# On the track: Comparing Distraction Caused by Interaction with Tertiary Interfaces with Cars on a Test Track

David Wilfinger, Alexander Meschtscherjakov, Nicole Perterer and Manfred Tscheligi
Christian Doppler Laboratory "Contextual Interfaces" ICT&S Center, University of Salzburg
Austria
[david.wilfinger, alexander.meschtscherjakov, nicole.perterer, manfred.tscheligi] @sbg.ac.at

## ABSTRACT

To understand the qualities and downsides of current multi-functional in-car HMIs, we are in the process of developing a study approach that allows us to assess these systems holistically. One of the major focus points of the approach is the measurement of distraction caused by visual load. Trying out the approach, we conducted two studies described in this paper, comparing three in-car systems each. In the description of the studies we focus on the visual distraction as one part of our results, besides subjective workload measures. We found differences in terms of visual load between the systems, but even more between single tasks. Additionally, we found different patterns of distraction. Some systems were more distractive than others, but over a shorter period of time. Based on our experiences in the studies we raise open questions how to handle differences in visual distraction caused by the systems when interpreting the gathered data. The main questions are concerning what an acceptable amount of distraction is and how a balance between fast and highly distractive and slow but less distractive task conduction can be found.

## Categories and Subject Descriptors

H5.m. [**Information interfaces and Presentation (e.g., HCI**]: Miscellaneous

## Keywords

Automotive User Interface, Evaluation, Visual Distraction

## 1. INTRODUCTION

In the field of Automotive HMI development it is challenging to compare the quality of various automotive user interfaces in terms of safety, performance, usability and workload aspects. In-car systems vary in their functionality, they provide different interaction modalities and the aesthetic design can be perceived highly different by people with different tastes. Likewise the appearance can be different: in one car a huge screen dominate the design or in another car all attention is directed to a central rotary device. It is further immanent that driving situations and its contextual parameters (e.g., weather, type of road, daytime) in which a system is used can also be very different. Thus, a general research questions is: "How can we compare the distraction caused by different automotive user interfaces in current cars from various manufacturers?".

We aim at developing an approach that allows a comprehensive assessment of the cognitive load caused by automotive user interfaces and its distraction from the driving task. The intended method package is supposed to built on already established approaches and combines and extends the features of best practices with findings from our recent research in automotive interface evaluation. For that we are working towards a test procedure with defined and comparable tasks and exact methods of measurement regarding workload, distraction, driving performance and task completion time. The aim is to develop a test procedure, which is straight forward and easy to reproduce by the automotive industry and related organizations. Test procedures already exist for in vehicle tasks, such as the SAE J2365 [1], which focus primarily on the navigation task. We nevertheless believe that it is necessary to study complex modern tertiary systems in their entirety, since these systems allow so much more interaction than the conduction of navigation tasks.

It is our goal to provide an approach for investigating in-car HMIs holistically. For that purpose we developed a list of foci, based on a review of related literature, such as efficiency, effectiveness, accessibility, user experience, qualities under dynamic context conditions, and distraction, besides others. For that purpose our approach combines appropriate qualitatively and quantitatively methods to investigate the aspects of systems, one of which is the distraction it causes from the road.

This paper focusses on the distractive potential of HMIs through their visual load, which is definitely one of the most important aspects of a system, since it affects safety directly. Additionally, we describe the subjective workload measures we used and the results we gained from these measurements. This position paper describes the first steps in our iterative approach along with two example studies, in which we applied the approach in order to inform its further development. Finally, it raises open issues that we discovered concerning researching cognitive load and distraction, as important part of our approach. These issues are what


*AutomotiveUI'12*, October 17-19, Portsmouth, NH, USA.
Adjunct Proceedings

we want to raise within the workshop, to discuss them and to gather input from the community on how to potentially solve them, so that the development of our methodological framework can progress further. Although our work focusses mainly on visual load, we believe that the issues we raise also contribute to the cognitive load discussion. Especially the dispute between having a high load over a short period of time and a lower load over a longer period applies to both visual and cognitive effort.

## 2. APPROACH

In order to compare the distraction caused by different automotive user interfaces in current cars from various manufacturers, several challenges have to be addressed. In the following we will describe our approach by means of a user study:

In a first step we analyzed, which tasks should be given to participants in order to study tertiary in-car systems as holistically as possible. We identified four main groups of functions to be representative for each system, which therefore should be in the focus of our investigation: *navigation*, *entertainment*, *communication* and *configuration*. Since most current systems group functions accordingly, example tasks from each functional group were selected. These involve typing in a navigation destination, selecting a radio station, calling a number from the phonebook, and changing audio settings in the vehicle. The tasks were chosen since they represent a typical feature of each functional group and are therefore available in all modern in-car systems.

