Traffic Light Assistant – Can take my eyes off of you

Michael Krause, Verena Knott, & Klaus Bengler
Institute of Ergonomics, Technische Universität München
Germany

Abstract

A traffic light assistant on a smart phone is assessed with an eye tracker in real traffic. The system is displayed on three different screen sizes on a nomadic device and gaze durations are measured. Another condition of the experiment includes an acoustic click when the display content changes to reduce glance frequency. The acoustic click as an auditory hint does not reduce the frequency of gazes, as expected. The gaze durations can get shorter as display size increases, but this does not necessarily reduce the percentage of time an in-vehicle information system (IVIS) is looked at. Subjective ratings indicate that display contents can even be shown too big. Overall, the gaze durations are in line with current limits, even when displayed on a small screen. A set of gaze histograms and calculated gaze metrics is provided to enable comparison with other experiments and IVIS.

Introduction

Stops at traffic lights and related braking and acceleration are potential aims for optimization. In the Bavarian pilot project KOLIBRI, which stands for Kooperative Lichtsignaloptimierung – Bayerisches Pilotprojekt [Cooperative Traffic Light Optimization – Bavarian Pilot Project], four partners evaluated the feasibility and chances of traffic light optimization and driver information on arterial roads. The project was managed by TRANSVER GmbH, a traffic engineering company. The Board of Building and Public Works in the Bavarian Ministry of the Interior was responsible for the two rural roads, which served as test tracks. BMW AG built up a demonstration car. The Institute of Ergonomics at the Technische Universität München handled the human factors and ergonomics issues within the project, especially the information for the driver about upcoming traffic signal states with smart phones.

The final information system is based on already installed GSM and UMTS networks. The traffic lights on the two test tracks - one in northern Munich with seven traffic light controlled intersections on seven kilometres, the other one near Regensburg with eight traffic light controlled intersections on a length of five kilometres - were equipped with radio data modems. A central server receives the process data from the traffic lights and tries to estimate upon historical data their behaviour within the next minutes. This probability forecast is send to the demonstration car or smart phones. These devices calculate a speed

recommendation, or give other appropriate information about the traffic light phases to a driver. Due to the in-vehicle use, special care must be taken for suitability while driving (ESoP, 2006).

**Related work – Traffic Light Assistance**

Early contribution to the field of traffic lights assistance and driver information can be found in the *Wolfsburger Welle* project from VW AG (Hoffmann, 1991; Voy et al., 1981). About the same time, Australian traffic engineers experimented with driver information and traffic lights, especially with dynamic speed recommendations signs along the road (Trayford et al., 1984; Bell 1984; van Leersum, 1985; Trayford & Crowle, 1989). In the CVIS project, a green wave in-vehicle information system was compared to dynamic signs along the road, in a simulator study (Duivenvoorden, 2007; Pauwelussen et al., 2008).

Thoma (2010) compared different human machine interfaces (HMI) for a traffic light information on-board system. The idea, to feed traffic light information into the adaptive cruise control, can also be found in Thoma (2010) and in the KOLINE project (Bley et al., 2011). Chen et al. (2009) gave a review on signal countdowns and showed the risks of using traffic light signal countdowns. Ferreira et al. (2010) dealt with the idea of virtual traffic lights (VTL) in the car, to replace the roadside infrastructure. Olaverri et al. (2012) evaluated HMIs on the head up display for the VTL in a simulator. The TRAVOLUTION project from AUDI AG (Braun et al., 2009) used WiFi to bring the traffic phase information to the car. This approach was also used in the German *aktiv* project (Hoyer, 2012), which used a mobile device to inform the driver. Data for WiFi connection ranges under different vehicle speeds can be found in (Iglesias et al., 2008). A traffic phase information system was also included in the German simTD project (2008). The SignalGuru project (Koukoumidis et al., 2011) used smart phones and image processing to estimate traffic light phases.

