Comparison between Occlusion Methods for Measuring Distraction Caused by Smartphone Interaction

Christian Jernberg
Pontus Larsson
Driver Environment & Human Factors, Advanced Technology & Research, Volvo Group Trucks Technology, SE-40508 Göteborg, Sweden (christian.jernberg@volvo.com)
Paper Number 15-0404

ABSTRACT

Research Question/Objective

The occlusion method is an established method for measuring the visual demand from in-vehicle interfaces. The method is usually employed by means of occlusion glasses which can be automatically made opaque or transparent. However, occlusion can also be achieved by turning on and off the screen of the interface being studied. It can be hypothesized though, that glasses-occlusion requires more visual re-orientation efforts after each occluded interval which may give longer Total Shutter Open Times (TSOTs) than the method where only the interface screen is occluded.

Methods and Data Sources

Two experiments were conducted, with 10 participants (5 female) in Experiment 1 and 14 participants (7 female) in Experiment 2. In both experiments, participants were seated in a truck mockup and performed tasks on a smartphone. Two occlusion conditions were employed: 1) Occlusion with glasses, and 2) Embedded occlusion (turning the screen of the smartphone on/off). Participants also performed a baseline condition with no occlusion. In Experiment 1, five different tasks were performed in each condition: destination selection, dialing a phone number, dialing a contact, changing radio frequency and setting the alarm clock. For each task, a note showing the data to be entered was posted next to the smartphone, e.g. the number to be dialed.
The four tasks used in Experiment 2 were: visual-manual text entry, make changes in a truck driving log, dial a phone number and select item from a map (using pinch/swipe operations). The participants did not have support from any note but had to remember the data to enter for each task.
For both experiments, TSOTs and Resumability (R) were submitted to separate ANOVAs to reveal differences between occlusion and tasks.

Results

The results of Experiment 1 showed a main effect of occlusion type on TSOT ($p < .05$), where Glasses resulted in longer TSOTs ($M = 8.1$s) than Embedded ($M = 7.4$s). However, the interaction between occlusion and task was also significant ($p < .01$), and suggested that mainly the phone dialing task caused the difference in TSOTs. An explanation to this could be that tasks requiring more visual-spatial reorientation are more difficult to perform with glasses occlusion. A simpler explanation could be that participants needed to look at the note while dialing – which can be done during the occluded intervals with the embedded method. The results from experiment 2 did not show any statistically significant effect of occlusion on TSOTs which suggests that the effects found in Experiment 1 were caused by the fact that participants had to look at the paper note while dialing.

Discussion and Limitations

The current experiments indicate that the embedded method gives TSOTs comparable to the glasses-occlusion method. However, one must be aware of the fact that there are practical differences between the methods which can lead to diverging TSOTs. The current experiments show that such divergence can occur when instructions are given visually.

Conclusions and Relevance to Session Submitted

The embedded occlusion method for evaluating distraction is nonetheless promising since it is easy to use and can be integrated in app developers' toolkit.

INTRODUCTION

The occlusion method is well adopted and established for measuring the visual demand of in-vehicle information systems (IVIS). The method dates back to the late 60s and was developed by Senders et al.[1] in order to find a suitable surrogate for the more expensive and time consuming eye-tracking method.

The International Organization for Standardization (ISO) state in their guideline for occlusion that when using the occlusion method, the system that are being tested should work as intended and be fitted in a vehicle, simulator buck, or mock up in a design which duplicates the placement of the system in its intended environment. The viewing angle and the placement of the controls in relation to the driver should be the same as in the finished product[2]. Before testing the system the participant should be allowed to practice until he or she can adequately handle the system. A rule of thumb is mentioned as between two and five test trials per function, including two with occlusion. After the practice, a total of five test trials should be performed with similar end state of the system. Each function should be practiced and evaluated separately before moving on to the next.

The ISO standard specifies three possible ways to use occlusion, and these are as follows:

1 – Occlusion glasses
2 – Display blanking
3 – Mechanically or electronically operated shutters.

All these methods have in common that they are not allowed to interfere with the manual controls of the IVIS and the change from occluded to un-occluded state (opaque to transparent) or vice versa cannot take longer than 20 ms. The length of the interval during which the shutter is open or closed is 1.5s.

