In-Vehicle Screen Density
Driver Distraction and User Preferences for
Low vs. High Screen Density in Integrated Displays

Hanna Johansson & Katarina Walter
LIU-KOGVET-D--05/19--SE
2005-09-15

Masters Thesis in Cognitive Science
Supervisor and examiner: Jonas Lundberg
Department of Computer and Information Science
Linköpings universitet, Sweden
ABSTRACT

Many information technology artefacts can be found in today’s cars. The interaction with these artefacts is the driver’s secondary task while driving the car in a safe way is the primary task. When designing interfaces for in-vehicle usage, measures have to be taken in order to make the interaction with the artefact suit the in-vehicle environment. One of these measures is to have the appropriate screen density level, which is the amount of information present on the screen.

This thesis compares the usability of two integrated in-vehicle display prototypes, one with low screen density and one with high screen density. The usability comparison considers both safety and user preferences. Safety was measured by a Lane Change Test (LCT) which measures distraction of a primary task while performing a secondary task, and user preferences was measured with a questionnaire. Before the comparison was made, controls and a graphical user interface were designed.

Results showed no significant difference in driver distraction between performing tasks on the high screen density display and the low screen density display. However, a vast majority of the users preferred high screen density over low. Furthermore, the distraction levels for both the high and the low screen density displays were below the proposed 0.5 meter limit for allowed driver distraction. The results indicate that in-vehicle displays can have a high level of screen density without imposing a level of distraction on the driver that is unsuitable for driving.
ACKNOWLEDGEMENTS

We wish to start by thanking Dan Nylander and Anders Hallén for providing us with the opportunity to write our thesis in the interesting field of in-vehicle usability and human factors at Volvo Car Corporation. Also, thanks to all employees at the department Human Factors Engineering and Ergonomics for showing interest in our work and making us feel welcome.

Thanks to Marianne Arkevall for introducing us to LCT, to Andreas Johansson for building the center stack prototype, and to all subjects participating in our study.

Finally we like to thank our supervisors; Anders Hallén and Staffan Davidsson at Volvo for giving us insight in the work and methods of the car industry, and Jonas Lundberg at Linköping University for encouraging and directing our thesis.

Hanna Johansson & Katarina Walter
Gothenburg, 17 August, 2005
# Table of Content

1 Introduction ........................................................................................................................................... 1

1.1 Purpose .................................................................................................................................................. 3
1.2 Method .................................................................................................................................................. 3
1.3 Delimitations .......................................................................................................................................... 4
1.4 Targeted Readers ................................................................................................................................. 5
1.5 References ............................................................................................................................................ 5
1.6 Thesis Overview .................................................................................................................................. 5

2 Volvo ....................................................................................................................................................... 6

2.1 Background to Thesis .......................................................................................................................... 6
2.2 Brief History of Volvo ........................................................................................................................ 6
2.3 Volvo Today .......................................................................................................................................... 7
2.4 Volvo Customers .................................................................................................................................. 8
2.5 Core Values .......................................................................................................................................... 9

3 Theoretical Framework ............................................................................................................................ 10

3.1 Cognitive Psychology ......................................................................................................................... 10
  3.1.1 Attention ....................................................................................................................................... 11
  3.1.2 Perception ..................................................................................................................................... 12
  3.1.3 Memory ......................................................................................................................................... 13
3.2 Human Factors .................................................................................................................................... 15
  3.2.1 Visual Demands of Driving .......................................................................................................... 15
  3.2.2 Mental Workload .......................................................................................................................... 16
  3.2.3 Measuring Driver Distraction ........................................................................................................ 17
3.3 Usability ................................................................................................................................................. 18
  3.3.1 Usability Heuristics ........................................................................................................................ 18
  3.3.2 Design Principles ............................................................................................................................ 20
  3.3.3 Guidelines for Usable Controls ..................................................................................................... 20
  3.3.4 Guidelines for Usable Interfaces ................................................................................................... 22
3.4 Screen Density .................................................................................................................................... 24

4 Method .................................................................................................................................................... 27

4.1 Heuristic Evaluation ............................................................................................................................. 27
4.2 Controls ................................................................................................................................................. 29
  4.2.1 Report from J. D. Power and Associates ..................................................................................... 30
  4.2.2 Control survey ............................................................................................................................... 32
  4.2.3 Center stack prototype ................................................................................................................ 34
4.3 The Display of Screen Density ............................................................................................................ 34
  4.3.1 Definition of Screen Density ........................................................................................................ 35
  4.3.2 QOC .............................................................................................................................................. 36
  4.3.3 Card Sorting Test .......................................................................................................................... 38
  4.3.4 Applying the Definition of Screen Density ................................................................................. 39
  4.3.5 Graphical User Interface ............................................................................................................... 40
4.4 Screen Density Experiment ............................................................................................................... 43
  4.4.1 Variables ...................................................................................................................................... 43
1 INTRODUCTION

Since the dawn of the automobile over a century ago, changes have constantly been made to improve it. These changes have been driven by safety, comfort, and technical development. In recent years, changes have been inspired by the possibility to install digital information technology in cars. Today, navigation systems, DVD-players, trip computers, mobile phones, and mini discs can be found as parts of the in-vehicle environment.