KOLIBRI differs from other projects in its heavy reliance on already installed mobile communication networks, instead of dedicated short range communications, with their coverage drawbacks and maintenance costs. The information presentation on the wide spread smart phones makes it possible to equip every car, regardless of make and year. The most important distinction is the user- and usability-centered approach. Its use while driving needs special attention. Many of the preceding or running projects showed a lack of consideration for user needs, thus focusing on technical feasibility only.

This paper will focus on a sub-aspect of “suitability while driving”: The gaze behaviour.

**Related work – Gaze behaviour**

Gaze behaviour in transportation is a long-researched topic. All major guidelines for in-vehicle information systems (IVIS) (i.e. JAMA, 2004; AAM, 2006; ESoP, 2006) address the visual demand with some gaze metrics (occlusion metrics or glance durations). For instance, the AAM states: “[…] single glance durations generally
should not exceed 2 seconds”. And the ISO 15005:2002, which is referenced in the ESoP, introduces a maximum dwell time of 1.5s to capture information.

There is also a research tradition on the determination of optimal font size (Beymer et al., 2008) and even the interaction between font size, display size and remembering was assessed (Sanchez & Goolsbee, 2010). Tullis (1984) used metrics like grouping and density, which are indirectly connected to display size, to predict search times and subjective ratings on alphanumeric displays. The well-known model SEEV (Wickens et al., 2001) has a special focus on visual attention allocation. The Distract-R approach (based on ACT-R) currently tries to expand this field with salience features (Lee et al., 2012). The advantages of an improved SEEV, called NSEEV, can be found by Steelman et al. (2011). NSEEV especially uses extended salience features and can be used for more than just steady-state distributions.

In the experiment presented here two topics related to visual attention are addressed:

- Given that smartphone screen sizes on the market are heterogeneous, how does the display size of presented graphical information on small devices affect gaze duration?
- Can an auditory hint (a click for new information on the display) reduce glances to the designed traffic light assistance IVIS and thus glance frequency?

Accordingly, it is possible to state the hypothesis that gaze duration should get shorter when the display size of small devices increases because the graphical content should be easier to perceive. On the other hand, this should only be true on a small scale, since the eye has only a narrow spot (fovea) for sharp seeing. So, if the display gets larger, one has to look around on the screen. This has been addressed by many usability studies (e.g. for websites, Nielsen 2006).

The discussion of the results of this experiment will try to connect the outcomes to the SEEV model. “SEEV derives its name from the four forms of attentional influence that it posits: channel salience, the effort needed to move attention between channels, the operator’s expectancy of the signal on each channel, and the task criticality or value of the information in each channel” (Steelman et al., 2011).
Method

Test environment

Test track
The individual test runs of the experiment were performed on federal road 13 (B13) in northern Munich. Over a length of about seven kilometres, the test track has seven traffic-light-signal-controlled intersections. Moreover, it is a section of the B13 with two lanes in each direction, and in addition, there are turning lanes in front of the individual traffic lights. A guardrail separates the carriageways structurally from each other. Under dry road conditions, the speed limits on the test track are mostly limited to 100 km/h, and reduced to 70 km/h in front of every intersection. In the various test runs used in the experiment, this route was driven either from the south (48° 14’ 48 N 11° 36’ 7 E) to the north (48° 17’ 56 N 11° 34’ 36 E), or in the opposite direction. At both ends of the test track, two turning points were used during the experiment to evaluate the performed test run subjectively by questionnaires and to explain the object of the following test run. The experiments avoided the rush hours, which have highly directional traffic load. Thus, the traffic density was about 500 cars/h, for each direction. In the experiment reported here, the traffic lights were acting on a coordinated scheme (green wave). The predetermined switching times are programmed into a nomadic device to inform the driver.

Test vehicle
For the study on the real test field, a BMW X5 (modified US model) was used as vehicle. In the experiment, a Samsung Galaxy Tab GT-P1000 (display: 7.0 inch, 1024 x 600, Android-Firmware-Version 2.3.6) was used to display the mobile KOLIBRI-traffic light assistant. The nomadic device was fixed to the right of the driver with a mount in the ventilation slots of the centre console of the test vehicle. The automated analysis of gaze data, which were generated by the Dikablis Eye-Tracking System (Ergoneers GmbH, Manching), also required the attachment of markers on the dashboard of the test vehicle (see Figure 1).

