally better at finding device borders when the reference changes over time and when the front camera of the eye-tracker is not sufficient to recognize content on the screens (because of brightness or other factors) [4].

The markers are placed on the sides/top/bottom of the screens in a way not to be obstructed (either by hands or bad angle) from the view of the eye-tracking camera while participants conduct the testing. The size of the markers vary depending on the closeness of the screen to the participant. The further screen requires larger markers, but not too large not to fit (therefore not detected) into the view of the eye-tracking camera. For a similar reason, the closest screen needs to have smaller markers. The placement of the markers can be seen in the diagram in figure 1.

4.2.2 Calibration

Calibration of the recording devices follows a structure, in which it is possible to synchronize created recordings. In our approach, we first calibrate the eye-tracking device, afterwards we start recording on the non-eye-tracking devices simultaneously and after a second passes, we start recording on the eye-tracking device. Similarly, the end of non-eye-tracking recordings ends a second after the eye-tracking recording stops. The synchronization error can be neglected as the starting time of each recording can be set manually in relation to the eye-tracking recording in a thorough analysis of the recordings. Moreover, before starting the study, we will display a marker on the screen. The participant will be asked to look at the screen, when the marker disappears, we will start conducting tests. This calibration marker will work for automatic synchronization during preprocessing.

4.2.3 Testing procedure

Finally, testing is done by doing task scenarios. In these scenarios, participants are asked to complete a secondary task requiring interaction with IVIS while conducting the main task driving. Both the start and the end of the tasks are logged to better aid the analysis of the tasks. Tasks have a correct path that the participant needs to complete the task successfully. However, the tasks end when the participant tells the moderator to end the task, which may result in failure in task completion. During tasks, participants are asked to think aloud about what they are about to do and how they feel about the interactions.

The experiment was conducted at the Faculty of Architecture and Design, Slovak University of Technology in Bratislava, and was moderated in the Slovak language. Five participants from ages 18 to 25 who have valid driving licenses and own vehicles, took part in the experiment. Two of these participants were female. The participants had a few minutes to learn to drive inside the simulator as well as try all the necessary interactions with the IVIS prototype. The IVIS prototype created for the experiment was created to mimic the appearance of CarPlay. The prototype supported two simple scenarios, which were also the tasks the participant had to complete:

- Call and hang up a contact in the driver's favorite contacts,
- Change the playing song and make it your favorite then return to the main screen

All the tasks began on the main screen, and participants were instructed to start interaction after the moderator finished saying the task goal. Moreover, the participants were told they could take as much time as they needed to complete the task.

Additionally, the latter task was repeated only 2 times, while the first was done 3 times depending on the mode of interaction used (shown in figure 3). Participants interacted with the IVIS prototype either by touch (Touchscreen IVIS), or by hardware control knob attached vertically (parallel) under the IVIS (with IVIS screen further from the driver) and with the control knob attached horizontally next to the driver and with the IVIS screen further from the driver. Therefore, there were a total of five tasks to complete.

After the driving test concluded, the participants were asked to fill out the NASA task load index [8]. The whole testing of one participant on average lasted for about 30 minutes.

4.3 Usability metrics

To answer our proposed research questions, we propose a set of metrics that are calculated over only single participant:

- **Time away from the main task (TAFTMT)** - how long a participant didn't have their gaze on the road during a task given by equation:

$$T_m = T - \sum_{i=0}^{n} t_i \tag{1}$$

  where T is total time of the task, t_i is duration of i^{th} fixation in array $t = (t_0, t_1, t_2, ...t_n)$ which holds all fixations on road during a task and $t \subseteq t_a$, where t_a are all captured fixations during task.

- **Time on the IVIS screen (TOTIS)** - how long a participant looked at the mounted IVIS screen during tasks given by equation:

$$T_s = \sum_{i=0}^{n} t_i \tag{2}$$

  where T_s is time on the secondary task, t_i is the duration of i^{th} fixation in an array $t = (t_0, t_1, t_2, ...t_n)$ which holds durations of all fixations on IVIS during a task and $t \subseteq t_a$, where t_a are all captured fixations during the task.