“Cars are driven by people. The guiding principle behind everything we make at Volvo, therefore is – and must remain – safety.”

(Assar Gabrielsson and Gustaf Larson, The founders of Volvo)

The driver’s primary task is to drive the vehicle in a safe manner. When a driver starts interacting with an in-vehicle digital artefact while driving, this interaction has to become the driver’s secondary task. It is of great importance that the interaction between driver and artefact has the smallest possible effect on the task of driving the car. Consequently, the interaction should not be designed in a way that forces the driver to focus solely on the artefact. Note the use of the word interaction. Interaction is something that occurs between driver and artefact when the driver is using the artefact. In order to design the artefact in the way argued above, it is the usage of the artefact that has to be considered, not only the isolated artefact. This thesis investigates one aspect of making the interaction with the digital artefact suit the in-vehicle environment. This aspect is called screen density and is described briefly later in this chapter, and in detail in chapter 3.4 Screen Density.

Interaction can be described as the communication between user and artefact. For example, users can communicate with digital artefacts by applying pressure to keys, by talking into microphones, or by moving in front of sensors. But an equally important part of the interaction is the different ways of communication that digital artefacts employ. A digital artefact can communicate to a user by sending sound through speakers, using motors to move, or most commonly, displaying information on a display. This thesis is investigating interaction by controls and display. More specifically, controls placed in the center stack, and an integrated display placed on top of the center stack (see Figure 1a for center stack, and Figure 1b for placement of
display and controls). In an integrated display, several functions communicate to the user through one single display as opposed to each function having a separate display. In this case: radio, navigation system, phone, CD, DVD, TV and hard drive (HD) communicate through one integrated display.

Figure 1a: Center stack within yellow frame

Figure 1b: Center stack seen from the side.

When designing an in-vehicle digital artefact to have minimum effect on the driving task, there are several measures to take. For example, measures can be taken to follow recommendations for the placement of controls or the height of characters on the display. Unfortunately, there are yet no recommendations concerning screen density for in-vehicle artefacts. Therefore, the aim of this report is to study screen density in an integrated display, and investigate which density level has the smallest effect on the driving task.
Screen density is the amount of information present on a screen and it is measured by dividing the number of characters on the screen with the number of possible character spaces on the screen. By doing this, screen density is expressed as a percentage. In this study, low density is defined as 13-33% and high as 32-53%. Earlier research on screen density has pointed in two directions; some experiments have shown that human performance is better with high screen density than low (Staggers, 1993), other studies have show the opposite (Cohen & Ivry, 1991; Somervell et al. 2002).

By investigating which screen density has the smallest effect on driving, this thesis explores the safety aspect of screen density for in-vehicle systems. In addition to this, it also investigates which screen density real end-users prefer. Previous studies have shown that users prefer medium to high density (Morrison et al 1989; Ross et al., 1994), or that they are equally satisfied with either type of density (Staggers, 1993). By combining safety and user preferences, this thesis aims to establish which density level has the best usability.

*Good design is not only a matter of styling the surface. It is just as important to make the product easy to understand and use. If the product is not functional, it can’t be beautiful.*

(The Volvo design philosophy)

1.1 Purpose

The purpose of this thesis is to compare the usability of two integrated in-vehicle displays, one with low screen density and one with high screen density. The usability comparison considers both safety and user preferences. Through this purpose the following problem statements have been formulated:

- What level of screen density, high or low, can be recommended for a safe in-vehicle display?
- What level of screen density, high or low, do users prefer?

Safety is tested through a Lane Change Test (LCT), and user preferences are measured with a questionnaire to real end-users.

1.2 Method

The method is based on these three assumptions:

I. The design of an interface is highly dependent on the way the user will interact with the system. **Accordingly, controls should be designed before**
the interface.

II. To facilitate a comparison between high and low density displays on basis of two usability aspects (safety and user preferences) all other aspects of the displays have to be usable as well. The reason for this is that a usability comparison between two displays with bad usability is not relevant and for that reason the results cannot be generalised to other usable systems. Consequently, in order to compare high or low density, two usable prototypes should be designed.

III. In order to reach a conclusion regarding screen density based on differences that occur in the comparison between high and low density displays, the only thing that varies between the two displays can be screen density. If other aspects besides screen density vary, differences in the comparison are not certainly due to screen density. Hence, all aspects (except screen density level) of the two prototypes should be equal.

With these assumptions set, the following line of work was carried out in order to compare a high density display with a low density display:

- Literature studies
- Contextual research with heuristic evaluation
- Design of controls
- Design of display
- Lane Change Test
- Qualitative subjective questionnaire

1.3 Delimitations

This thesis investigates an 8 inch colour display with the proportions 9:16, placed on top of the center stack. The display is integrated and this thesis does not look into the advantages and disadvantages of an integrated display compared to several displays.

In the design of the display interfaces, consideration is taken neither to resolution and brightness of the display nor the colour and contrast of the layout. The focus of the interface design is grouping of information, and legibility.

The interaction with the display is done through permanent controls that are reachable to the driver and placed in the center stack. Hence, this thesis will
not evaluate or give recommendations about controls on the steering wheel. The reason for this is that Volvo has decided that controls on the steering wheel should be redundant to controls on the center stack.

