



# Efficiency and User Experience of Gaze Interaction in an Automotive Environment

Benedikt Lux<sup>1</sup>, Daniel Schmidl<sup>1</sup>, Maximilian Eibl<sup>1</sup>, Bastian Hinterleitner<sup>2</sup>,  
Patricia Böhm<sup>1</sup>, and Daniel Isemann<sup>1</sup>(✉)

<sup>1</sup> Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Germany  
daniel.isemann@ur.de

<sup>2</sup> Audi Electronics Venture GmbH, Sachsstraße 20, 85080 Gaimersheim, Germany

**Abstract.** Over the last years the number and quality of information and entertainment systems in automobiles has been rising constantly. This presents the challenge to provide safe and user-friendly interaction techniques, the implementation of which can lead to a higher level of efficiency, safety and user experience. One novel and promising approach is to use the drivers gaze as input for interaction with infotainment systems. We implemented a virtual car interior model to test the efficiency and user experience of gaze interaction with automotive infotainment systems. In a user study with 20 participants we compared a gaze-based interaction style to a haptic interaction technique. The usability of both techniques turned out to be very similar, while the user experience and the efficiency varied in parts. We used an eye-tracking device to investigate gaze behavior, but due to some technical problems with the device our quantitative findings are not as reliable and robust as we would have hoped for and have to be interpreted with care. Our qualitative data indicated a preference for gaze interaction.

**Keywords:** Gaze interaction · Eye-tracking · Haptic interaction  
In-vehicle information systems · User experience · Efficiency  
Automotive environment

## 1 Introduction

Due to technical developments and new user demands, in-vehicle information systems (IVIS) have become more versatile and advanced. IVIS update traffic information, control the driver's comfort in the car and display multimedia content. This development in the car interior goes hand in hand with the growth of complexity of these systems [19]. Research has shown that this rise of complexity leads to a bigger mental workload for the driver especially when using poorly designed IVIS [14]. Inappropriately designed IVIS can not only cause minor driving errors like losing track of the lanes [26] or slower reaction times while driving [5] but are also among the main reason for traffic accidents in the United States [17]. With 8.8% for all fatalities in car accidents in the US in 2015 distraction-affected fatalities were the main reason for casualties in traffic.

One way to prevent distraction related accidents is the deployment of additional security systems in the car. A different approach is to improve the interaction with the IVIS while driving and hence minimize driver distraction. Therefore, many different design strategies have been explored over the last years, including larger infotainment screens and additional displays. In order to minimize driver distraction while interacting with IVIS we developed a gaze interaction based prototype and compared this novel interaction style with haptic control in a user study.

The prototype comprises a virtual car interior model and four infotainment system components that users could interact with: a navigation system, a speedometer, a climate control system and a telephone interface. In a within-subject study users had to complete typical scenarios (e.g. dismiss a call) with both test conditions (haptic vs. gazed based control) and usability and workload were assessed.

## 2 Related Work

There are many studies on the use of eye tracking systems in cars, for example for interacting with a car's dashboard [18] and for the regulation which supports the driver's awareness by warning him according to his eye movements [12]. Much research focuses on the cognitive workload of the driver and his driving performance. Beyond that it is crucial to examine the driver's user experience of IVIS [4]. For these reasons our user study assesses performance metrics as well as subjective measures. We use the System Usability Scale (SUS) and the User Experience Questionnaire (UEQ) to deepen the understanding about the user's attitude and reaction towards our system.

The approach of interacting with systems based on gaze direction like a mouse pointer is known to be an error-prone solution because of the so-called "Midas Touch" problem [10]. Because of this eye-tracking interaction alone is often not regarded as a robust way to communicate with systems not least because of the lack of technical precision and unpractical use of the gaze for interacting with systems, such as IVIS.

Consequently, most gaze-based interaction systems are combined with other modalities like speech and haptic systems. While auditory displays have been found to have the lowest impact on workload and driving performance it compounds the task efficiency [23] and appears to be slower [11]. Although the combination of eye tracking with touch has shown to be the fastest way to complete car dashboard tasks it is not the most practical one with regard to driver's reaction times and the feeling of being in control of the steering wheel. One convenient way for a multimodal gaze direction-based IVIS is the combination with a steering wheel button system. Kern et al. [11] show the advantages of this combined approach: It is possible to interact with screens which are hard or unpractical to reach while having the hands on the steering wheel.

For the evaluation of these multimodal interaction components a fundamental restriction is the limit of the eye-tracking system which requires a minimum angle for accuracy [6]. For the scientific purpose eye-tacker SMI RED the angle is  $0.4^\circ$  which is only reached with good calibration. Furthermore, there is a need for additional camera systems for eye detection across a wide range of the IVIS. Dobbstein et al. have shown

that a full landscape eye-tracking system still presents serious issues because of technical problems with multiple eye-tracking systems [4].