- **Time on a task (TOT)** - how long a participant took to complete each task given by equation:

$$T_i = t_{Ei} - t_{Si} \tag{3}$$

  where i is task number and T_i is total time on the task, t_{Ei} and t_{Si}, denoting end and start timestamp of a task i.

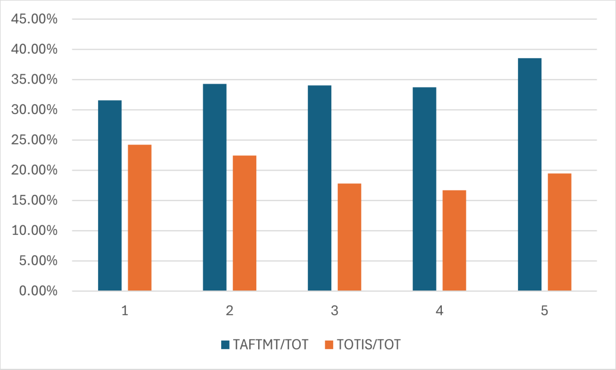


Figure 4: Average time away from the main task (blue bar) and average time on the IVIS screen (orange bar) against the total time in relation to the total time spent on a task given in percentage.

- **Time to first fixation on IVIS (TTFFOI)** - time to first fixation on the IVIS screen from the beginning of the task given by equation:

$$T_{fi} = t_{Fi} - t_{Si} \tag{4}$$

  where i denotes task number and t_{Si} denotes the timestamp of the beginning of the task and t_{Fi} denotes the timestamp of the first fixation on the IVIS screen during the task.

5 Results and Discussion

All of the participants successfully completed each task provided with each mode of interaction. Therefore, the metrics that defined the task completion criteria are: **Number of successful tasks** and **Task completion rate** are all rated at their maximum value of **5** and **100%** respectively. Because of this, the standard usability metrics are not enough to capture the complexity of the vehicle interactions.

Self-reported subjective measure of NASA task load index (shown in table 1). We only use the short version of the NASA task load index, which doesn't weigh the importance of each category, mainly because we find each category as relevant to the tasks at hand. We found that our participants reportedly found the tasks either cognitively demanding or non-demanding. The values for *Physical*, *Temporal* demand, *Performance*, *Effort* and *Frustration* fell mostly into Low (1-7) to Medium (8-14) cognitive load category with exception of one participant which reported High (15-21) cognitive load for most of the categories.

From our metrics (shown in table 2) we found that participants during all of the tasks spent approximately 20% of the time on the IVIS screen which was approx. 50% of the all of the time away from the main task within all

Table 1: NASA task load index result reported by the participants.

	Mental	**Physical**	**Temporal**	**Performance**	**Effort**	**Frustration**
Mean	9,8	7	9,4	8,2	11,4	9,6
STD	5,38	4,24	6,37	5,64	4,03	5,99

	Task 1	Task 2	Task 3	Task 4	Task 5
TAFTMT	5.25 (2.62)	8.61 (3.65)	7.83 (2.78)	6.02 (3.34)	9.02 (3.94)
TOTIS	4.03 (2.61)	5.62 (3.75)	4.09 (3.04)	2.97 (2.53)	4.55 (2.81)
TOT	16.65 (3.54)	25.10 (9.3)	23.02 (5.96)	17.84 (4.65)	23.39 (6.72)
TTFFOI	1.33 (1.06)	4.08 (5.29)	4.18 (5.03)	1.76 (1.61)	1.53 (1.26)

TAFTMT - Time away from the main task, **TOTIS** - Time on the IVIS screen, **TOT** - Time on a task, **TTFFOI** - Time to first fixation on IVIS.