This thesis does not investigate interaction by remote control, touchpad, touch screen, or voice. These delimitations were proposed by Volvo. Furthermore, the design of controls will not consider destination entry. The reason for this is that the task of entering text into a system while driving is extremely advanced. To find an input method that is usable could therefore be considered as a separate thesis and hence to immense to accomplish within this study.

1.4 Targeted Readers
This thesis is primary written for HMI-students, e.g. students in cognitive science, interaction design or computer system developing. The secondary readers are people working with in-vehicle HMI at Volvo Car Corporation.

1.5 References
The majority of the references used in this thesis are related to in-vehicle systems, literature from other domains have not been studied to any greater extent. However, literature on screen density has been studied thoroughly regardless of domain. For example, screen density experiments on clinical nurses (Staggers, 1993), learners (Morrison et al, 1989), and a simulated power plant (Burns, 2000).

1.6 Thesis Overview
Chapter 2 gives a short background of Volvo for those who are interested in the company. Chapter 3 describes the theoretical framework of the thesis, where screen density research, as well as related research on HMI and Human Factors, is covered. The method is described in chapter 4, and the results in chapter 5. In chapter 6 the method and the received results are discussed, followed by recommendations for further work and conclusions in chapter 7.
2 VOLVO

This chapter describes the background of this thesis and gives a concise description of Volvo. The references are Volvo Cars’ homepage (2005), Volvo Group’s homepage (2005), and 2004 Pocket guide (2004).

2.1 Background to Thesis

This thesis is written in collaboration with the Human Factors Engineering and Ergonomics department at Volvo Car Corporation in Torslanda, Gothenburg. The reason for this collaboration is that there is a current trend in the car industry to integrate several functions, that earlier had one display each, into one single display. When this is done a large amount of information has to share the same space, and therefore Volvos assignment for us was to investigate how this could be done in the best possible way, primarily to see how different amounts of information affect safety. While performing introductory literature studies on menu breadth/depth and structuring of information, screen density arose as the most suitable area to investigate. Since earlier research on this topic never reached any clear conclusions it made it even more interesting to study.

Only one earlier study (Somervell et al., 2002) has investigated screen density on a secondary task. This is comparable to investigating screen density on an in-vehicle display while the subject is driving. However, Somervell’s study measured performance on the secondary task. To investigate safety in a car, performance of the preliminary task (driving) has to be measured. A study on this would therefore be research not yet done. For this reason, a study on screen density would not only serve Volvo’s interests but also be motivated from an academic point of view.

2.2 Brief History of Volvo

The brand Volvo was born in 1927 when the first series-built car left the company’s works in Gothenburg. The first model was called Jakob (see Figure 2.1) and from this point forward models like PV, Amazon, Volvo 240, and Volvo XC90 (see Figure 2.2) have been seen on the roads.
Assar Gabrielsson and Gustaf Larsson are the two men behind Volvo. Both worked at SKF in Gothenburg and in 1924 they started their own corporation. SKF helped with both financing and the name “Volvo AB”. Volvo is Latin and means “I roll”. Sales figures for the first model Jakob were not very impressive during the first year, only 297 cars were sold. The model evolved and two years later 1383 cars were sold and among these 27 cars were exported. Last year 466 036 Volvo cars were made and today there are 5 million Volvo car owners around the world. The corporation has grown but the core values have stayed the same during the development. Gabrielsson and Larsson wanted the guiding principle behind everything made at Volvo to be safety.

2.3 Volvo Today
In 1999, Ford Motor Company, the world’s second largest car maker, bought Volvo Car Corporation. Since then, Volvo Car Corporation is a subsidiary to Ford Motor Company, and is a part of a division called Premier Automotive Group (PAG). The quality brands within this group are Jaguar, Land Rover, Aston Martin and Volvo. Within this group, Volvo is “Centre of Excellence for Telematics”, and within all Ford owned cars Volvo is “Centre of Excellence for Safety”. This means that Volvo Car Corporation’s research within advanced safety solutions has a strong influence to all brands within Ford Motor Company.
Volvo Group is still a Swedish owned company with the departments; trucks, aero and marine. However, the Volvo trademark is owned by both Ford Motor Company and Volvo Group through the company Volvo Trademark Holding AB.

Volvo Car Corporation still develops and manufactures its cars and it is at Volvo’s headquarters in Gothenburg, Sweden, that most the models are produced. Out of Volvo Cars 27000 employees worldwide 20000 are based in Sweden where five of Volvos six factories are.

2.4 Volvo Customers
The four largest markets are the US, Sweden, Great Britain and Germany. Out of the 415,046 Volvo cars sold in 2003; 134,602 were sold in the USA, 47,928 in Sweden, 39,135 in Great Britain, and 30,285 in Germany.