Figure 1. Attachment of the nomadic device and the markers in the test vehicle
Human Machine Interface (HMI)

After initial experiments regarding various HMIs in the static driving simulator (Krause & Bengler, 2012a – 2012b), the optimized concept in Figure 2 is used for the experiment in real traffic.

![Figure 2. Display states of the KOLIBRI-traffic light assistant](image)

The favoured display concept of the KOLIBRI-traffic light assistant on the nomadic device, which was used in this study, includes a speed recommendation in form of a speed carpet. The green area represents the speed recommendation that supports the driving behaviour for a green wave (Figure 2-1 and Figure 2-2). The illustrated vehicle position corresponds to the current speed. The white vehicle indicates that the driver is within this recommendation while driving at current speed. The black car of Figure 2-2 points to the contrary. The position of the pointer of the “Heuer” traffic light assistant in the right upper corner of the application shows the current cycle state - green or red light - of the next traffic light. The speeding display appears when the speed limit is violated (> 10 km/h), see Figure 2-3. If the vehicle is outside the criteria for calculating a speed recommendation (i.e. if the driver has to move too slowly or too fast to reach the green wave), Figure 2-4, called “Vorbereiten auf Halt” (“Prepare to Stop”), is displayed. It shows the driver that the next traffic light will be red at arrival. The combination of the Heuer traffic light and the countdown appears at a speed less than 5 km/h in front of a red light signal system (Figure 2-5). This function displays the waiting time, until the traffic light changes to green again.

Independent variables: variation of size and click

This experiment used variations of the KOLIBRI-traffic light assistant in the display size. Furthermore, a variant was used with an acoustic signal. Figure 3 illustrates the relative proportions of the screen sizes of the display concepts of the assistance system in the individual test runs. In the previously conducted studies in the static driving simulator, the traffic light assistant was presented on a smartphone (Samsung Galaxy Ace S5830) with a 3.5-inch diagonal (Krause & Bengler, 2012a – 2012b). So the experiment could be compared in real traffic with the appropriate simulator study, one display of the variations on the tablet should correspond to the smartphone display in size. The display concept HMI_M has the characteristic of a 3.5-inch diagonal. The large display HMI_L was adapted in such a way that the screen of the tablet is utilized optimally (width of 900px). Derived from these
dimensions generated for the HMI_L display, the exact sizes of the other KOLIBRI-displays result from the specified scaling factors.

<table>
<thead>
<tr>
<th>Display Size</th>
<th>~ 2,5&quot;</th>
<th>~ 3,5&quot;</th>
<th>~ 6,4&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic Click</td>
<td>HMI_S</td>
<td>HMI_M</td>
<td>HMI_L</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>HMI_M_C</td>
<td>-</td>
</tr>
</tbody>
</table>

1) Scaling factor of the edge lengths
2) Resolution: 170 dpi

Figure 3. Relative proportions of the screen sizes

The smallest display concept HMI_S corresponds to half of the display surface of the HMI_M-system. The fourth variant of KOLIBRI-assistance, indicating a change between different states by an acoustic click (HMI_M_C), is identical in size to HMI_M-display concept. The HMI prototype was implemented in Adobe AIR. The size scaling was performed inside AIR (vector graphics). There was nothing like a border around the resized graphic. The resized graphics were just centred on the black screen. The short click was played when the display changed between one of the states of Figure 2-1 to Figure 2-5 (for instance, if the vehicle travelled too fast and Figure 2-3 appeared or when the car drove out of or into the speed recommendation, e.g. from Figure 2-1 to Figure 2-2). The volume of the tab was set to maximum. The display content was recalculated for every new GPS value, typically at 1Hz.