Table 2: Calculated metrics from all the tasks performed by participants during simulator-based usability testing. Values are reported in second in format ”Mean (Std)”.

tasks. The total time on a task didn’t decrease meaningfully even when participants gained proficiency with the IVIS prototype. This can be explained by the mental workload required to learn how to use different ways of interacting with the IVIS. Many participants started to look at the IVIS screen as soon as the task was given often below the 2-second threshold (in 56% of all tasks).

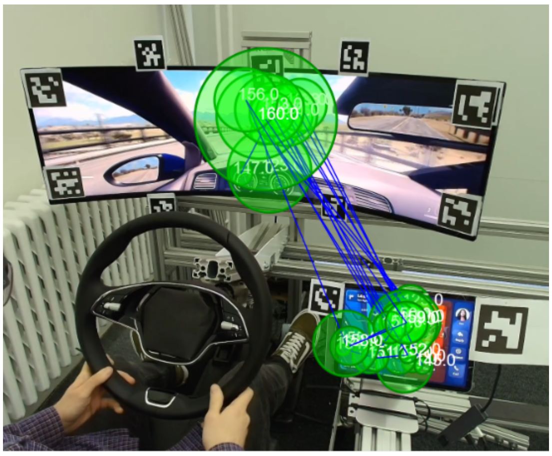
We also noticed that participants who used the hardware buttons parallel to the screen didn’t look at the IVIS screen (approx. 18%) as often as when it was touch-based (approx. 23%). But on the other hand, they spent around the same amount of time away from the screen (shown in Figure 4). This could be attributed to an error in remapping the views as well as the interaction style. During the visual analysis, we noticed that the participants didn’t often focus on the IVIS screen during interactions due to closeness to the main task screen, but had it in their peripheral vision.

Figure 5 shows a sample of fourth participant’s heatmap and scan path. We chose them as they represent the maximal time needed to complete the task. The fourth participant’s focus was mainly on the main task and focus on the IVIS screen was mainly on the control button’s locations and the middle of the screen where the system responded to the participant’s actions. From the scanpath, it is possible to see that the participant often switched from the main task screen to the IVIS screen. This participant had total of 33 fixations during the task from which 21 (63%) were placed on the IVIS screen.

Q1: It is possible to propose a measure of distraction for the effectiveness of the design of infotainment systems.

In our experiment, we used the same IVIS prototype

(a) Heatmap



(b) Scanpath

Figure 5: Task 1: Heatmap (a) and Scanpath (b) visualization of the fourth participant. On the heatmap, we can see attraction points (red denoting more attractive points); on the scanpath, it is possible to see the switching of participant focus between the main task and IVIS.

.

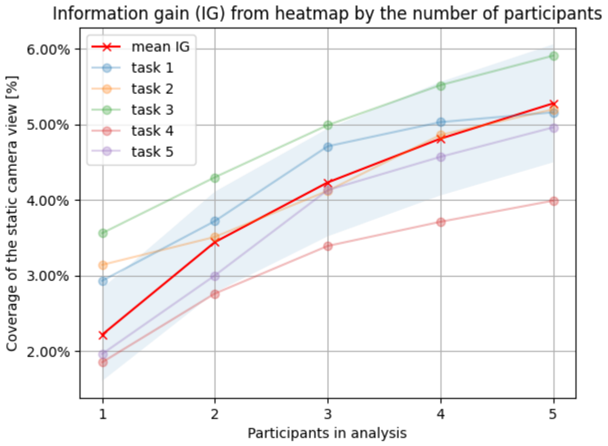


Figure 6: Information gain given by the number of participants. The average information gain in a task through all the permutations of participants' order is shown by lines from task 1 to task 5. Average information gain across tasks is shown by line mean IG.

with different modes of interaction with the system. From our results, we saw that moving the IVIS screen closer to the main screen task resulted in fewer fixations located on the IVIS screen. Also having the control knob parallel to the IVIS screen and visible in peripheral vision resulted in smaller time away from the main task as opposed to the knob being next to driver away from the peripheral vision.