Volvo does work with usability on the basis of a typical user. They have a ruff description on the average user, who is a 46 years old male of unknown nationality, but they do not design the cars with this user in mind. Instead they have a range of products which are specified for different marked areas. The Volvo Intelligence Centre decides which customers the cars should be designed for, and evaluates what these customers like.\(^1\)

---

1 Lena Persson, Departement of Product Planning, Volvo Cars Corporation, Torslanda
2.5 Core Values

When Assar Gabrielsson and Gustaf Larsson designed and manufactured the first ever Volvo car they strived to design it with the human being in center. This vision has evolved in Volvos core values today: safety, environment and quality. Volvos vision is to be the worlds most wanted and most successful high class car brand. Their mission is to create the safest and most exciting car experience for modern families. This mission contains the essence of Volvos historical heritage: *Volvo. for life.*
3 THEORETICAL FRAMEWORK

During the last few years the focus when developing in-vehicle information systems has shifted from getting the technology to work, to the interaction between driver-system and the real needs of the drivers (Stevens, 2000). Many researchers welcome the shift in focus. Tsimhoni & Green (2001) argue that the continued growth of in-vehicle systems depends upon successfully addressing concerns of safety and usability. Nilsson et al. (1998) state that the goal with new in-vehicle systems ought to be to support the drivers in their driving. In order to do this, they argue that the graphical interface has to be designed on basis of human capacity and resources for collecting information. If this is not done, the new technology might distract the drivers instead of supporting them.

To address concerns about safety and usability, information about human factors and usability is needed. To design a system on basis of human capacity and resources, knowledge about cognitive psychology is needed. The aim of this chapter is to give a theoretical framework that can be used as a starting point in the design of an in-vehicle information system. 3.1 Cognitive psychology explains theories of how the human mind works, and 3.2 Human Factors describes the limitations of the human mind due the in-vehicle environment. 3.3 Usability explains how an in-vehicle system is made usable, and gives various guidelines to control and interface design. The last section in this chapter, 3.4 Screen Density, presents other research made in the area of screen density.

3.1 Cognitive Psychology

“cognitive psychology deals with how people perceive, learn, remember, and think about information.”

(p. 2, Sternberg, 1999)

This section briefly describes human cognition, which is the area that aims at describing how the mind works. Due to the complexity of the human mind it is neither possible nor desirable to explain all parts of human cognition within the frames of this thesis. For that reason, focus is placed on the areas of
importance to the process designing in-vehicle information systems such as attention, perception, and memory.

### 3.1.1 Attention
Humans always have an enormous amount of information available through their senses, memories and cognitive processes. However, there are limits in the ability to process this information and therefore the most interesting piece of information has to be selected. When actively processing the selected piece, attention is directed to that piece and the rest of the available information is dimmed out. This is called *selective attention* and it is used when focusing on some stimuli and ignoring other (Sternberg, 1999). When interacting with an in-vehicle information system while driving, the driver has to pay attention to both system and driving. This is called *divided attention* (Sternberg, 1999). However, when the total demand of both tasks become too excessive, one of the tasks must suffer (Wickens, 2000). That is to say, the driver has to select which task to focus on. An arising problem in this situation is that humans sometimes select the inappropriate aspects of the environment to process (Wickens, 2000). If a driver pays too much attention on the system and therefore too little on driving, the driver could be an increased safety risk – both to himself/herself and others (Barrow, 1991).

Tijerina (1999) categorizes driver distraction in three groups; general withdrawal of attention, selective withdrawal of attention, and biomechanical interference. General withdrawal of attention happens when the driver is tired (eyelid closure) or inattentive (glances away from the road). This is noticeable by degraded vehicle control and degraded object and event detection. Selective withdrawal of attention is caused by attention to thoughts, and decision making based on expectations instead of factual traffic situation. When this happens the vehicle controls stays unaffected but the driver’s ability to detect objects and events is degraded. Biomechanical interference refers to the driver changing position from the neutral seated position. When the driver leans over to manipulate a system with one hand, distraction might be caused by the change in position, and the occupation of the hand may degrade the ability to execute manoeuvres.

When comparing visual output to auditive output from system to driver, a key aspect that benefits a visual output system is that the driver can choose when to look at the display (Burnett, 2000). Presumably, the driver will look at the output when the traffic situation is calm instead of the system, unaware of the traffic situation, forcing output on the driver.
3.1.2 Perception

Humans constantly receive sensations from the environment and these sensations need to be recognised, organised, and interpreted. All these cognitive processes are together called perception (Sternberg, 1999).

The visual sensations are of great importance while driving a car. Many traffic accidents happen because the driver falls short in visual perception (Nilsson et al., 1998). Humans tend to perceive visual information in a way that most simply organise it. That is to say, we form coherent wholes out of separate parts even if the parts are disparate. This phenomenon is described by the Gestalt law of Prägnanz (Sternberg, 1999). Here are brief descriptions of some of the Gestalt principles that all support Gestalt law of Prägnanz:

- **Figure-ground** – When looking at a visual field, some objects seem highlighted (the figures) against others, which are fused into the background. The picture below, Figure 3.1.2a, can both be seen as a vase or two faces from the side.

  ![Figure-ground](image)

  **Figure 3.1.2a: Figure-ground**

- **Proximity** – Objects that are close together seem to form a group (see Figure 3.1.2b).