Gaze tracking

Within the scope of this study, the Dikablis Eye-Tracking System (Ergoneers GmbH, Manching) was used for experimental recording of the gaze data and D-Lab 2.1 for analysis. Three areas of interest were defined (Figure 4): “windscreen” (or “front”), “speedometer” and “smartphone” (or “tablet”). The “smartphone” area had a fixed size and was not adjusted when display content was scaled.
Figure 4. Areas of Interest (AOI)

**Questionnaires**

In addition to the questionnaire for collecting demographic information, the following standardized written survey instruments were used in this study:

**NASA-Task Load Index**

The method for collecting of ratings of the workload, called NASA-TLX, was developed by the Human Performance Group at NASA Ames Research Center (Hart, 1986; Hart & Staveland 1988). In this study, we used a German translation (Seifert, 2002). An Overall Workload Index (OWI), based on a non-weighted average of the ratings of the subscales, was used for comparison (Hart, 1986).

**System Usability Scale**

The System Usability Scale (SUS) represents a reliable and cost-effective questionnaire with ten items (each on a 5-point-Likert-scale) for a global rating of the usability of interactive systems (Brooke, 1996). In the experiment, the SUS was used after each full run combined with the KOLIBRI-traffic light assistance. A careful German translation was used, combined with the English original text. There is some effort in Germany (Reinhardt, 2012) to get a uniform and verified translation. The translation used here was involved in different experiments at our institute, so we published the translated items and condensed SUS results from these studies (Krause & Bengler, 2013) to make comparisons possible.

**Test procedure**

After the test subjects were welcomed and informed about the KOLIBRI project, they indicated their consent to participate in the study. In addition, a demographic questionnaire collected personal data. Afterwards, the explanation of the maximum 120 minutes ongoing trial followed. After successful calibration of the eye tracking system and vehicle-specific instructions, the explanation of the KOLIBRI-traffic light assistant with its functions followed.
After these preparations, the acclimatisation test run took place using the KOLIBRI-traffic light assistance, which included the acoustic click. In addition, five test runs in randomised order were realised in the BMW X5 on the test track B13. A baseline run (BL) without the use of a traffic light assistant was compared with four other runs combined with various KOLIBRI systems that differed in their display sizes (HMI_S, HMI_M, HMI_L). Moreover, in another test run (HMI_M_C), the KOLIBRI assistant with a 3.5-inch diagonal and an acoustic click was tested on the test track. The signal happens at each change in display to illustrate a display change to the driver.

The individual test runs were randomly done in one direction (south-north or north-south). To assess the subjective workload during the test runs, both the NASA-TLX and the SUS questionnaire (except for the baseline run) had to be completed by the test subject after finishing the run.

Participants

The sample of this study included 22 test persons (5 females; mean age = 28.6 years; st. dev. = 8.4). All test subjects had a valid driving license. No test person of the study had a colour deficiency or important eye disease for the test procedure. However, 32% of the subjects reported the use of a visual aid. Nine percent of the participants were not familiar with driving an automatic car and two persons of the KOLIBRI study did not know the real test track in northern Munich. With a proportion of 45 percent, almost half of all participants of the experiment had already participated in a KOLIBRI study in the driving simulator. 63 percent of the subjects had participated in a previous study that was carried out with the use of the traffic light assistant in real traffic. In this study, reaction times were collected by handling a tactile detection task.

Results and discussion

Glance duration and display size

The traffic lights assistant reported the current display state to the eye tracking system. So it is possible to split the gaze analysis up. An average trial run in the “HMI_L”, “HMI_M” and “HMI_S” conditions lasted 417 seconds (SD: 40s). From this time, only sections were analysed when the vehicle was moving. This was achieved by considering the display states “speed recommendation” (see. Figure 2-1 and Figure 2-2) which was, on average, active for a total time (summed up segments) of 200 seconds (SD: 53s) during a test run. And “prepare to stop” (Figure 2-4), which was, on average, active for a total time of 124 seconds (SD: 66s). Figure 5 shows the results.