Q2: Objective measures are more effective in the formalized evaluation of the usability testing of complex driving tasks in comparison with self-reported subjective measures.

From the calculated NASA Task Load Index scores (table 1), participants gave varying answers ranging from non-demanding to very demanding. Participants are often unable to report their own cognitive demands as the questionnaires assessing it are given afterward when they might misremember their experiences. From our result, it can be seen that both participants four and five reported high mental demand yet participant five was much quicker in most of the tasks. Participant Four also reported high temporal demand but took the longest time to complete the tasks.

Q3: The information gain from visual scene converges with an increasing amount of participants.

To answer this question, we created an aggregated heatmap based on fixation count visible from the static view camera (similar to figure 5b). And calculated the covered area by the heatmap. We also took all the permutations of participants' order in which they can be added to the analysis. The result of this analysis is shown in figure 6. It is possible to see that the IG starts to slow down at the 5-participants mark. Therefore, it is possible to use at most 5 participants to conduct such usability testing, but a few more participants can bring slightly more information gain.

6 Conclusions

This study has explored advanced usability testing methodologies for automotive user interfaces focusing on reducing driver distraction caused by in-vehicle infotainment systems. We developed a comprehensive framework for evaluating user interactions with in-vehicle infotainment systems by applying eye-tracking methodologies to objectively measure distraction. This framework combines objective measures, such as gaze metrics, time to complete a task, and task completion, with subjective evaluations, such as participant feedback in the form of cognitive load evaluation, to provide a holistic understanding of user experience.

The experiment in the study was aimed at validating the framework's viability to conduct and analyze usability testing of IVIS, and shows that the simulators are a viable alternative to field testing. Our framework has proven itself to be able to measure form of distraction caused by IVIS, and can be used for further studies of IVIS.

This research addresses key challenges in the usability testing of automotive systems, limitations brought by the scope of our experiment, which showed the need to evaluate more elements like the hardware button used in the experiment.

Future work should focus on advanced visualizations of gaze metrics and development of specialized tool for evaluation of the results obtained using the proposed framework.

References

[1] Stefan Becker, Parrish Hanna, and Verena Wagner. *Human Machine Interface Design in Modern Vehicles*, pages 1–16. John Wiley Sons, Ltd, 2014.

[2] Anna Bolder, Stefan M. Grünvogel, and Emanuel Angelescu. Comparison of the usability of a car infotainment system in a mixed reality environment and in a real car. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, VRST '18, New York, NY, USA, 2018. Association for Computing Machinery.

[3] William Kenneth Boyd. An analysis of user experience and distraction in automative an analysis of user experience and distraction in automative infotainment units infotainment units. Number 73, 2020.

[4] Adam Bykowski and Szymon Kupiński. Feature matching and aruco markers application in mobile eye tracking studies. In *2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, pages 255–260, 2018.

[5] Riccardo Crescenti, Piercarlo Dondi, Marco Porta, Cristiano Resta, and Gianfranco Rotondo. An eye tracking based evaluation protocol and method for

in-vehicle infotainment systems. In *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4, 2023.

[6] Tina Cvahte Ojstersek and Darja Topolsek. Eye tracking use in researching driver distraction: A scientometric and qualitative literature review approach. *J Eye Mov Res*, 12(3), September 2019.

[7] Yili Liu Fred Feng and Yifan Chen. Effects of quantity and size of buttons of in-vehicle touch screen on drivers' eye glance behavior. *International Journal of Human–Computer Interaction*, 34(12):1105–1118, 2018.

[8] S. G. Hart. Nasa task load index (tlx). volume 1.0; paper and pencil package. 1986.

[9] Michael Hildebrandt, Jens-Patrick Langstrand, and Hoa Thi Nguyen. Synopticon: A real-time data fusion platform for behavioral research. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings*, AutomotiveUI '19, page 494–497, New York, NY, USA, 2019. Association for Computing Machinery.