  ![Proximity](image)

  **Figure 3.1.2b: Proximity**

- **Similarity** – Objects that are similar seem to form a group (see Figure 3.1.2c).
• **Continuity** – Disrupted or discontinuous forms seem to have continuous forms (see Figure 3.1.2d).

\[
\begin{align*}
\text{=} & \quad \text{+} \\
\therefore & \quad \therefore
\end{align*}
\]

**Figure 3.1.2d: Continuity**

• **Closure** – Uncompleted forms seem complete (see Figure 3.1.2e).

\[
\begin{align*}
\text{[} & \quad \text{[} \\
\text{]} & \quad \text{]} \\
\text{[} & \quad \text{[} \\
\text{]} & \quad \text{]}
\end{align*}
\]

**Figure 3.1.2e: Closure**

• **Symmetry** – A collection of objects seem to be forming mirror images around a central point. In Figure 3.1.2f the brackets are not perceived as four separate objects but as two sets of brackets around a central axis.

\[
\langle ( \ ) \rangle
\]

**Figure 3.1.2f: Symmetry**

The Gestalt principles are useful for understanding how humans perceive groups of objects to form integrated wholes (Sternberg, 1999).

### 3.1.3 Memory

Traditionally memory has been described to consist of two parts; short- and long-term memory. The short-term memory holds memories for seconds up to a couple of minutes. Long-term memory holds memories for hours, years, and in some cases for life (Sternberg, 1999). Miller (1956) found that the limitation of our short-term memory is 7 chunks, plus or minus 2. A chunk
can be a single item or several items put together to form a meaningful whole, e.g. a telephone number can be remembered as eight chunks “5 5 6 5 9 9 7 0” or as four chunks “55 65 99 70”. Other researchers have in addition to this found that any delay or interference can cause the 7 chunk capacity to drop to 3 (Sternberg, 1999).

Baddeley, among others, describes memory from a different perspective (Sternberg, 1999). In this perspective, working memory plays a significant part. Working memory is part of long-term memory, and short-term memory is part of working memory. Working memory holds the recently activated parts of long-term memory and moves these parts in and out of short-term memory. Baddeley suggests that the working memory contains a central executive, a phonological loop, and a visuospatial sketchpad (see Figure 3.1.3). These parts have limitations in the amount of information they can process. The central executive co-ordinates attentional activities and governs response (Sternberg, 1999). The phonological loop comprises a temporary phonological store in which auditory memory can be held over a period of a few seconds. (Baddeley, 2000). The visuospatial sketchpad has the capacity to store a single complex pattern (e.g. a detailed picture) but has clear limitations when it comes to remembering serial patterns (e.g. several pictures in a row) even if the serial patterns are simple (Baddeley, 2000).

![Diagram of Baddeley's working memory](image)

**Figure 3.1.3: Simplified picture of Baddeley's working memory**

14
3.2 Human Factors

“Designing machines to accommodate the limits of the human user is the concern of the field called human factors.”

(p. 2, Wickens & Hollands, 2000)

In-vehicle information systems bring a new dimension to the car environment (Roessger, 2000). Today a driver’s attention is no longer only focused on the road ahead and interacting with steering wheel, pedals and gear lever, attention is also on the inside of the car when the driver is interacting with the information system. One problem with this is that drivers do not fully compensate for the demands of the in-vehicle information system by maintaining an appropriate level of driving safety (Tsimoni & Green, 2001). Even though this is the case, the technological development in this area is hard to stop, so the ergonomic community has to escort the development (Roessger, 2000). The ergonomic community has to make sure that the cost of using the systems is not higher than the benefit (Roessger, 2000). That is to say, the mental capacities and amount of attention the driver has to spend to get information from the system has to be lower than the benefits of the retrieved information. From a human factors perspective, safety must always be of primary importance when designing in-vehicle information systems (Burnett, 2000). Nilsson et al. (1998) states that the amount of information as well as timing, placement, and presentation are crucial for providing best possible conditions for the driver.

3.2.1 Visual Demands of Driving

The task of driving is highly dependent upon visual attention (Wollter & Törnqvist’s, 2002). Consequently, if the driver misinterprets visual cues, or misses important visual stimuli it can have devastating consequences. It is also established that many car accidents are caused by visual errors (Nilsson et al., 1998). Due to this fact, several studies have been done to establish the visual demands of driving, and how driving is affected when the driver performs different tasks within the car in addition to driving. There is evidence that the probability to deport from a lane while driving increases when the driver is studying detailed maps (Tsimhoni & Green, 2001) or writing text into a navigation system (Green, 2002). Green (2002) also reports on studies showing that the number of accidents increases when the driver is using a mobile phone or other systems with high visual demands.
Roessger (2000) states that the reduction of visual load has to be the major focus when designing interfaces for in-vehicle information systems. Designers has to be held responsible since the drivers themselves not fully compensates for the demands of an in-vehicle information system (Tsimhoni & Green, 2001). Designers have to be aware that bringing an information system into a car dramatically changes the consequences of small differences in time and accuracy of visual search of the interface (Everett & Byrne, 2004).

3.2.2 Mental Workload
According to Nilsson et al. (1998) the risk for mental overload is already obvious in the traffic environment, and this risk will increase with in-vehicle information systems. In fact, several studies have showed that driver workload is affected by the addition of a display in the vehicle (Barrow, 1991). Workload is a person’s experienced load of demands, and is task-specific and person-specific (Rouse et al. in de Waard, 1996). The experienced load is affected by the driver’s capabilities, motivation, strategies, mood and state (de Waard, 1996).