## 3 Experiment

### 3.1 Prototype

To explore the potential of gaze interaction in an automotive environment we implemented a virtual car interior model and four infotainment system components that users could interact with: a navigation system, a speedometer, a climate control system and a telephone interface. The model with the different emulated screens was presented on a wide TV screen to our study participants. The prototype contained one display for each component with several UI elements. Furthermore a number of tasks typical for in-car interaction (e.g. decline an incoming call from a friend on the phone display) were implemented. To compare gaze interaction to haptic interaction the prototype allowed two interaction styles: In the “eye tracking” condition the users could switch the displays by looking at them. In the “haptic condition” four specific buttons on the steering wheel could be selected to activate each display.

### 3.2 Setup

The experiment took place in the Future Interaction Lab of the University of Regensburg. In our setup the test persons were sitting in front of a 48” TV set with a distance from the screen of around 70 cm. The TV screen had a Tobii eye tracking device mounted underneath it. Between the users and the TV screen a gaming steering wheel was placed (see Fig. 1). Only the wheel’s hardware buttons were used during the experiment because there was no need for the users to steer. The participants were asked to perform eight tasks overall, four in each condition. One condition was using the eye tracker to select the displays on the screen, the other one was to select them by pushing a button.

### 3.3 Experiment Design

As dependent variables we assessed the Task Completion Rates and Times as well as subjective measures. The independent variables were the gaze interaction and haptic interaction. For our experiment, we chose a within-subject design, meaning that each of the participants had to perform tasks in both conditions. Half of the participants started with condition one, the other half with condition two in order to balance out learning effects.

In the beginning the participants were welcomed and introduced to our team. After a short explanation of the experiment, the eye tracker was calibrated. Afterwards, the prototype environment was started and the participant was advised to select the emulated displays by gazing at them, to familiarize themselves with the system and also to make sure the eye-tracker was calibrated correctly. Then the test coordinator read the tasks and the participant started with the execution. The time on task was assessed for each



**Fig. 1.** Test setup in front of a 48" TV with a steering wheel and an eye tracker beneath the TV.

task. After all tasks in the first condition, test users filled out several questionnaires, namely the System Usability Scale [22], the User Experience Questionnaire [13] and the NASA Task Load Index (NASA-TLX) [8] and were interviewed to gather qualitative data. The procedure was repeated with the second condition. In the end users were given a demographic questionnaire.

We used a mixed-methods approach, collecting qualitative and quantitative data to get a deeper understanding of the new interaction technique. Besides the interview items from the questionnaires mentioned before and additional demographic questions, we also measured Task Completion Times (TCT) and Task Completion Rates (TCR). The quantitative data was used to identify significant differences regarding efficiency, effectiveness or satisfaction whereas the qualitative data gave insights on the subjective behavior and perceived problems by the user.

### 3.4 Tasks

The test tasks were designed as interrupted tasks in which the user had to switch between the different displays two to four times, as it may occur in a real life scenario. To achieve this, we implemented events triggered by the action of the user. For example, the user changed the temperature in the car which triggered an incoming call on another display, which had to be declined. By using these interrupted tasks, it was possible to make users switch between displays during the test.

As shown in Fig. 2 four displays were presented on the screen. As is common in modern cars, the cruise control was positioned above the steering wheel in the Head-Up-Display (HUD). The navigation system was placed behind the steering wheel. In the center console we positioned the infotainment menu and the climate control. If selected by the user pressing one of the shoulder keys of the wheel or using gaze direction, the



**Fig. 2.** Display setup in the prototype with speed control, navigation system, infotainment system and climate control. (Color figure online)

displays were highlighted in red for providing visual feedback. Once selected the user was able to navigate inside this display by pressing the up and down button on the wheel. By pressing the “X” button on the wheel the user was able to select a button on the display.

The tasks consisted of small actions like the user having to start the navigation to a friend’s home address, setting the air conditioning to a specific level, setting the speed of the car’s cruise control or declining an incoming call.

### 3.5 Participants

13 male and 7 female subjects participated in the test, resulting in 20 subjects overall. The average age was 28, most participants were students at the University of Regensburg. Only one person did not have a driver’s license, and 13 owned a car. 14 users indicated that their technical affinity is off a normal to high level.

## 4 Results

As mentioned above, a mixed methods approach was chosen in recording the results of our user study. After every one of the 20 subjects had completed the user test with each interaction concept, they had to fill out some questionnaires. For the collection of quantitative data the metrics of the SUS and UEQ were selected to measure the level of

usability and user experience and the NASA-TLX to give an indication of the cognitive workload while using one of the two interaction techniques. The times users needed to complete a single task were recorded (task completion times, TCT). In addition to the quantitative data we also collected some qualitative data. Especially with rather small samples or samples that do not fully represent the whole target group, as in our case, it has been recommended to explain quantitative results with qualitative observations in a mixed-methods approach [21]. This should lead to a data analysis with greater explanatory potential [3] and help us answer the research questions whether the gaze interaction technique is more efficient than the haptic one and whether it records higher levels of usability and user experience. To gather such qualitative data the participants were asked questions about their experiences during the user test in short interviews. Furthermore, we added questions to our demographic questionnaire that give insights into the driving experience and the use of technical interaction components in cars.