For investigating the distractive potential of each task, we ask participants to conduct them while driving on a test track. To reduce the effects of a first time usage, all tasks are trained once while the car is parked before they are conducted and recorded while driving. While this might not sufficiently reflect the experience users have with systems after using them for a longer period of time, our approach did not allow for a longer training phase.

For the driving task we propose the usage of a test track for several reasons: It increases safety and makes the conduction of an experiment easier, since all equipment can be stored at a central location. Additionally it allows a controlled experiment, with environmental conditions being relatively stable. The test track on the premises of our industrial partner represents a circle of 64 meters in diameter with an integrated part shaped like an "eight", which allows changing the direction of the curve that is driven (Figure 1).

**Figure 1: Test track used in both studies.**

Speeds up to 50 km/h are safely possible in trial conditions. Driving in a circle represents an easy and reproducible, nevertheless not undemanding driving task, which could be established as a standard for testing car interfaces while in motion.

To measure the distraction caused by the visual load affected by interacting with tertiary systems in the vehicles we use eye tracking technology, recording the eye movements between road and systems. Additionally, we used self reporting tools to give the participants the possibility to express their own perception after the trial.

## 3. STUDIES

The following paragraphs describe two studies we actually conducted with the above mentioned approach. The aim of the description is not to reveal all details of the results, but to focus on findings that are interesting to present to the community, not so much as finished results but in the form of challenges that we currently face. These findings form topics that we believe pose valuable issues for discussion in a workshop dealing with cognitive load.

### 3.1 Study 1

The goal of the first study was the comparison of three state of the art centralized in car systems on sale by German automobile manufacturers in May 2011 (BMW iDrive, Audi MMI, Mercedes COMAND). The three systems consisted of a central screen and a rotary knob with additional buttons to interact with the menu and functions shown on the display. The main differences of the systems are the menu structure and logic, the mapping of the rotary knob with menu functionality and the usage of context keys. For the study we invited 12 users, split evenly into three age groups of four people each (Group 1: 20y - 35y; group 2: 36y - 50y, group 3: 50y - 65y). Each of the groups consisted of two women and two men. We chose a within subject design, each user therefore operated each vehicle and system. Participants conducted tasks from the groups mentioned above, namely navigation, communication, entertainment, and settings.

Within our sample, we found high differences in the mean task duration, tasks conducted with the BMW system took on average 51.16s (SD: 29,1s), while the tasks with the Audi required 65,98s (SD: 39,7), resulting in tasks with the BMW only requiring 72% of the task duration required for the Audi system. Especially remarkable were the differences in the radio task (Audi: 43.22s, SD: 66s; BMW: 25.93s, SD: 25.4s) and in the phone task (Audi: 55.22s, SD: 28.4s; BMW: 35.14s, SD: 22.9s)).

Overall the eye tracking data showed a high level of distraction caused by all systems, with visual attention being directed on the system about 50 percent of the time (Audi: 49%, BMW: 54%, Mercedes: 53%). In combination with the task duration we computed the lowest eyes of the road time during the task conduction with the BMW system.

What left us with open questions was that we found that participants had the highest mean gaze duration on the display (Audi: 0.94s, SD: 0.34s; BMW: 0.99s, SD: 0.34s; Mercedes: 0.85s, SD: 0.23s) while conducting tasks with the BMW iDrive system. Nevertheless, the total eyes of the road time was the shortest with the BMW system due to its short overall task durations. The BMW system therefore was more distractive while the tasks were conducted, but less distractive in terms of overall eyes off the road time. Tasks were simply faster to conduct, which reduced the overall distraction.

We after conducting tasks with each system, we handed out the NASA RTLX questionnaire, which is based on the TLX scale by [3], to the participants. On the scale from 0 - 100 the systems got the following ratings: (Audi: 28.8, SD:17.8; BMW: 25.7, SD:13.4; Mercedes: 37.9, SD: 24.2). Similar to the eye tracking data, the TLX shows that the BMW system caused the lowest workload on the user side. What differs to the eye tracking interpretation is the fact that the overall task load with the Audi system was seen to be lower than with the Mercedes system.