Lack of attention to the road and distraction are both major contributing factors in many road accidents (Burnett, 2000). According to Harbluk (2002) it contributes to over 20% of motor vehicle crashes. The lack of attention and distraction are not only due to the driver looking somewhere else but also depends on the driver’s mind being on something else. As John Lee expressed, quoted in ElBoghady (p. 2, 2000) “the fact that you’re looking at the road doesn’t mean that you’re thinking of the road”. Harbluk (2002) agrees with this fact and declares that the distraction using in-vehicle systems “may arise not from the manual manipulation of these devices, but rather the cognitive consequences of their use”. When a driver performs tasks with increased cognitive load it results in changes in the drivers’ visual behaviour, vehicle control and subjective assessments of workload, safety and distraction (Harbluk, 2002).

When workload becomes too high, drivers adapt task management strategies (Wickens & Hollands, 2000). In an article Jansen (2000) is commenting on the European Commission’s statement of principles on human/machine interface. He is pointing out that adaptation of driver behaviour to a new situation on a strategic level could result in risk compensation. Therefore, adaptation of driver behaviour should be one of the central themes when designing HMI for in-vehicle use. Four types of adaptation are possible (Wickens & Hollands, 2000):
• Allowing performance of tasks to degrade
• Performing tasks in a more efficient way, sometimes at expense of accuracy
• Shedding tasks of lower priority in order to operate in a more optimal fashion
• Shedding tasks that should be performed, resulting in nonoptimal operation of tasks

3.2.3 Measuring Driver Distraction

Since safety is the most important issue when designing HMI for in-vehicle use it is also important to have a method for measuring the amount of distraction the system is imposing on the driver’s attention. Several methods for the measuring of how much a system is distracting the driving are used in related research. In this section the methods 15-second Rule, Occlusion Method, and Lane Change Test will be described briefly.

The Society of Automotive Engineers, ITS Safety, and Human Factors Committee develops standards to ”assure product safety and usability, both by design and assessment” (p. 1, Green, 1999). The committee has put together standards for what drivers should be allowed to do, and recommendations for procedures to calculate if a task complies with the 15-second rule. The 15-second rule states that any task in the system should not take more than 15 seconds to complete when measured as a continuous task. Timing starts when the driver’s hand leaves the steering wheel or moves towards the device and ends when feedback is provided for the last-switch actuation (Green, 1999).

Opposed to the 15-second rule, the Occlusion method is based on the idea that the completion of each portion of a task should not take more than 1-2 seconds to complete. It is argued that only considering the total completion time does not take the portioning of the task into consideration and hence does not simulate a real traffic situation. Instead, the Occlusion method simulates a situation where the driver has to shift visual attention between car periphery and the in-vehicle system. This is done by having the subject look at the system for 1-2 seconds and then blocking the subject’s sight for about 3 seconds (Bauman et al., 2004).

The Lane Change Test, shortened LCT, is based on the notion that driver distraction is best measured by calculating the distraction on a primary task while performing a secondary task. It is the distraction the secondary task is imposing on the primary task that is relevant, not the time it takes to complete
the secondary task or a portion of a the secondary task. The LCT measures the difference in driving performance between driving a car in a simulator, and doing this while performing a secondary task, e.g. navigating in a menu structure on a display in the center stack. (ISO proposal for LCT, 2005)

3.3 Usability

"Usability: the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."

ISO 9241-11 (p. 537, Bevan, 2001)

According to Burnett (2000) usability is one of the most important aspects of in-vehicle information system design. In order to achieve usability both ergonomic and cognitive aspects must be considered. Usability, both ergonomic and cognitive, is achieved by doing the following (Kunimitsu et al., 1999):

- Supply information in an integrated, related and easily understandable form.
- Optimize and simplify the input method (operation flow) according to quantity, quality and urgency of information.
- Make all operations consistent to reduce the burden of memorisation. By doing this manual-free operation becomes possible.

Murphy (2001) believes that there are two approaches to usability: by evaluation or by principles. In order to guide a design that will result in a system with the qualities listed above, section 3.1 Usability describes a method for evaluation as well as numerous principles for designing usable controls and interfaces.

3.3.1 Usability Heuristics

Jacob Nielsen is well known for his heuristics which are used to evaluate the usability of systems. They are called "heuristics" because their nature is more like rules of thumb than specific usability guidelines (Nielsen, 1993). The heuristics are briefly described below.

- **Visibility of system status** – The system should always keep users informed of what is going on, through appropriate feedback within reasonable time.
• **Match between system and the real world** – The system should speak the users’ language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

• **User control and freedom** – Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

• **Consistency and standards** – Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

• **Error prevention** – The system should be carefully designed to prevent problems from occurring in the first place.

• **Recognition rather than recall** – Make objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

• **Flexibility and efficiency of use** – Accelerators, unseen by the novice user, may often speed up the interaction for the expert user such that the system can be used by both inexperienced and experienced users. Allow users to tailor frequent actions.

• **Aesthetic and minimalist design** – Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

• **Help users recognize, diagnose, and recover from errors** – Error messages should be expressed in plain language not using codes, precisely indicate the problem, and constructively suggest a solution.