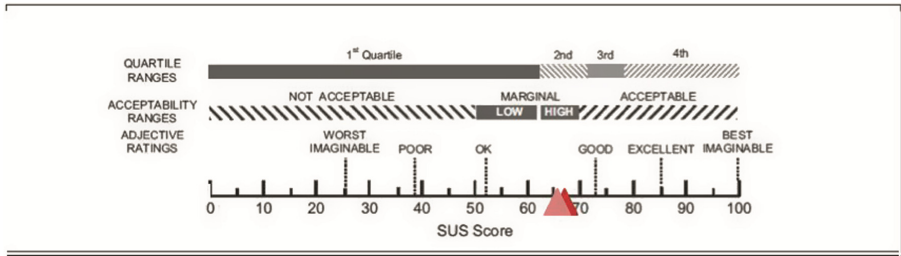
#### 4.1 System Usability Scale

The System Usability Scale (SUS) aims to identify users subjective rating of the usability of a tested system [2]. After rating the 10 oppositional word-pairs of the SUS on a Likert scale, the calculation of the SUS value leads to a number between 0 and 100 indicating the usability of a system. By admission of its inventors, the SUS represents a “quick and dirty solution” [2] which reveals only a rough tendency of the usability. With the SUS score of the haptic interaction ( $M = 67.000$ ,  $SD = 14.022$ ) and the score of the gaze interaction being not significantly different ( $M = 65.625$ ,  $SD = 14.478$ ), it is not possible to draw a comparative conclusion. To decide if these values are indicators for good or bad usability we compared them to SUS values from a larger range of tested systems, namely 2324 systems tested in a study by Bangor et al. [1] and 324 systems tested in a study by Sauro and Lewis [15] (Table 1). Compared to the Bangor et al. study the SUS values of our systems are below the mean and in the third highest quartile. Compared to the Sauro and Lewis study our SUS values are above both mean and median and in the second highest quartile. Hence, the score of the haptic and gaze interaction techniques are between the mean SUS Scores of those two studies. This may indicate an average level of usability.

**Table 1.** Basic information on SUS scores from the Bangor et al. and Sauro & Lewis studies.

	Bangor et al.	Sauro & Lewis
N	2324	324
Mean	70.1	62.1
1 <sup>st</sup> Quartile	55.0	45.0
Median	75.0	65.0
3 <sup>rd</sup> Quartile	87.5	75.0

A popular graphical outline of how to interpret the SUS score was created by Rauer [20]. The placement of the SUS results of our study according to this graphical interpretation scheme shows a similar assessment of our SUS scores (Fig. 3). However,



**Fig. 3.** Placement of the SUS scores of the gaze (light red) and haptic (dark red) interaction in Rauer's graphical interpretation scheme. (Color figure online)

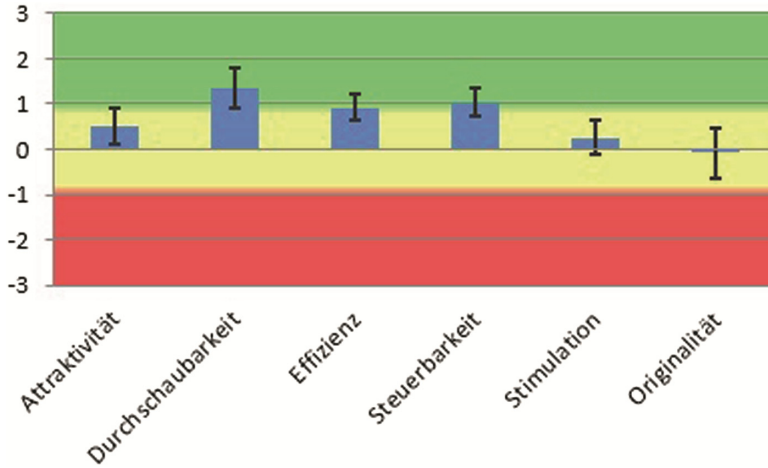
Rauer's interpretation scheme is not referencing scientific findings or literature and should therefore be viewed with caution.

## 4.2 User Experience Questionnaire

As the SUS is only a metric to indicate a tendency of the level of usability we also used the User Experience Questionnaire (UEQ). The UEQ collects a wide range of data and allows for deeper and more detailed interpretation of the results. It consists of 26 items from six different dimensions and contains not only indicators for usability, but also for users' feelings and emotions while interacting with the system. We used the UEQ to measure the user experience of our two interaction concepts. Participants had to fill out the questionnaire and the analysis of the resulting data was done with a spread sheet, which is available for download at the UEQ online documentation [9].