The two most relevant measurements used in this type of evaluation are the Total Shutter Open Time, TSOT, and the Resumability Ratio, R. The TSOT value represents an estimation of how long the driver would have to avert the gaze from the road to perform a task and R is a measurement of how easy it is for the driver to continue a dialogue after it is interrupted. R is calculated as the ratio between the TSOT and the total task time unoccluded (TTT_{unoccl}).

The occlusion method is accepted by ISO, the Society of Automotive Engineers (SAE)[3], the National Highway Traffic Safety Administration (NHTSA)[4], the Alliance of Automobile Manufacturers (AAM)[5] and the Japan Automobile Manufacturers Association (JAMA)[6] to determine if a system is safe enough to be used while driving. However, the criteria for how the method should be used might vary slightly. In the ISO standard the criteria is up to the person who is performing the evaluation or developing the interface. JAMA on the other hand say that the limit for total task time is 8 seconds, which make the maximum acceptable TSOT 7.5 seconds. For the NHTSA and AAM guidelines the maximum TSOT are twice as long, 15 seconds.

Occlusion should be optimal for evaluations in the early phases of the design process[7]; the main reason being that it is comparably cheap and simple to perform. Therefore, it is possible to have a larger number of participants for several design loops and iterations compared to when the eye-tracking method is used for example. So even though the occlusion method is not as precise in measuring where the most problematic areas of an IVIS are as the eye-tracking method, it can still give an indication of which concept that is the least distracting when comparing different ones[8]. Furthermore, if the display blanking variant of occlusion can be integrated in software development packages, it should even further simplify the process of iteratively developing and testing new in-vehicle interfaces.

In this study, the focus lies mainly with comparing the occlusion glasses to display blanking (hereafter referred to as “embedded occlusion”), and it is hypothesized that there are some differences between these which may lead to that the different methods result in different TSOT- and R values. For example, with the glasses there should be a short period of time where the eyes need to adjust to a new focal distance; something not needed when it is only the screen of the IVIS that is turned blank. Also, while the glasses do not interfere with the manual controls, it covers these as well as the hand of the operator. This is not the case with embedded occlusion and thus it should be easier to resume a task or even continue a task during the occluded interval (so called ‘blind operating’) since there is still the ability to see where the fingers are positioned relative to the edge of the IVIS. This hypothesis is tested in two experiments presented in the following paragraphs in which a set of apps in a regular smartphone is evaluated using both the glasses-occlusion method and a smartphone-embedded method.

METHODS AND DATA SOURCES

Two experiments were conducted, between which the main differences were that Experiment 2 included tasks of higher difficulty and that participants did not receive any visual instructions during the tasks.

Experiment design

For both experiments, a within-group factorial design was used with Occlusion type (None/ Glasses / Embedded), Task (1-5 for Experiment 1 or 1-4 for Experiment 2) and Task repetition (5) as independent variables. The order of conditions was balanced between participants in order to avoid order effects.

Participants

In Experiment 1, 10 participants (5 female) took part. Their mean age was 39.5 years (*SD* = 9.25). Two participants were over 50 years old. In Experiment 2, 14 participants (7 female) took part. Their mean age was 36.7 years (*SD* = 6.98). All participants were recruited from within the company.

Apparatus and equipment

Apparatus and equipment was identical for the two experiments. A Samsung I9001 Galaxy S Plus smartphone was used as the main interface with which the participants would perform the task. The smartphone was positioned in a cradle mounted to the dashboard of a truck cab mockup. The truck cab mockup consisted of a truck seat, in which the participants were seated during the whole experiment, a steering wheel, and a complete dashboard (see Figure 1).

Occlusion was accomplished with PLATO visual occlusion spectacles for the glasses-occlusion conditions (see Figure 2) and with a custom app for the embedded occlusion conditions. Timing of the tasks was accomplished with a regular stop watch app in a smartphone.

Figure 1. Truck cab mockup used in both experiments


Figure 2. Person wearing occlusion glasses.

<u>Tasks</u>

In both experiments, five variations of each task were to be completed (which is according to the ISO standard), for example: call five different phone numbers or select five different destination addresses.

In Experiment 1, five different tasks were to be conducted in each occlusion condition. For each task in Experiment 1, a note showing the data to be entered was posted next to the smartphone, e.g. the number to be dialed. The tasks were:

1. Change radio frequency: In the smartphone's built-in radio app, the participants had to repeatedly tap either a left or right arrow until a frequency given by the paper note had been reached. In total, the arrow had to be tapped 31 times to complete each variation of the task.