• **Help and documentation** – Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user’s task, list concrete steps to be carried out, and not be too large.
Nielsen’s (1993) purpose with the heuristics is to use them as a systematic approach to user interfaces and finding its usability problems.

3.3.2 Design Principles
Design principles are based on theory, knowledge, experience and common sense and are used by designers to improve the usability of a design. The best know design principles are written by Norman (Preece et al., 2002). The principles are described in short below (Norman, 2002):

- **Visibility** – Make functions noticeable to the user.
- **Feedback** – Let the user know what has been accomplished by sending information back to the user.
- **Constraints** – Reduce the possibility for the user to make a mistake by restricting the available options at a given point.
- **Mapping** – Provide a logical relationship between controls and the effect they impose on the world.
- **Consistency** – Give the user similar operations and similar elements to perform similar tasks through out the system.
- **Affordance** – Provide the system with attributes that let the user know how to use the system.

3.3.3 Guidelines for Usable Controls
The Commission of the European Communities (2000) gives recommendations on designing in-vehicle systems with retained safety. The recommendations concern parts of the system being involved in the interaction between driver and system. Below is an extract of principles concerning controls.

- **The system should be designed in such a way so that the allocation of driver attention to the system displays or controls remain compatible with the attentional demands of the driving situation.** (principle 2.1.2)
- **The driver should always be able to keep at least one hand on the steering wheel while interacting with the system.** (principle 2.4.1)
- **System controls should be designed such that they can be operated without adverse impact on the primary driving task.** (principle 2.4.4)

When it comes to arrangement of controls it is important to group and locate them logically (Holtelius, 2002; Roessger, 2000; Volvo Cars Dept 93202,
Holtelius (2002) recommends that the functions should be arranged with consideration to users’ goals and not technical structure. Roessger (2000) agrees with this and also extends the recommendation to include all input devices. He suggests that colours can be used to group information and to emphasize the most important information.

It is important for the driver to be able to identify different controls in the in-vehicle environment. This can be done through different coding methods, e.g. location, shape, texture, size, mode of operation, labels and colour (Andersson & Svensson). Roessger (2000) states that aural input is to prefer. If this is not possible, haptic feedback should be provided by using different switches or knobs (Roessger, 2000; Andersson & Svensson, 2002). According to Roessger, tactile feedback reduces the visual load. Sound feedback should not be used due to relatively high sound level in a car (Andersson & Svensson, 2002).

The design of the individual controls should be intuitive. E.g., clockwise turns, moving from left to right, and upwards motion should lead to an increase (Volvo Cars Dept 93202, 2002). Nilsson (2005) also found users believe that + and - are good cues for increase/decrease. Another example is that a button pressed in/out (Andersson & Svensson, 2002), a switch up/down and the symbol ‘1 0’ should be used for the operation turn on/off (Nilsson, 2005).

Kunimitsu et al. (1999) tested best placement of screen, best placement of switch, and best switch. They found that the best placement of screen is on the upper surface of the dashboard (forward in the vehicle). With the screen located as described, the switch should be placed where the conventional audio instrument and control panel are installed. With the screen and switch located as described, the best type of switch is the seesaw

![Seesaw switch](image)

**Figure 3.3.3: Seesaw switch**

Holtelius (2002) has made a study that compares four different center stack controls. The comparison is based on user reactions on understanding of function layout and general appearance. The user reactions were collected through a focus group consisting of 8 potential Volvo S80 buyers living in southern California. The subjects were between 41 and 65 years old. During the focus group session the subjects had the control plates available as paper prototypes. The results from the study shows that users prefer big buttons and clear graphics. Further more, the number of controls is not an issue for users,
however, grouping of controls is. Separating into smaller function groups and logical grouping helps the user to remember where to locate the wanted function. When using a multifunction knob, users accept toggling to reach sound controls e.g. find fader, balance, treble and bass. On the other hand, direct access to AM, FM and CD is wanted.

As far as the look and feel design of the controls goes, Holtelius (2002) states that they should be big with clear graphics. According to Andersson and Svensson (2002) coding of controls can be done by location, shape, texture, size, mode of operation, labels, and colour. If a rotary knob is equipped with a push function, hints that the function exists should be provided. Preferred size of a button is 20 x 15mm, and preferred diameter of a rotary knob is 20 mm.

3.3.4 Guidelines for Usable Interfaces

The design of the interface is an emerging key area that has significant impact on the overall usability and safety of in-vehicle systems (Stevens, 2000). Important aspects of the design include minimizing the distraction potential of the interfaces, and managing the information flow to the driver (Tsimhoni & Green, 2001). Below is an extract of the recommendations for interface design provided by Commission of the European Communities (2000).