A single factor (display size), repeated measurement ANOVA for “speed recommendation” reported a significant difference F(2;42) = 13.4, p < 0.001, whereas the same test for the state “prepare to stop” was not significant F(2;42) = 0.646 p = 0.529. A Bonferroni corrected post-hoc test revealed the data in Table 7.
can take my eyes off of you

Figure 5. Mean glance durations with different display sizes

Table 7. Bonferroni corrected post hoc test for glance durations in display state “speed recommendation” (*p < 0.05)

<table>
<thead>
<tr>
<th>Display Size Comparison</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMI_S – HMI_M</td>
<td>0.004*</td>
</tr>
<tr>
<td>HMI_S – HMI_L</td>
<td>&lt; 0.001*</td>
</tr>
<tr>
<td>HMI_M – HMI_L</td>
<td>0.923</td>
</tr>
</tbody>
</table>

For the “speed recommendation” display state, the same ANOVA (display size) was carried out on glance frequency and glance percentage of time. Both were not significant (average glance percentage of time to the phone during “speed recommendation” was HMI_S: 12.1%; HMI_M: 10.2%; HMI_L: 11.2%; average glance frequency during “speed recommendation” was HMI_S: 0.174Hz; HMI_M: 0.163Hz; HMI_L: 0.186Hz). It seems like a larger screen can reduce glance durations up to some point, where information is easy accessible and additional increasing has only a minor effect (maybe a further enlargement can even worsen glance durations). At the same time, a larger information presentation can slightly increase the glance frequency, which even counteracts potential benefits. In terms of SEEV, the top-down features (expectancy and value) should not be affected by resizing. The effort to move the gaze to the HMI should also not be highly affected (same location). The slightly increasing gaze frequency could be explained by increasing salience with display enlargement. What is hard to explain and model with SEEV would be the decreasing glance duration. In NSEEV, the parameters “inhibition-of-return”, or the fixation duration maybe could be used. An easier approach would be to add a bottom-up parameter, which accounts for information retrieval costs within one area of interest.

Glance frequency and acoustic click

For the analysis of glance frequencies, the data are analyzed similarly to the antecedent section and treated separately for the display states of “speed recommendation” and “prepare to stop”. Figure 6 shows the results. The hypothesis was that the acoustic click could reduce check gazes and lower glance frequency.
However, a one-tailed t-test revealed no significance in both cases. For the state of “prepare to stop”, even the direction of the hypothesis is wrong.

Table 8. One-tailed t-tests for the reduction of glance frequencies by acoustic click

<table>
<thead>
<tr>
<th>Condition</th>
<th>t(21)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>speed recommendation</td>
<td>0.577</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>prepare to stop</td>
<td>-0.414</td>
<td>&gt;0.05</td>
</tr>
</tbody>
</table>

A two-tailed t-test for paired samples between the states of “speed recommendation” and “prepare to stop” (i.e. Figure 6-1 versus Figure 6-2) reveals: t(43) = 4.7; p <0.001. So the system design, and information presentation (to show in some situations “prepare to stop”), reduces the glance frequency significantly.

One explanation with SEEV for the inefficiently working acoustic help could be top-down expectancy, together with a sufficient bottom-up visual salience. The display Figure 2-1 contains the Heuer sign in the upper right corner, an item with low informational bandwidth (it behaves like a clock and is predictable) and the dynamic speed recommendation with a medium to high information bandwidth. The display of Figure 2-4 also contains a Heuer clock and a static message with extremely low informational bandwidth. Maybe, if visual salience meets the requirements imposed by expectancy, there is not much room for improvement by the acoustic sound.

Figure 6. Mean glance frequency with and without acoustic click

System Usability Scale

The SUS scores for the different system variants are reported in Figure 7. Single factor, repeated measurement ANOVA showed no significance: F(3;63)=1.207, p=0.314.