2. Select destination: In the navigation app "Waze", the participants had to tap to open the menu, tap "navigate", scroll through a list and select one of five pre-defined destinations and then finally tap to start navigation.

3. Set the alarm clock: In the smartphone's built-in clock app, the participants had to first tap "create alarm", then tap the hour digits and through the appearing on-screen keyboard enter one out of five given times (20:00, 00.00, 17:00, 10:00, 18:00), tap "next", tap "repetition of alarm", tap "every day" and finally tap "Save" twice.

4. Dial a phone number: In the smartphone's regular phone interface, the participants had to manually enter one out of five phone numbers – all containing a long push "+" and 13 digits (e.g.: +49 157 3943 3782) and then tap the green phone symbol to call the number.

5. Dial a contact: In the smartphone's regular contact list, the participants had to scroll to one out of five names, tap the name, and finally tap the contact's phone number. All names started with the letter "M" and the participants had to scroll down approximately 30 contacts before the target contact became visible.

In Experiment 2, four different tasks were to be conducted in each occlusion condition. In this experiment, no note displaying the target data (i.e. phone number, destination address etc.) to be entered was used; the participants had to remember the data to be used in each task, and this was given verbally by the test leader before each task. The tasks were:

1. Enter destination: In the navigation app "Waze", the participants had to tap to open the menu, tap "navigate", tap in the search field and manually enter one out of five addresses given by the test leader with the on-screen keyboard, then tap the search icon, select the appropriate destination address from the list appearing and tap "drive" twice to start navigation. The addresses were Importgatan, Planetgatan, Teknikgatan,

Kvillegatan, and Londongatan (all selected from the vicinity of the test site so that no time lag would occur after the search/navigate icon had been tapped.)

2. Change time in log: In the app “TimeBook” (a tool for truck drivers to record their working and resting times), the participants had to tap “list” to open a list with logs, then tap one out of five specific logs (identified by the number of tasks they contained – between 21-25) , scroll to the last task in that log, tap the task, tap “end time” for the task and tap the “+” icon twice to extend the task by 2 hours and finally press “enter”.

3. Dial a phone number: In the smartphone’s regular phone interface, the participants had to manually enter one out of five phone numbers – all containing a long push “+”, the numbers “495” and then 10 digits. Since the participants had to remember the number sequences by heart, logical number sequences were used (e.g.: +495 24 68 10 12 14). When the number had been dialed, the participants had to tap the green phone symbol to call the number.

4. Call dealer: In the app “DealerLocator” (a tool for truck drivers to find nearby truck dealers), the participants had to use multiple swipe / pinch-zoom operation to navigate a digital map, starting from Göteborg, Sweden and find one out of five dealers in the Barcelona, Spain region. Once the dealer had been located, the participant had to tap the dealer marker, and finally tap the dealer’s phone number to make a call.

Procedure

The procedure was identical for both experiments. Participants arrived individually to the test lab. They were firstly introduced to the study, the smartphone, the occlusion glasses and the embedded occlusion app. Then, the tasks to be performed during the test were demonstrated.

The actual test then started. Before each task, participants practiced the task (using other data entries than the ones used in the experiment trials) until they could perform them without error. They then got to practice the task also with occlusion (either embedded or glasses depending on the condition) before an actual trial was conducted during which the task completion time was recorded by the test leader. The test leader instructed the participants verbally to start each task by saying “start… now” and the participant notified when he/she was finished with a task by saying “done”. If the task was not completed without errors, they were asked to perform the task once more. After all tasks had been completed, participants were debriefed and thanked for their participation.

RESULTS

Using the acquired Total Task Times (TTT), the measures Total Shutter Open Time (TSOT) and Resumability ratio (R) were first calculated according to:

$$TSOT = 0.75 \cdot \left(INT\left(\frac{TTT}{1.5}\right)\right) + 0.5 \cdot MOD\left(INT\left(\frac{TTT}{1.5}\right), 2\right) + MOD\left(\frac{TTT}{1.5}, 1\right) \cdot \left(1 - MOD\left(INT\left(\frac{TTT}{1.5}\right), 2\right)\right) \quad (1)$$

$$R = \frac{TSOT}{TTT_{unoccluded}} \quad (2)$$

TSOT and R values were then submitted to separate 2x5x5 - occlusion (embedded/glasses) x (task number) x (task repetition) and 2x4x5 - occlusion (embedded/glasses) x (task number) x (task repetition) - repeated measures Analyses of Variance (ANOVAs) for Experiment 1 and Experiment 2 respectively.