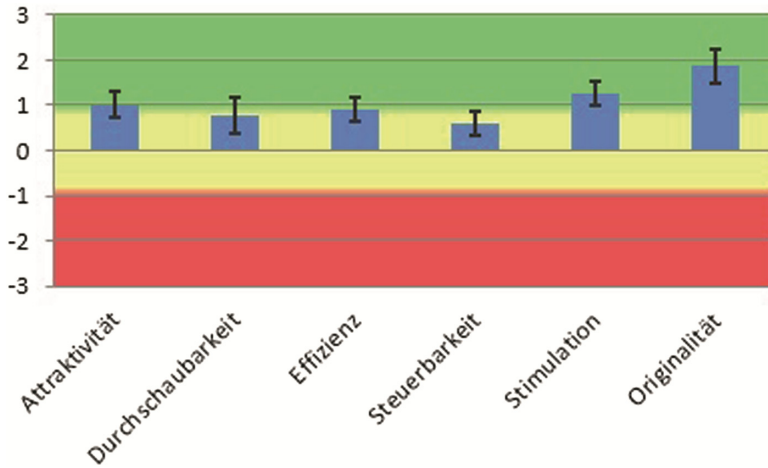
The six main dimensions of the UEQ are *attractiveness*, *perspicuity*, *efficiency*, *dependability*, *stimulation* and *novelty*. While perspicuity, efficiency and dependability are indicators for ergonomic quality, stimulation and novelty show the hedonic quality. Attractiveness displays the overall impression of the research object [13].

Figures 4 and 5 show the UEQ mean scores per dimension of both interaction techniques. The UEQ score for attractiveness of the gaze interaction ( $M = 1.02$ ,  $SD = 0.68$ ) shows a higher value than the haptic interaction ( $M = 0.52$ ,  $SD = 0.91$ ). But this difference is not statistically significant ( $t(20) = .056$ ,  $p < .05$ ). The dimension of perspicuity shows a higher score for the haptic interaction ( $M = 1.34$ ,  $SD = 1.01$ ) than for the gaze interaction ( $M = 0.76$ ,  $SD = 0.90$ ) but is also not significant ( $t(20) = .065$ ,  $p < .05$ ). Efficiency shows almost equal UEQ scores for haptic interaction ( $M = 0.93$ ,  $SD = 0.68$ ) and for gaze interaction ( $M = 0.91$ ,  $SD = 0.63$ ) and is therefore not significant either ( $t(20) = .953$ ,  $p < .05$ ). A significant difference is found in the UEQ dimension dependability ( $t(20) = .048$ ,  $p < .05$ ). Here the haptic interaction technique shows a higher score ( $M = 1.05$ ,  $SD = 0.74$ ) than the gaze interaction technique ( $M = 0.61$ ,  $SD = 0.61$ ). On the other hand the gaze condition achieves a significantly ( $t(20) < .001$ ,  $p < .05$ ) higher score ( $M = 1.26$ ,  $SD = 0.62$ ) than the haptic condition ( $M = 0.26$ ,  $SD = 0.90$ ) in the dimension of stimulation. The largest significant difference ( $t(20) < .001$ ,  $p < .05$ ) and hence the possibly greatest advantage for the gaze interaction is seen in the dimension



**Fig. 4.** UEQ scale of the six dimensions of the *haptic interaction* showing mean scores (blue bars) and confidence intervals (black lines). See text for English equivalents of the dimensions. (Color figure online)

of novelty. Here the score of the gaze condition ( $M = 1,86, SD = 0.9$ ) is much higher than the score of the haptic condition ( $M = -0.09, SD = 1.22$ ).



**Fig. 5.** UEQ scale of the six dimensions of the *gaze interaction* showing mean scores (blue bars) and confidence intervals (black lines). See text for English equivalents of the dimensions. (Color figure online)

To sum up the haptic interaction achieves better results in perspicuity (German: “Durchschaubarkeit”) and dependability (German: “Steuerbarkeit”), whereas the gaze interaction produces higher values for attractiveness (German: “Attraktivität”), stimulation (German: “Stimulation”) and novelty (German: “Originalität”). The dimension



of efficiency (German: “Effizienz”) is nearly equal for both interaction techniques. The only confidence interval which seems critical is the one of the novelty dimension of the haptic interaction. This value should not be over-interpreted because of possible misunderstandings with this item. Remarkable are especially the very high values for the originality of the gaze interaction and the perspicuity of the haptic one.

### 4.3 NASA – Task Load Index

To assess the workload experienced when using the two interaction techniques we used the well-established NASA-TLX. This task load index includes 6 items, measuring *mental*, *physical*, and *temporal demand*, the overall *effort*, *frustration level* and the subject’s satisfaction with their own *performance*. As Fig. 6 shows, the individual NASA-TLX dimensions for gaze and haptic interaction were comparable. To test for significant differences in the individual dimensions Wilcoxon tests for paired samples were used. The mean score of mental demand of the gaze interaction (8.35) was lower than the mean score of the haptic condition (9.1). The difference was not statistically significant ( $z = -.692$ ;  $p = 0.489$ ). The mean of the dimension physical demand for the gaze condition (6.7) was also lower compared to the haptic condition (7.5). The difference was not statistically significant ( $z = -1.195$ ;  $p = 0.232$ ). Regarding temporal demand the mean score for gaze (7.1) was higher than the one for haptic (6.4). The Wilcoxon test did not show a significant difference ( $z = -7.51$ ;  $p = 0.453$ ). The mean score for effort in gaze interaction (7.7) was very similar to the mean for haptic interaction (7.8). As expected the Wilcoxon test was not significant ( $z = -2.65$ ;  $p = 0.791$ ). Regarding performance the mean scores were relatively high with 14.1 for gaze and 13.1 for haptic interaction. The difference was not statistically significant ( $z = -1.136$ ;  $p = 0.256$ ). The mean frustration for gaze (6.75) was higher than the one for haptic