## 3.2 Study 2

A second comparative study was conducted in spring of 2012. The study goal was again to compare interactive in-car systems and their effect on distraction. This time we compared three european compact class vehicles (Volkswagen Golf VI, Opel Astra, Ford Focus). Again we chose tasks out of the four functional groups of the systems (Navigation, Entertainment, Settings, Communication). Distraction was again measured with an eye tracking system. Different to study 1, we also included a mobile navigation system (TomTom) in each the three cars to compare it to the built in systems (see Figure 2). In study 2 we choose a between subject design, inviting 36 participants, 12 for each car. Participants were balanced in terms of gender and divided into three age groups (20-35, 36-50, 51-65 years old). The between subject design was chosen to minimize participant burden. In study 1 one test lasted up to four hours for each participant leading to exhaustion.

**Figure 2: Volkswagen Golf VI cockpit equipped with mobile navigation system and eye tracker.**

Summing up all tasks the Volkswagen system allowed the fastest average task completion time (VW: 44.31s, SD: 31.8s; Opel 46.02s, SD: 48.47; Ford 53.7s, SD: 52.6s). When analyzing the time, in which users on average focussed their visual attention on the displays, we found that although tasks could be conducted faster with the VW system they lead to more eyes of the road time than tasks with the Opel system (VW: 26.55s, Opel 22,95s, Ford 31.18s). In terms of eyes off the road, the VW system therefore distracted the users more, than the Opel system. Tasks with the Opel system took longer, but could be conducted with much shorter gazes (compared to the VW system). Therefore the distractive episodes were shorter and attention to the road could be paid more often and for a longer time. This was supported by shorter average duration per gaze with the Opel system compared to the others (VW: 0.81s, Opel: 0.71s, Ford: 0.87s).

It is important to note that the distraction caused by the systems strongly varied over all tasks. Thus, differences in the control design and philosophy become apparent (e.g., touchscreen vs. rotating knob). The VW system (touch screen) allowed a much faster conduction of the navigation task (i.e. type in letters), scrolling lists was nevertheless much easier with the Opel system (rotary knob).

We compared the built in systems of the Volkswagen, Opel, and Ford with each other and with a mobile TomTom navigation system to asses how distractive the mobile system was compared to the built in ones. Since the mobile system only supported the navigation task (type in a destination) we could only compare values form this tasks over the systems. We found that the navigation task was on average faster to conduct with the mobile navigation system (TomTom: 75,94s, Cars: 112,4s). This also lead to a lower duration of eyes of the road (TomTom: 43.15, VW: 56.61, Opel: 57.87, Ford: 77.74). Nevertheless, the average glance duration was longer than with the built in systems, resulting in a higher distraction during the phases of interaction for a task. The average glance duration during the navigation task with the Opel system, for example, was 0.77 seconds, with the mobile navigation system a glance on the screen took 1.16 seconds on average. The causes for this difference remain unclear, one cause could be the mounting position on the windshield, that required more effort in hand-eye coordination when reaching over for an input, but in the same moment allowed the users to keep the eyes on the system for longer, since the could see the road behind the navigation system.

In study 2 we asked the participants to rate their subjective experienced effort on a self assessment scale after they completed a task. The scale reaches from 0 to 220, participants tick a certain point on the scale based on their perceived effort, low ratings are representing a low effort. Comparing the three cars tested regarding the subjective experienced effort of conducting the different tasks, no significant difference could be found (VW: 38, Opel: 45, Ford: 44). Certainly, it can be stated that the navigational task with the Opel system (mean: 104) led to a marginally higher experienced effort in comparison with the VW system (mean: 56) and the Ford system (mean: 83). Regarding the other tasks, no further significant differences could be found regarding the subjective experienced effort of the three systems.

# 4. DISCUSSION

Both studies left us with the open question: “Is a higher level of visual load and distraction for a shorter moment more or less beneficial than a lower load over a longer period of time?” We are aware that there is not one right answer to that question and that there will be boundaries, in between one answer might be better than the other. But where are those boundaries? Is it the 2 second per glance rule or the rule that a task should not last longer than 25 seconds [2]? We perceive this question as highly relevant for our design activities in the automotive field, having to decide whether to design for a fast but potentially more demanding, or slow and less demanding conduction of tasks.

Our findings in study 2 show this conflict, resulting from different choices in input modalities. The touchscreen interface of the Volkswagen allowed a faster conduction of the tasks, but due to the necessary hand-eye coordination and the bad responsiveness of the screen, glances lasted longer and therefore the distraction in each moment of the task was higher then with the other systems. The same is true for the mobile navigation system, which allowed the tasks to be conducted faster, but was more visually demanding during the interaction. Based on our two studies we hypothesize that the touch screen interfaces we investigated were more efficient, but more demanding than the ones using a rotary knob or directional pads.