Experiment 1 results

Considering TSOT, a statistically significant main effect of occlusion was found: $F(1,8) = 8.323$, p = .02. Post hoc test using Bonferroni's adjustment for multiple comparisons showed that the Glasses occlusion condition resulted in longer TSOTs than the Embedded occlusion ($M = 8.122$, $SE = 0.423$ compared to $M = 7.397$, $SE = 0.436$).

A main effect of task was also found: $F(4,32) = 61.843$, $p < .001$ (Sphericity assumed). Post hoc test showed that task 4 (dial phone number) resulted in statistically significantly longer TSOTs than all the other tasks (Bonferroni's method used to adjust for multiple comparisons). The effect of task on TSOT is shown in Figure 3 below.

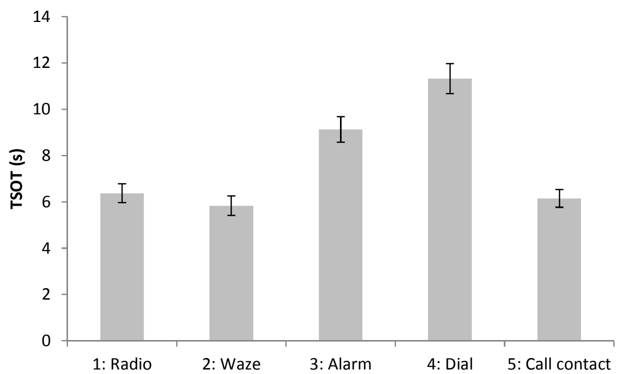

Figure 3. TSOT as function of Task for Experiment 1. Whiskers show standard error.

However, a statistically significant interaction effect between occlusion and task was also found: $F(4,32) = 6.143$, $p = .001$, Sphericity assumed. The effect is visualized in Figure 4 below, where it can be seen that it seems it is mainly Task 4 (Dial a phone number) for which the occlusion types differ.

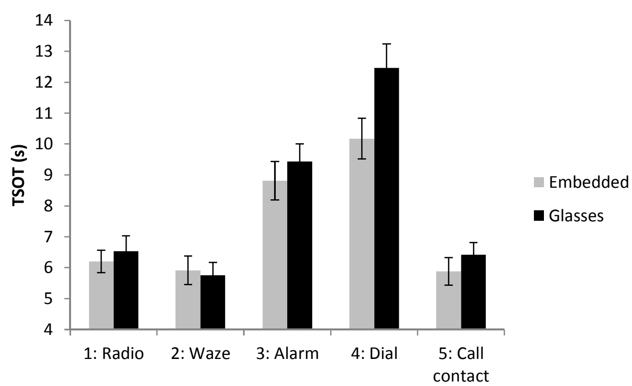

Figure 4. Interaction between occlusion and task for TSOT, experiment 1. Whiskers show standard error.

For Repetition, a main effect was found ($F(4,32) = 6.495$, $p = .001$), and from the result of a Bonferroni-adjusted post hoc test, one could see that this effect was due to repetition 4 resulting in a statistically significantly lower TSOT compared to repetition 1 ($M = 7.485$, $SE = 0.423$ vs $M = 8.313$, $SE = 0.516$, respectively, $p = .01$).

For R, no main effect of occlusion was found, although a trend could be noted indicating that Glasses occlusion could give higher R than embedded occlusion ($M = 0.849$ vs $M = 0.791$, $p = .067$). A similar trend was found for the interaction between task and occlusion. A statistically significant main effect was however found for task: $F(4,28) = 7.579$, $p < .001$. Figure 5 below shows that it was mainly the difference between Task 1/ 2 (Radio / Waze) compared to Task 3 (Alarm) and the difference between Task 1 and 4 (Dial) which caused this effect (Bonferroni's method used to adjust for multiple comparisons). No other effects were found.