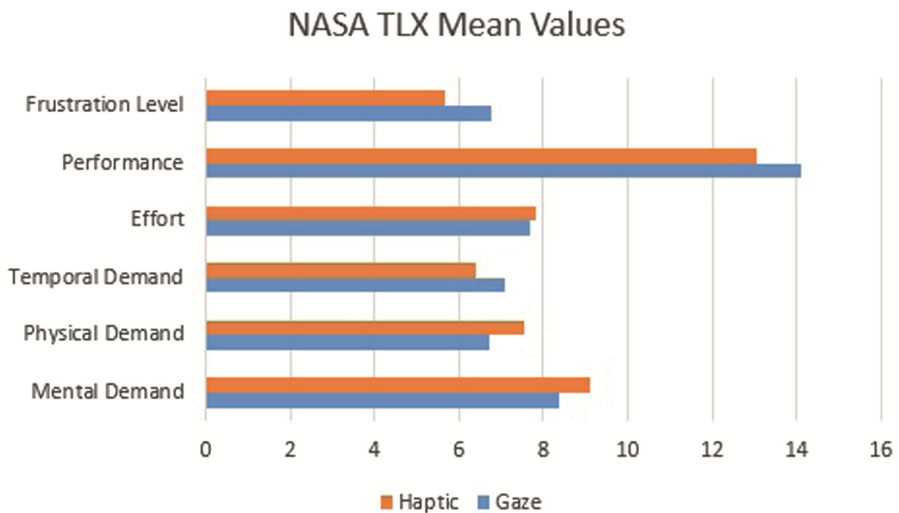


Fig. 6. The NASA-TLX item values show small differences.

interaction (5.6). The Wilcoxon test was not significant ( $z = -6.79$ ;  $p = 0.497$ ). As the mean scores of the dimensions were very similar and the Wilcoxon tests did not show any significant differences, the workload for gaze and haptic interaction would appear to be roughly equal.

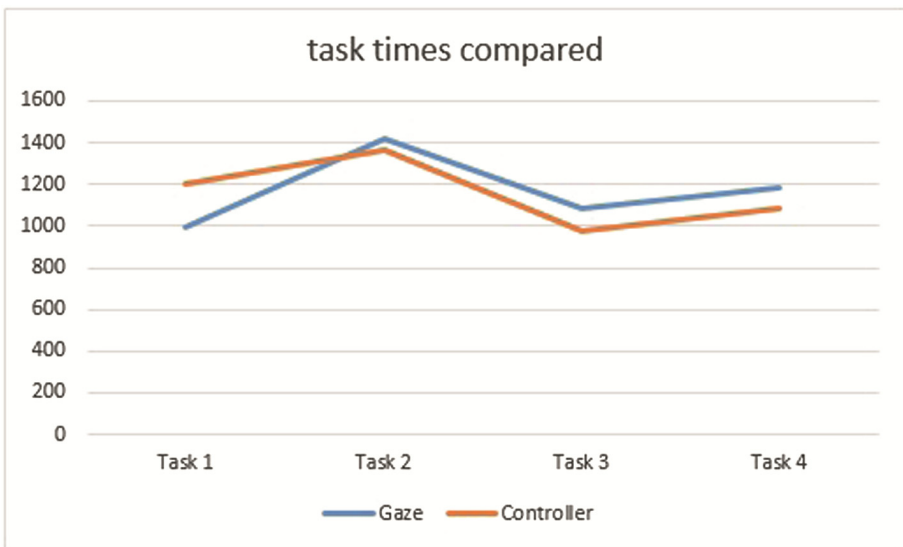
#### 4.4 Task Completion Rates

Every test person was able to complete every task using haptic and gaze input, so the task completion rate was 100% for both conditions.