Looking at the absolute values HMI_M seems to be the favorite. According to Bangor et al. (2009), a score of 86.4 can be connected to the adjective “excellent”. There seems to be a drop in the SUS-Score for the HMI_L. In terms of SEEV, this could be a mismatch between top-down parameters (expectancy and value) and the bottom-up parameter salience. The subjects feel it is too large.
NASA-TLX

Figure 8 shows the results of the NASA-TLX questionnaire. Single factor, repeated measurement ANOVA (Greenhouse Geisser corrected, due to Mauchly test) revealed $F(2.9;60.7)=4.577$, $p=0.006$. A Bonferroni corrected post-hoc test was significant for: baseline/HMI_S ($p=0.004$), baseline/HMI_M_C ($p=0.048$) and HMI_S/HMI_M ($p=0.047$). Looking at the absolute values, HMI_M imposes the lowest additional subjective demand compared to the baseline run.

Figure 9 depicts the results from the NASA-TLX sub-dimensions. The most subjective impact for the smartphone information comes from the “mental” and “temporal” dimensions.

![SUS Scores](image)

**Figure 7. SUS Scores of the different system variants**

![NASA-TLX Overall Workload Index (OWI)](image)

**Figure 8. NASA-TLX Overall Workload Index (OWI)**
User preferences

The test subjects were asked which variant they preferred (forced choice).

Table 9 contains the results. This led to a split decision between HMI_M and HMI_M_C.

Table 9: Number of users that prefer a specific system variant

<table>
<thead>
<tr>
<th>Variant</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMI_S</td>
<td>3</td>
</tr>
<tr>
<td>HMI_M</td>
<td>8</td>
</tr>
<tr>
<td>HMI_M_C</td>
<td>8</td>
</tr>
<tr>
<td>HMI_L</td>
<td>3</td>
</tr>
</tbody>
</table>

The comments of the participants were written down and grouped. Table 10 lists the remarks expressed most often. The most frequent answer reveals that display components can be even subjectively too large. The often sounding signal can be a result of driving style: Some participants drove around the border of the speed recommendation and got a lot of clicks when they drove in and out of the velocity carpet (despite an implemented hysteresis).

Table 10: Top three comments with number of occurrence

<table>
<thead>
<tr>
<th>Comment</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation of HMI_L is too large</td>
<td>7</td>
</tr>
<tr>
<td>Acoustic click is a good idea</td>
<td>6</td>
</tr>
<tr>
<td>Acoustic signal sounds too often</td>
<td>5</td>
</tr>
</tbody>
</table>

General gaze histograms and metrics

For the following histograms, the gaze data were exported from D-Lab and processed by Matlab. An additional minimal threshold of 120ms for single glances was introduced. The histograms only show data when the smart phone was in the

Figure 9. NASA-TLX sub-dimensions
“speed recommendation” display state (Figure 2-1. Figure 2-2). For the baseline condition, the display was blanked (black), but the smartphone reported to the eye tracking system if a speed recommendation would have been shown to get comparable road sections. Figure 10 shows the distribution of gazes to the smartphone. The distribution data in the upper right corner reveals that the 85th percentile of all gazes from all participants is 0.88 seconds.

Figure 10. Gazes to the phone, condition HMI_M

The smartphone shows a speed recommendation and can induce people to combine gazes to the phone with gazes to the speedometer. So Figure 11 holds the times, participants do not look to the area “windscreen” (called “Eyes off the road”). Figure 11 shows that the 85th percentile value of this metric (1.12 seconds) is also below the limit of current major guidelines. Figure 12 holds the same metric for the baseline condition.

Figure 11. Eyes off the road, condition HMI_M

Figure 12. Eyes off the road, Baseline
Figure 12. Eyes off the road, condition Baseline

Figure 13 and Figure 14 hold the distribution of gazes through the windscreen for the conditions HMI_M and baseline (both under the condition that the display shows a speed recommendation). Both figures show an exponential decay. For modelling, the lowest and highest bin from the figures were excluded. An exponential curve fitting led to:

\[ \text{occurrence}(t) = 233 \cdot e^{-\frac{t}{2.28s}} \] for Figure 13 with a fitting of \( R^2=0.945 \) and

\[ \text{occurrence}(t) = 144 \cdot e^{-\frac{t}{2.82s}} \] for Figure 14 with a fitting of \( R^2=0.932 \).