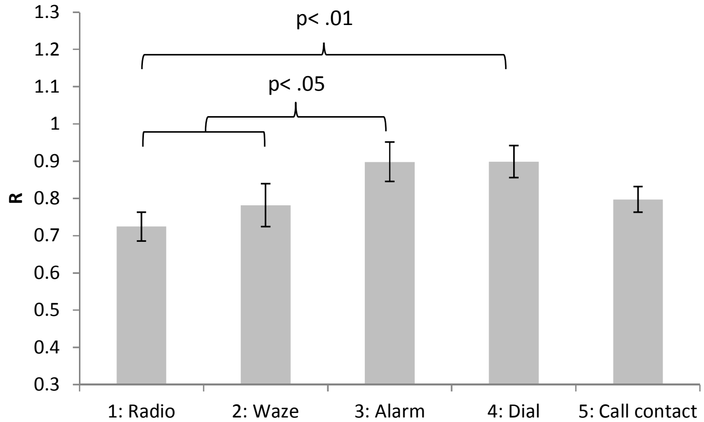

Figure 5. R as function of task for R, Experiment 1. Whiskers show standard error.

Taken together, the Dial and Alarm tasks seemed to require more glance time and were also more difficult to resume compared to the other tasks. The Dial task seemed to be the reason why a main effect of occlusion on TSOT was found. The reason for this effect may be that more difficult tasks, which require more visual reorientation than simpler tasks, are more difficult to perform with the glasses occlusion. A much simpler explanation could be that the Dial task required the participants to look at the paper note to be able to complete the task (the data to be entered for the other tasks were much simpler to remember). This could obviously not be done while the glasses were shut, compared to the Embedded occlusion where participants could read the number to be dialed also during blanked-sceen intervals.

These explanations were the reason for conducting Experiment 2, in which more difficult tasks were included and the participants did not have the paper note to read off.

Experiment 2 results

There was no effect of Occlusion on TSOT in experiment 2 ($F(1,13) = .160$, $p = .695$, $M = 15.163$ for embedded occlusion and $M = 15.708$ for glasses occlusion, $p = .695$). A main effect of Task on TSOT was however found: $F(1.637, 21.284) = 56.4$, $p < .001$ (Greenhouse-Geisser corrected). Post hoc test using Bonferroni's method of adjustment showed that all tasks differed from each other ($p < .01$); means and standard errors are visualized in Figure 6 below. Here it can also be seen that the Dial task, when performed with numbers in a logical sequence, resulted in lower mean TSOTs compared to Experiment 1, when the participants had to read on the paper note (of course not statistically tested here however). Interestingly, a similar effect of repetition as found in Experiment 1 was also found in Experiment 2: $F(4,52) = 4.401$, $p = .004$. Again, it seemed that Repetition 4 resulted in the lowest TSOTs, but also that the last repetition gave longer TSOTs than the others. The general trends for repetition for both Experiment 1 and 2 are shown in Figure 7.

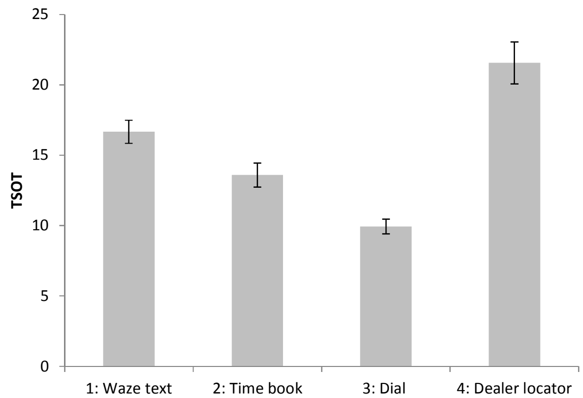

Figure 6. TSOT as function of task, experiment 2. Whiskers show standard error.

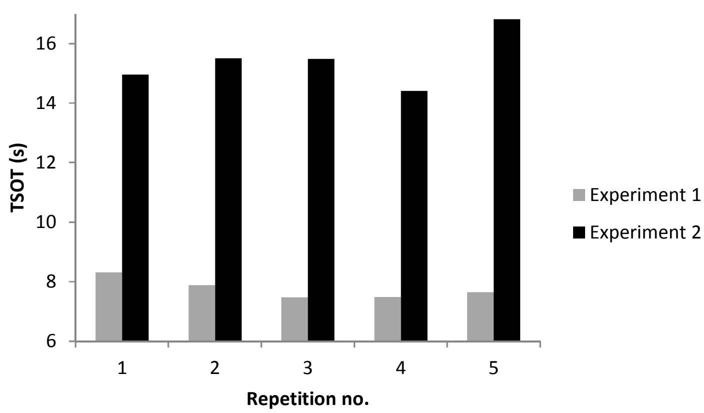

Figure 7. TSOT as function of repetition comparing both experiments.