As stated by Green [2], rules like the 15-seconds-rule can be applied to tasks and systems, that have similar characteristics to the navigation task, that the rule was developed for. However, it can not be applied to any tasks. We therefore see a necessary discussion on what is acceptable in terms of workload and distraction, based on the kind of task. Differences in task concern the frequency of their conduction, the immediacy in which the task has to be conducted, and so on. Or is it, as Green suggests, the total eyes of the road time, which has to be taken into account. We nevertheless argue that a maximum of 10 seconds off the road time could also result in systems, that cause a constant distraction from the road over 10 seconds, without braking the rule.

We were also struggling with the fact, that the tasks for each car were very heterogenous in terms of distraction. Our data clearly showed, that no system was superior to the other throughout all tasks. One could argue that already a single task, that brakes established rules, makes a system unacceptable. This argument is supported by the finding that not only the average visual distraction should be considered but unusual exceptions in response to a traffic situation should be taken into account, since these are more likely to cause crashes [4]. Another opinion could be, that some functions, which were identified to be most distractive, should only be possible while the car is parked. But would that not lead to even more distraction, caused by user frustration and uncertainty, which functions are available in which situation? Other studies of ours have shown, that users would simply switch to using their smartphones, if the requested functionality is not available on the in-car system.

We also found that users experienced the workload caused by the tasks as less intense than the eye tracking data shows. In fact, all task load ratings are relatively low given the fact, that users were distracted from the road more than 50% of the task time. We can conclude that the perceived workload with the evaluated systems may be lower than the real workload, potentially leading to a dangerous gap between perceived and real risk of system usage.

Finally the question has to be raised whether the driving realism with the test track is sufficient for making statements about the distractive potential of in-car interfaces. Users were driving in a constant circle and could follow the track without much load caused by the driving task. No unexpected events were to occur, a safe driving style on the test track therefore could be maintained without a high amount of attention on the track. We therefore suspect that tasks were completed faster than under normal driving conditions. On the other hand our study ignored the concept of plastic time [5], meaning that interaction with in car systems can be interrupted and resumed, based on mini episodes during a trip (e.g., stopping at a red light). The way tasks were conducted in the experiment, therefore might not have represented the way users would conduct the in-car tasks in reality.

## 5. SUMMARY

In this paper we presented an approach we use to compare different in-car HMIs based on their usability, the user experience they cause and how they distract users from the main driving task. We propose the usage of example tasks from the areas of navigation, communication, entertainment, and configuration. Focussing on the distraction from the road, which we measured, we presented two studies in which we compared the multifunctional interfaces of three currently available cars each. We found differences between the vehicles, which allowed a statement about the distraction caused by each system, in comparison to the others. Nevertheless we were left with open questions. The main open question concerns the conflict between tasks, that are fast to conduct but more distractive and task that last longer, but distract less in every moment. Additionally, we were confronted with the question which level of distraction can be considered as acceptable and how to deal with systems, which provide functions that are beyond that acceptable range as well as functions, that follow established rules like the 15 second rule.

Since we are constantly refining our study approach as well as the interpretation of measured data, which is an important aspect of proposing a study concept, we would like to discuss these issues in the workshop with other members of the community. We believe that this would help us to direct our future efforts in understanding distraction caused by visual load in the vehicle.

## 6. ACKNOWLEDGMENTS

The financial support by the Federal Ministry of Economy, Family and Youth, the National Foundation for Research, Technology and Development and AUDIO MOBIL Elektronik GmbH is gratefully acknowledged (Christian Doppler Laboratory for “Contextual Interfaces”).

## 7. REFERENCES

[1] P. Green. Estimating compliance with the 15-second rule for driver-interface usability and safety. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 43, pages 987–991. SAGE Publications, 1999.

[2] P. Green. Potential expansion of the 15-second rule. In *NHTSA Driver Distraction Expert Working Group Meetings, Summary and Proceedings*, 2000.

[3] S. Hart and L. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Human mental workload*, 1:139–183, 1988.

[4] W. Horrey and C. Wickens. In-vehicle glance duration: Distributions, tails, and model of crash risk. *Transportation Research Record: Journal of the Transportation Research Board*, 2018(-1):22–28, 2007.

[5] T. Rattenbury, D. Nafus, and K. Anderson. Plastic: a metaphor for integrated technologies. In *Proceedings of the 10th international conference on Ubiquitous computing*, pages 232–241. ACM, 2008.