#### 4.5 Task Completion Times

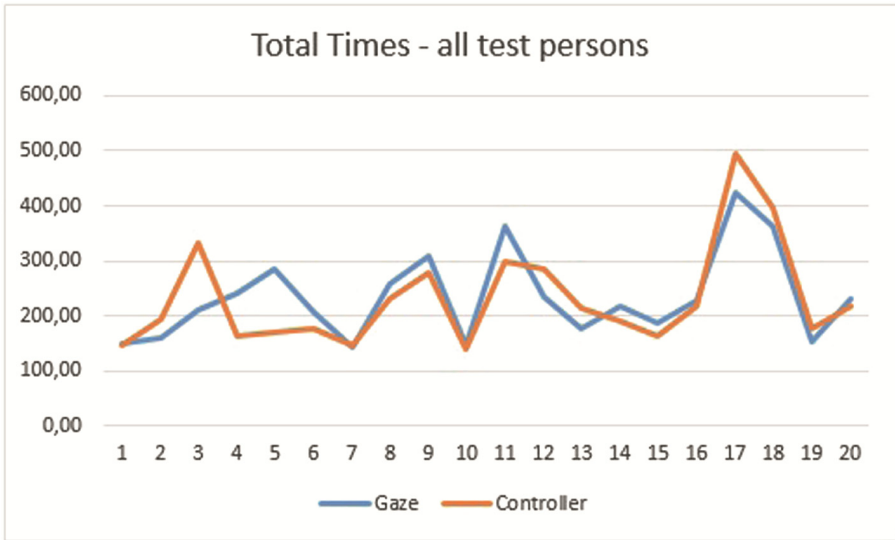
The time a person needs to complete a task is a very important part of an interaction technique in an automotive environment. The more time a driver has to spend on completing a task with an interaction system, the greater the distraction from his primary task namely driving safely and focusing on the traffic. So measuring the TCT of every Task with both interaction techniques was indispensable.

Figure 7 illustrates the total task completion times for both conditions and shows that except for the first task, the haptic interaction was faster than the gaze interaction. To examine whether these differences are statistically significant we analyzed the average completion times and used Wilcoxon tests to detect significant differences. The mean task time for task 1 with gaze interaction (49.5 s) was lower than with haptic interaction (60.1 s), but the difference was statistically not significant ( $z = -1,493$ ;



**Fig. 7.** The total of task completion times per task in seconds summed over all 20 participants. The comparison between haptic interaction (orange line) and gaze interaction (blue line) mostly shows shorter TCTs for the haptic interaction on average. (Color figure online)

$p = 0.135$ ). For task 2 the mean time on task for gaze (71 s) was higher than for haptic (68.2 s). The Wilcoxon test did not show a significant difference ( $z = -5.6$ ;  $p = 0.575$ ). The mean task time for task 3 was 53.3 s for gaze and lower for haptic with 48.7 s. The difference was statistically not significant ( $z = -1.8$ ;  $p = 0.07$ ). In Fig. 8 it can be seen, that the differences between conditions per subject are mostly small. Only for subjects 3, 4, 5, 11 and 17 there are more noticeable differences. It is perhaps noteworthy that the difference between the fastest and the slowest subject is nearly 350 s.



**Fig. 8.** The total completion time per person in seconds summed over all 4 tasks. The comparison between haptic interaction (orange line) and gaze interaction (blue line) shows mostly small differences between both interaction techniques. (Color figure online)

#### 4.6 Structured Interview

After collecting the quantitative data from all subjects under both research conditions the subjects were asked to answer some short questions about their experiences and impressions during the experiment. The goal of this short structured interview was to get answers that are succinct and easy to compare. The potential drawbacks which such an interview method entails, such as incentivizing certain answers, were kept in mind and mitigated as far as possible [16].

The interview comprised seven free-text questions:

1. Which interaction technique did you find more pleasing and why?
2. Could you briefly compare both techniques?
3. Did you find the separation of the four displays intuitive?
4. Would you rather prefer interacting with one display?
5. Can you imagine using gaze interaction in a car?

6. Which concerns do you have about gaze interaction?
7. Do you have other remarks to both techniques?

These questions were appended with two Likert scale questions, which inquired how intuitive switching between displays was with each technique:

8. How intuitive did you find switching between the displays with gaze interaction?
9. How intuitive did you find switching between the displays with haptic interaction?

The answers to these questions were collected and afterwards structured by a qualitative content analysis. The analyzed answers were assigned to the inductively developed categories which were counted based on how often they occurred.

### **Question 1**

14 of the 20 subjects answered that they would prefer the gaze interaction. 4 subjects would rather choose the haptic interaction. The remaining 2 gave no usable answer.

### **Question 2**

The answers to the second question revealed four perceived advantages of the gaze interaction: The technique was seen as innovative (mentioned 4 times), faster (2), less stressful (2) and more intuitive (1).

### **Question 3**

17 subjects found the separation and placement of the four displays intuitive. 2 had the opposite opinion and 1 did not give a clear answer.

### **Question 4**

Only 3 subjects would prefer to interact with a single display, 16 would not and 1 answer was not usable.

### **Question 5**

All suspects gave a clear answer if they could imagine gaze interaction in a car. 3 could not, 17 subjects said, they could.

### **Question 6**

The purpose of the final free-text question was to generate further thoughts or ideas by the subjects. A noteworthy remark was made by participant 10, who opined that a car manufacturer could outdo competitors by using gaze-interaction in a car: "You could surely score bonuspoints against your competition, if you used gaze interaction (Translated from German: "Mit einer Blickinteraktion könnte man sicher gegenüber den Wettbewerbskonkurrenten punkten"). Subject 18 who was completely against both systems and preferred haptic hardware buttons for discrete functions "I find it cumbersome to first select a display in order to then interact with something there. I would rather control things directly" (Translated from German: "Ich finde es umständlich erst ein Bildschirm auszuwählen, um dann dort etwas zu bedienen. Ich würde lieber die Dinge direkt steuern").