No statistically significant effect of Occlusion on R was found: $F(1,12) = .013$, $p = .912$. Moreover, there were neither any statistically significant effects of Task ($F(3,36) = 2.348$, $p = .089$) or Repetition ($F(4,48) = 1.599$, $p = .190$) on R.

DISCUSSION AND LIMITATIONS

In Experiment 1, a statistically significant effect of occlusion on TSOT was found. However it was suspected that this effect was only due to one of the tasks, namely “Dial phone number”. One possible explanation was that the use of visual instructions (a paper note with the number to be dialed) caused this effect. In case of the embedded occlusion, the phone number could obviously be read also during screen blanking, but when wearing the glasses the number can only be read during shutter-open intervals. Another explanation was that the occlusion types generate different results only for more difficult tasks which require visual reorientation, in turn which could be more difficult when the whole scene is blanked out as in the glasses-occlusion situation.

Experiment 2, in which there was no paper note to be read from, did not result in any differences in TSOT depending on occlusion type. Experiment 2 also included more difficult tasks, especially one involving interacting with a map which one would expect to be more difficult for the glasses occlusion - but even the measures obtained from this task did not seem to be affected by occlusion type. Thus, we can be fairly certain that the two occlusion methods will not generate different results, if visual instructions are avoided. It is important to note however, that the current set of experiments has not considered visual instructions and task difficult as factors in a factorial design (since this was not the purpose/hypotheses of the experiments from the beginning). To be able to draw firm conclusions, another experiment should be conducted where visual instructions and task difficulty are independent variables.

CONCLUSIONS AND RECOMMENDATIONS

Our conclusion is that the occlusion method is a good way to early on determine if a system is too distracting for a driver or not. Even though it is not as accurate as eye-tracking, the simplicity with which it can be performed makes it ideal for app-developers to use while developing automotive applications. With the embedded occlusion method, they could test their systems before presenting them to a vehicle manufacturer and thus save time and resources for both parties.

However, a recommendation would be to add in the ISO guideline that the occlusion method should not allow visual instructions (apart from such that are presented in the actual interface under study) during the trials as this could interfere with the results of the evaluations.

ACKNOWLEDGEMENTS

The work presented in this paper was conducted as a part of the ongoing project SICS (Safe Interaction, Connectivity and State), co-financed by the Swedish Governmental Agency for Innovation Systems (VINNOVA).

REFERENCES

[1] Senders, J.W., Kristofferson, A.B., Levison, W.H., Dietrich, C.W. and Ward, J.L. 1967. “The attentional demand of automobile driving.”
Highway Research Record, No 195: 15-33

[2] ISO 16673:2007. ”Road vehicles – Ergonomic aspects of transport information and control systems – Occlusion method to assess visual demand due to the use of in vehicle systems.”
Available for purchase at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=38035. Accessed March 3, 2015.

[3] SAE J2364, 2004. “Navigation and Route Guidance Function Accessibility While Driving”.
Available for purchase at: http://standards.sae.org/wip/j2364/. Accessed March 3, 2015.

[4] NHTSA: National Highway Traffic Safety Administration, 2013. “Guidelines for Reducing Visual-Manual Driver Distraction during Interactions with Integrated, In-Vehicle, Electronic Devices, Version 1.01 Docket No. NHTSA-2010-0053”.
Available at: https://www.distraction.gov/download/11302b-Visual_Manual_Distraction_Guidelines_V1-0_010815_v1_tag.pdf. Accessed March 13, 2015.

[5] Driver Focus-Telematics Working Group, 2006. “Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems.”
Alliance of Automobile Manufacturers, Washington, DC.

[6] JAMA: Japan Automobile Manufacturers Association Inc., 2004 “Guidelines for In-Vehicle Display Systems - Version 3.0.”
Available at: http://www.jama-english.jp/release/release/2005/jama_guidelines_v30_en.pdf. Accessed March 13, 2015

[7] Young, K, Lee, J.D. and Regan, M.A. 2008. “Driver Distraction: Theory, Effects and Mitigation”
CRC Press, Boca Raton, FL: 123-134

[8] Kroon, M. 2006. “Comparison Between Visual Occlusion and Eye-movement based visual demand assessment of in-vehicle information systems”
Departure of Computing Science, Umeå University, Sweden