The constant (233 or 144) adjusts for the relative differences, but the difference in mean lifetime (2.28s or 2.82s) should be an interpretable gaze metric of a system. The baseline has a shallower curve (mean lifetime 2.82s) with relatively longer gazes to the front. The more often used term with exponential decay is half-life (= ln(2) * mean lifetime), which is 1.95 seconds for Figure 14. The visual demand of a system lowers the mean lifetime and, accordingly, the half-life of front gazes. Or to put it another way: Gazes through the windscreen are a kind of eyes-off-the-system glances.

Krause and Bengler (2012b) used an artificial metric to compare different HMI variants: As measure of severity, the average gaze frequency was multiplied with the
85th percentile value (both values are derived from the data pool of all participants, not with an intermediate step of individual averages). Thus, the severity measure for the “natural” behaviour (e.g. speedometer checks) in Figure 12 would be 0.23Hz x 1.04s = 0.24. The severity with speed recommendations from the smart phone (Figure 11) would be: 0.32Hz x 1.12s = 0.36. We interpret the difference of 0.36 – 0.24 = 0.12 as an additional visual load imposed by the smartphone. A similar value can be obtained by calculating the severity of Figure 10: 0.16Hz x 0.88s = 0.14.

The multiplication of frequency by mean glance duration gives a percentage amount of time. For instance, according to the data in Figure 10, the participants show 0.16Hz x 0.63s = 0.1, so 10% of the time to the phone. The multiplication of frequency by the 85th percentile of glance duration maybe can include some information about how heavy the tail of the glance distribution is. The long gaze durations in the tail are important for driving safety (Horrey et al., 2007).

Conclusion

An increase in display size can reduce glance duration, but it does not necessarily reduce the percentage of time an IVIS is looked at. Furthermore, subjective ratings indicate that display contents can even be shown too big. The gaze durations for the traffic light assistant are in line with limits of major guidelines, even when displayed on a small smartphone or must be scaled by height due to landscape mode.

An acoustic click (as hint for display changes) does not reduce glance frequency as expected. Subjective ratings are divided. The click will be implemented in the final system and can be enabled by the user. This will also include a function to suppress fast successive clicks for some driving styles. Appropriate information presentation design has shown to reduce glance frequency for this system more effectively than acoustic feedback.

The subjective ratings from NASA-TLX dimensions indicate that participants feel a slight mental and temporal demand when using the system. This was assessed in another experiment with a tactile detection task (TDT), to get objective data.

The different comments of the participants motivated to carry out some minor tweaking of the system. With a view of the already excellent SUS-scores, it is clear that, with each iteration of the engineering process, it gets much more difficult to score higher.

A set of gaze histograms and calculated gaze metrics was provided to enable comparison with other experiments and IVIS, as well as discussion with researchers.

Outlook

From the results of this experiment one can argue: Maybe the interface and information are within the limits for gaze behaviour, but how strong is the mental demand on the driver? Another experiment was carried out to test this and measure the mental demand of the information system with a tactile detection task (TDT).
The TDT is a tactile variant of the detection response tasks (DRT), which are currently in a standardization process by ISO TC22 SC13 WG.8.

In the experiment reported here, the traffic lights were acting on a predetermined and coordinated scheme (green wave). Afterwards, the traffic lights on the test roads were equipped with data modems to transmit their current state and were switched back to a traffic-adaptive control scheme. Based upon the transmitted process data, a central server tried, based upon historical data, to estimate how the traffic light probably will act within the next couple of minutes. The server sent this forecast to a demonstration car and smart phones. In the last field experiment, we tested this improved app and examined if the estimated switching times can be used to inform a driver.

Acknowledgement

The authors would like to thank Philipp Schneider and Johannes Henel for preparing the gaze data.

References