[10] Suhwan Jung, Jaehyun Park, Jungchul Park, Mungyeong Choe, Taehun Kim, Myungbin Choi, and Seunghwan Lee. Effect of touch button interface on in-vehicle information systems usability. *International Journal of Human–Computer Interaction*, 37(15):1404–1422, 2021.

[11] Tirza Jung, Christina Kaß, Dieter Zapf, and Heiko Hecht. Effectiveness and user acceptance of infotainment-lockouts: A driving simulator study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 60:643–656, 2019.

[12] Helene Høgh Larsen, Alexander Nuka Scheel, Toine Bogers, and Birger Larsen. Hands-free but not eyes-free: A usability evaluation of siri while driving. In *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, CHIIR '20, page 63–72, New York, NY, USA, 2020. Association for Computing Machinery.

[13] Nade Liang, Jing Yang, Denny Yu, Kwaku O. Prakah-Asante, Reates Curry, Mike Blommer, Radhakrishnan Swaminathan, and Brandon J. Pitts. Using eye-tracking to investigate the effects of pre-takeover visual engagement on situation awareness during automated driving. *Accident Analysis Prevention*, 157:106143, 2021.

[14] Benedikt Lux, Daniel Schmidl, Maximilian Eibl, Bastian Hinterleitner, Patricia Böhm, and Daniel Isemann. Efficiency and user experience of gaze interaction in an automotive environment. In Don Harris, editor, *Engineering Psychology and Cognitive Ergonomics*, pages 429–444, Cham, 2018. Springer International Publishing.

[15] Tyron L. Louw Mickaël J. R. Perrier and Oliver M. J. Carsten. Usability testing of three visual hmis for assisted driving: How design impacts driver distraction and mental models. *Ergonomics*, 66(8):1142–1163, 2023. PMID: 36259259.

[16] Viktor Nagy, Gábor Kovács, Péter Földesi, Dmytro Kurhan, Mykola Sysyn, Szabolcs Szalai, and Szabolcs Fischer. Testing road vehicle user interfaces concerning the driverrsquo;s cognitive load. *Infrastructures*, 8(3), 2023.

[17] Frederik Naujoks, Sebastian Hergeth, Katharina Wiedemann, Nadja Schoemig, Yannick Forster, and Andreas Keinath. Test procedure for evaluating the human-machine interface of vehicles with automated driving systems. *Traffic Injury Prevention*, 20:S146–S151, 08 2019.

[18] NHTSA. Distracted driving in 2021. *TRAFFIC SAFETY FACTS: Research Note*, 2023.

[19] Oskar Palinko, Andrew L. Kun, Alexander Shyrokov, and Peter Heeman. Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, ETRA '10, page 141–144, New York, NY, USA, 2010. Association for Computing Machinery.

[20] Gina Maria Schmidbauer-Wolf and Markus Guder. Usability and ux of a gaze interaction tool for front seat passengers: Evaluation of a gaze controlled optical feedback system in a car. In *Proceedings of Mensch Und Computer 2019*, MuC '19, page 677–681, New York, NY, USA, 2019. Association for Computing Machinery.

[21] David L. Strayer, Joel M. Cooper, Madeleine M. McCarty, Douglas J. Getty, Camille L. Wheatley, Conner J. Motzkus, Rachel M. Goethe, Francesco Biondi, and William J. Horrey. Visual and cognitive demands of carplay, android auto, and five native infotainment systems. *Human Factors*, 61(8):1371–1386, 2019. PMID: 30950645.

[22] Rao Tummala, Klaus-Juergen Wolter, Venky Sundaram, Vanessa Smet, and Pulugurtha M. Raj. New era in automotive electronics, a co-development by georgia tech and its automotive partners. In *2016 Pan Pacific Microelectronics Symposium (Pan Pacific)*, pages 1–4, 2016.

[23] Arun Ulahannan, Paul Jennings, Luis Oliveira, and Stewart Birrell. Designing an adaptive interface: Using eye tracking to classify how information usage changes over time in partially automated vehicles. *IEEE Access*, 8:16865–16875, 2020.