## 5 Interpretation

The SUS scores of both interaction techniques are around the mean value of the compared 2648 scores. They do not differ much from one another, so there is no clear difference between the two systems. The UEQ data by contrast reveals some differences. Taking a closer look at Figs. 4 and 5, it would appear that the two patterns are nearly axis-symmetric. The gaze interaction shows good results in the dimensions that the haptic interaction does not and vice versa. If one could improve the scores of the ergonomic dimensions of the gaze interaction, one could improve the whole UEQ result and it would probably also lead to a higher SUS score. The replies from the subjects' interviews about the concerns of the experienced gaze-interaction indicated that there were several problems with the Tobii EyeX eye tracking device during the experiment. Several times the tracked gaze was lost or too imprecise. This caused longer TCTs and with the TCTs being negatively correlated to the ergonomic UEQ items, the gaze interaction showed a bad performance [24]. This might also be a reason for worse UEQ and SUS results. Repeating the experiment with more robust eye tracking functionality would surely lead to better results for the gaze interaction.

This interpretation also explains why most of the subjects could imagine to use a gaze interaction system in a car, although the quantitative test results do not reveal a big advantage for this interaction technique.

Also the TCTs of Task 1, the only task where gaze interaction was faster than haptic interaction, support this theory, because it was the task with the shortest distance for the users gaze movements and therefore had the smallest error rate for the eye tracker.

A closer look at the dataset of subject 11 also underlines this theory. His individual SUS scores for gaze (SUS score: 50) and haptic interaction (SUS score: 75) showed the greatest difference among all subjects. He was one of the few people who rated gaze interaction badly in the UEQ dimensions efficiency, stimulation and novelty. His TCTs were all in all about 100 s slower with gaze interaction than with haptic interaction. To question b) from the structured interview he answered, that the eye tracking was very inconsistent and imprecise, so that he interacted with the wrong display regularly. The interaction with the steering wheel might have been much easier, because using the gaze interaction had been too inconsistent and imprecise. He had regularly interacted with the wrong widget. In spite of all this he could imagine to use the gaze interaction in a car, because it would be more exciting if it worked better.

## 6 Conclusion

In our user study we wanted to find the answer to two research questions. The first was if a gaze interaction technique in an automotive environment is more efficient than a haptic one. To answer this question one could point to the average or total TCTs and come to the conclusion that the gaze interaction technique is not more efficient. However a closer look at the times of each task shows, that for example the first task was completed faster using gaze interaction. Also 9 out of 20 subjects were faster with the eye-tracker. Especially subject 17 and 18, the two elderly subjects (>70 years), who

had to completely learn both techniques from scratch, had more problems with the haptic interaction and were faster with the gaze interaction. If we keep the problems with the eye-tracker in mind, the overall results could possibly have been better. Even the subjective efficiency of the suspects was not significantly different, as the UEQ results indicated.

The second research question was, if the eye-tracking system is more user-friendly. Therefore we wanted to estimate the usability and user experience of both interaction styles. The SUS indicated almost no difference in usability with the two scores being very similar (SUS of gaze interaction: 65.625, SUS of haptic interaction: 67). The results of the UEQ demonstrated that both techniques offer a completely different experience of interaction. Almost every dimension of the UEQ is good in one interaction technique and bad in the other one. Where the gaze interaction scores in the hedonic dimensions of attractiveness, stimulation and novelty, the haptic interaction points to the ergonomic dimensions perspicuity and dependability. Only the dimension of the efficiency is nearly equal for both techniques. The interview showed clearly that 14 out of the 18 usable answers of subjects to the question which system they would rather use, expressed preference for gaze interaction.

Overall our study cannot provide a completely accurate answer to our research questions. While our quantitative data shows a small advantage for the haptic interaction, the qualitative data points in the opposite direction and the hardware problems mentioned above may even have hampered a clearer endorsement of gaze interaction.

The variation in distances between different subjects and the eye tracker as well as the precision of the eye tracker itself caused problems with the gaze estimation. In addition, some subjects had more problems than others because of their height or their glancing behavior. It was obvious that corresponding to the two glancing stereotypes of Fridman et al. [7], the eye tracking device had more problems with so-called *owls* and less problems with so-called *lizards*.

Despite these restrictions we could identify some positive tendencies and aspects of gaze interaction. Therefore it may be worth trying to pursue the development of this interaction technique further. A key factor would be the use of a more stable and precise eye-tracker. There is a high chance that this would improve the results of future studies researching gaze interaction. The statements made in the interviews clearly showed that gaze interaction holds some potential as an interaction technique in future cars.

