

Methodical aspects of text testing in a driving simulator

Sundin A^{a,*}, Patten C.J.D.^b, Bergmark M^a, Hedberg A^a, Iraeus I-M^c, Pettersson I^c

^a*Ergonomics/HMI, Semcon AB, 417 80 Gothenburg, Sweden*

^b*VTI, Swedish National Road and Transport Research Institute, 581 95 Linköping, Sweden*

^c*Volvo Car Corporation, 405 31, Gothenburg, Sweden*

Abstract. A test with 30 test persons was conducted in a driving simulator. The test was a concept exploration and comparison of existing user interaction technologies for text message handling with focus on traffic safety and experience (technology familiarity and learning effects). Focus was put on methodical aspects how to measure and how to analyze the data. Results show difficulties with the eye tracking system (calibration etc.) per se, and also include the subsequent raw data preparation. The physical setup in the car where found important for the test completion.

Keywords: HMI, Methodology, Driving simulator, Eye Tracking System

1. Introduction

The presented simulator study was part of the ongoing Swedish research & development project EFESOS (Environmental Friendly Efficient Enjoyable and Safety Optimized Systems), run 2010-2013. The project is managed by Volvo Car Corporation (VCC) and is collaboration between VCC and seven other research partners whereof three are technical universities. The project is financed by the industry partners and the Swedish authorities; Swedish Transport Administration, Swedish Energy Agency and Swedish Governmental Agency for Innovation Systems, as part of the FFI Strategic Vehicle Research and Innovation program. See www.efesos.net for further information.

The study was carried out in the EFESOS sub-project USI (User Information & Interaction), Work Package 1 Multi-Modal Interaction, Task 2 - Texting and Integrated Nomadic Devices, and sub-project METOHMI Work Package 2 Evaluation methodology, Task 2 - Generic methods for the evaluation of infotainment systems and services. The work package leader Semcon, group Ergonomics/HMI, was responsible for the development and programming of a new touch screen, Semcon and VTI responsible for

the majority of the planning and data analysis of the test and the simulator activities was carried out jointly between Semcon and VCC in the VCC HMI Usability Lab .

The purpose of the test was concept exploration and comparison of existing user interaction technologies for text message handling with focus on traffic safety, ease of use and enjoyment. The driving simulator that was used was newly built by VCC, therefore focus was also put on methodical and practical aspects of calibration, design and data analysis.

2. Methodology

2.1. Driving simulator

The fixed-based driving simulator used in this study had a 180 degree screen with five projectors and with half a Volvo car mockup in the centre. The cockpit had fully functional seats, steering wheel and pedals and it had an automatic gearbox. The driving simulator was newly installed in the Volvo HMI Usability Lab (Figure 1).

* Corresponding author: anders.sundin@semcon.com



Figure 1: Volvo Car Corporation's HMI Usability Lab.

2.2. Test procedure and design

30 test persons conducted a test where a predefined text message was entered into two text input concepts of an in-vehicle information system via Touch Pad and Touch Screen. The drivers texted while driving where a longer text was the task, e.g. found in texting (sms), email, blog or Facebook update, in contrary to shorter texts usually used in navigation tasks navigation tasks. The touch screen and touch pad were mounted in the cockpit positioned as preferable as possible with respect to physical ergonomics (Figure 2). A pilot test was carried out a week before the actual tests in order to check the setup, methods, questionnaires and Flash simulation as well as the equipment in the HMI lab such as eye tracking equipment, simulator and test rig.

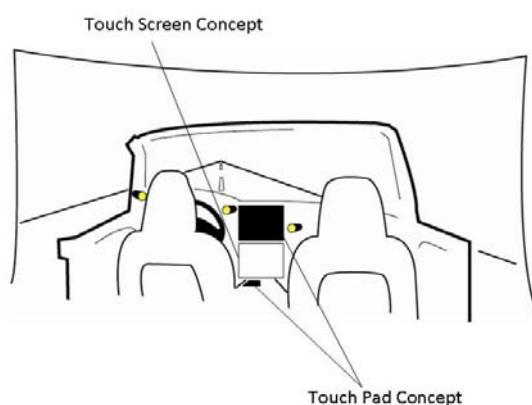


Figure 2: A sketch outline of the simulator and basic placements of the Touch Screen and Touch Pad concept.

2.3. Driving scenario

The driving scenario was a car-following scenario on a two lane highway with other interacting vehicles.

2.4. Hypotheses

Evaluation aspects incorporated

- Safety
- Ease of use
- Enjoyable, engaging control

The overall research question was in what extent safety, ease of use, and enjoyment are fulfilled in the Touch Screen and Touch Pad concepts during in-vehicle texting. 14 hypotheses were put forward on both primary and secondary task.

2.5. Test participants

There were 30 test participants recruited from the VCC facility and Semcon office in Gothenburg. The mean age was 40,4 years, and had previous experience of sending text messages.

2.6. Data collection and analysis

Collected data from the simulator was Standard deviation of lateral position (SDLP), Time headway (TH), and Task completion time. From an Eye Tracking System: Percentage road centre (PRC), Glance frequency and Glance duration. Furthermore was Keystrokes per character – (KSPC) measured. Finally, a questionnaire of 34 questions using a five step answering scale gave ratings on safety, ease of use and enjoyment.

2.6.1. Statistical analysis

Parametric t-test and non-parametric Wilcoxon statistical analyses were performed using the SPSS 17.0 statistical program.

3. Results

Approximately 68% of the tests resulted in data used in data analysis. At least 23 out of 34 test persons could be used in the data handling. Three test persons got simulator sickness, for six test persons it was not a correct test setup and for two test persons there were poor eye-tracking data as result.

3.1. Eye Tracking System

The results regarding methodical aspects show that the Eye Tracking System created problems in setup and in data output e.g. sequences errors in the data, omitted data and glance frequency errors.

The definition of road center was not fully made for the test. Furthermore, a reference task was not used, as intended. As the test consisted of two concepts it was possible to compare the two, but with a reference task, e.g. the Haste test, it would be possible to benchmark against other tests and concepts.

The touch pad ($t(1, 20) 8.130, p = <.001$) and touch screen ($t(1, 21) 6.566, p = <.001$) are significantly different from the baseline condition but there was no statistically significant difference between the touch screen and the touch pad for the road centre eye tracking metric.

Analyses of glance frequency, glance duration and total glance time was not successfully collected and the data deemed invalid.

3.2. Time on task

There was a significant difference for the Time on task metric analysed with a paired t-test ($t(1, 21) 2.072, p = .05$) between the touch screen interface and the touch pad interface (figure 3).

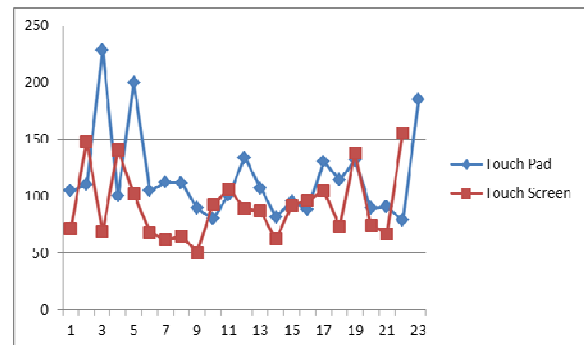


Figure 3: Time on task metric in seconds between two different text entry interfaces.

Mean time on task for the touch pad was 112.9 seconds and touch screen was 91.6 seconds.

To complete the texting task, participants needed on average 21.3 seconds more time using the touch pad interface compared to the touch screen interface.

3.3. Physical set-up

The importance of planning the physical setup in the car was found to be an important aspect to be able to carry out the test, whereas positioning of the concepts in height and distance to the driver were found most crucial for each concept. Also the positioning of the eye-tracking cameras was important to be able to carry out the calibration and collect the eye-tracking data from test persons with glasses etc.

4. Discussion

The reason for found glance frequency errors was that the translation time was not counted for in the test. A glance is defined by the time it takes to refocus (translation time) + focus on new area + refocus on first area (translation time). The test set-up only gave time spent on the different areas. The translation times was not considered.

To be able to effectively capture data, later to be used in analyses and be a foundation for decisions how to proceed with different concepts, it is essential to do detailed planning early. This early planning should emphasize on which data source to be used, what amount of data is the minimum, in what way will the data be analyzed etc. The aim is to minimize time spent in the data analysis phase and avoid data output errors e.g. sequences errors in the data, omitted data and glance frequency errors.

The road centre eye tracking metric was not significant between the two interfaces, however the time on task metric was significantly different. This discrepancy could imply that the eye tracking road centre requires improved calibration in order to portray a more correct evaluation.

Due to the problems with data acquisitions, the possibility to answer all stated hypotheses was limited.

References

- [1] M. Bergmark, Visual performance of different drivers and alternative positioning of displays and buttons in vehicles, Essay Vehicle ergonomics, Chalmers Tekniska Högskola
- [2] ISO9241-410-2008preview