## References

1. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. *Int. J. Hum.-Comput. Interact.* **24**(6), 574–594 (2008)
2. Brooke, J.: SUS – a quick and dirty usability scale. *Usability Eval. Ind.* **189**(194), 4–7 (1996)
3. Creswell, J.W., Creswell, J.D.: *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. Sage Publications, Thousand Oaks (2017)
4. Dobbstein, D., Walch, M., Köll, A., Sahin, Ö., Hartmann, T., Rukzio, E.: Reducing in-vehicle interaction complexity: gaze-based mapping of a rotary knob to multiple interfaces. In: *Proceedings of the 15th International Conference of Mobile and Ubiquitous Multimedia*, pp. 311–313. ACM, December 2016

5. Donmez, B., Boyle, L.N., Lee, J.D.: Safety implications of providing real-time feedback to distracted drivers. *Accid. Anal. Prev.* **39**(3), 581–590 (2007)
6. Feit, A.M., Williams, S., Toledo, A., Paradiso, A., Kulkarni, H., Kane, S., Morris, M.R.: Toward everyday gaze input: accuracy and precision of eye tracking and implications for design. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1118–1130. ACM, May 2017
7. Fridman, L., Lee, J., Reimer, B., Victor, T.: ‘Owl’ and ‘Lizard’: patterns of head pose and eye pose in driver gaze classification. *IET Comput. Vis.* **10**(4), 308–314 (2016)
8. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: *Advances in Psychology*, North-Holland, vol. 52, pp. 139–183 (1988)
9. Hinderks, A.: UEQ Online: Downloads. <http://www.ueq-online.org/index.php/user-experience-questionnaire-download/?lang=de>. Accessed 17 Jan 2015
10. Jacob, R.J.: What you look at is what you get: eye movement-based interaction techniques. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 11–18. ACM, March 1990
11. Kern, D., Mahr, A., Castronovo, S., Schmidt, A., Müller, C.: Making use of drivers’ glances onto the screen for explicit gaze-based interaction. In: *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 110–116. ACM, November 2010
12. Langner, T., Seifert, D., Fischer, B., Goehring, D., Ganjineh, T., Rojas, R.: Traffic awareness driver assistance based on stereovision, eye-tracking, and head-up display. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3167–3173. IEEE, May 2016
13. Laugwitz, B., Held, T., Schrepp, M.: Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) *USAB 2008*. LNCS, vol. 5298, pp. 63–76. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-89350-9\\_6](https://doi.org/10.1007/978-3-540-89350-9_6)
14. Lee, J.D., Caven, B., Haake, S., Brown, T.L.: Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers’ attention to the roadway. *Hum. Factors* **43**(4), 631–640 (2001)
15. Lewis, J.R., Sauro, J.: The factor structure of the system usability scale. In: Kurosu, M. (ed.) *HCD 2009*. LNCS, vol. 5619, pp. 94–103. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-02806-9\\_12](https://doi.org/10.1007/978-3-642-02806-9_12)
16. Mayring, P.: Qualitative inhaltsanalyse. In: Mey, G., Mruck, K. (eds.) *Handbuch Qualitative Forschung in der Psychologie*, pp. 601–613. VS Verlag für Sozialwissenschaften, Wiesbaden (2010). [https://doi.org/10.1007/978-3-531-92052-8\\_42](https://doi.org/10.1007/978-3-531-92052-8_42)
17. National Highway Traffic Safety Administration: 2015 motor vehicle crashes: overview. *Traffic safety facts research note*, pp. 1–9 (2016)
18. Poitschke, T., Laquai, F., Stamboliev, S., Rigoll, G.: Gaze-based interaction on multiple displays in an automotive environment. In: *2011 IEEE International Conference Systems, Man, and Cybernetics (SMC)*, pp. 543–548. IEEE, October 2011
19. Purucker, C., Naujoks, F., Prill, A., Neukum, A.: Evaluating distraction of in-vehicle information systems while driving by predicting total eyes-off-road times with keystroke level modeling. *Appl. Ergon.* **58**, 543–554 (2017)
20. Rauer, M.: Quantitative Usability-Analysen mit der System Usability Scale (SUS). *Nachrichten, Tipps & Anleitungen für Agile, Entwicklung, Atlassian Software (JIRA, Confluence, Stash, ...)*. Seibert Media (2011). <https://blog.seibert-media.net/blog/2011/04/11/usability-analysen-system-usability-scale-sus/>. Accessed 30 Mar 2017

21. Ritchie, J., Lewis, J., Nicholls, C.M., Ormston, R. (eds.): *Qualitative Research Practice: A Guide for Social Science Students and Researchers*. Sage, Thousand Oaks (2013)
22. U.S. Department of Health & Human Services: System Usability Scale (SUS) (2013). <http://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>. Accessed 8 May 2015
23. Weinberg, G., Harsham, B., Medenica, Z.: Evaluating the usability of a head-up display for selection from choice lists in cars. In: *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 39–46. ACM, November 2011
24. Zwahlen, H.T., Adams, C.C., DeBals, D.P.: Safety aspects of CRT touch panel controls in automobiles. In: *Vision in Vehicles II. Second International Conference on Vision in Vehicles*, pp. 335–344 (1988)