- *The system should be designed to support the driver and should not give rise to potentially hazardous behaviour by the driver or other road users.* (principle 2.1.1)
- *Visually displayed information should be such that the driver can assimilate it with a few glances which are brief enough not to adversely affect driving.* (principle 2.3.1)
- *The driver should be able to control the pace of the interaction with the system.* (principle 2.4.5)

According to Murphy’s (2001) general advice for graphical user interfaces, the interface should display only the required information and other information should be hidden until specifically requested. This is strengthened by the design principles frequently used in University of Michigan Transportation Research Institutes’ interface design guidelines (Green, 1996). These design guidelines state that the interface should be kept simple and that the design should be consistent. The importance of consistency is also pointed out by Roessger (2000) as an important design feature. She claims that users get irritated when functions change from one subsystem to another.
Users tend to prefer information provided by a in-vehicle system to be limited, clear and well structured (Ek & Myhrman, 2004). One way to achieve this when designing a graphical user interface with a considerable amount of information is to organise the information in a menu. There has been a lot of research on how a menu is best structured. Miller (in Larson & Czerwinski, 1998) found that menus structured $8^2$ is better than $2^6$, $4^3$ or $64^1$. This is to say breadth 8 and depth 2 is preferable for a menu. Tolle et al. (1987) made similar findings. They found that $8^2$ is better than both $2^8$ and $64^1$.

Items in a menu can be sorted or unsorted. The items can be sorted by alphabet, function or frequency. McDonald et al. (1988) have found studies that show that users have trouble understanding menus sorted by frequency. Card (1982) has found that if the user knows the exact word to search for, alphabetical sorting of menu items is to prefer. Papa and Cooke (1997) state that when the user do not have the exact word, alphabetical sorting is not helpful since the user is looking for a semantic meaning of a word and not for the syntactic meaning. Kurtenbach and Buxton (1994) have discovered that sorted menus are faster than unsorted only when the user is novice. When the user becomes familiar with the system, it does not matter whether or not the menu is sorted.

There are many different guidelines concerning graphical user interfaces, in short GUI. Stevens et al. (2002), recommends a font without serifs. They also state that a mixture of upper and lower case letters is to prefer. The NASA-STD-3000 man-systems integration standards (in Andre et al, 1998) recommends that the font should include lower case characters and allow for descendents. Which size the letters should have depends on the distance between eye and screen. According to Smith’s (1979) James Bond rule the letters should be at least 5.95 mm high when the distance between eye and screen is 850 mm. The rule is the following: 0.007 radians multiplied with the distance between eye and display in mm gives the height of the displayed letter in mm. The relationship between thickness of letter and character height should be from 1:6.25 to 1:12.5. The relationship between broadness of letter and character height should be 0.6:1 to 0.8:1. The distance between letters should at least be as broad as the letters are thick (Stevens et al., 2002). Size of font should according to Ford’s internal standard regulations be at least 3.0 mm for letters and 6.0 mm for symbols, and according to ISO 15008 the smallest letter should at a distance of 850 mm be at least 5.93 mm (Stevens et al., 2002).
Colours should be used to group information and to emphasise the most important information (Roessger, 2000). Colour combination of red and green should be avoided, and blue and yellow be encouraged. The reason for this being that colour blind people have trouble distinguishing green from red, but almost everyone can distinguish blue from yellow (Ware, 2004).

3.4 Screen Density

Screen density is the amount of information present on a screen (Woodruff et al. 1999). Woodruff et al. (1998) states that screen density should be taken into account when designing applications to avoid the risk of affecting the user’s navigation in the system in undesirable ways. What amount of information that gives the most usable system is an important question. High density may lead to a clutter where certain information items are occluded by other information. Low density, on the other hand, may lead to inefficient use of available screen space (Woodruff et al. 1999). Most research on the area try to determine which density levels that allows the fastest, most accurate information detection performance when the information detection is the primary task (Staggers, 1993; Cohen & Ivory, 1991; Burns, 2000). Some researchers examine user preferences for varying screen densities in primary tasks (Morrisson et al., 1989; Ross et al., 1994). Unfortunately, only a limited amount of research is available on density level when the screen is for a secondary task (Somervell et al, 2002).

During the 40 years of studies in the field of screen density, experiments can be found under many different terms. Some examples of different terms are information density, overall density, information packing, information quantity and screen complexity (Staggers, 1993).

Not only can the research on screens density be found under different terms, the term screen density is defined in several different ways. One way is to define it as closeness of objects on a screen. Assuming a finite number of objects on a screen, the density is low when the objects are spread apart and the density is high when the objects are close together (Burns, 2000; Cohen & Ivry, 1991). From this point, this type of density will be referred to as distribution. High distribution is when the objects are spread apart and low distribution when they are close together. Another way to define density is the amount of objects on a screen; low density is few objects and high density is many objects (Woodruff et al., 1998; Somervell et al., 2002). This can be expressed as the proportion of the screen displaying information (Everett &
Byrne, 2004; Morrison et al., 1989; Staggers, 1993; Stevens et al., 1994). See Table 3.4a for an overview of how the terms will be used in the thesis.

**Table 3.4a: How the terms *distribution* and *density* will be used in this thesis.**

<table>
<thead>
<tr>
<th></th>
<th>Definition</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Distribution</strong></td>
<td>The spreading of objects</td>
<td>Spread apart</td>
<td>Close together</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>The amount of objects</td>
<td>Few</td>
<td>Many</td>
</tr>
</tbody>
</table>

When screen density is defined as the proportion of a screen displaying information, density level is expressed as a percentage of available character spaces that are in use. This measurement is most suitable for character-mode screens (Tullis, 1997). Variations on this definition include whether or not to include graphics or highlight characters in the calculation (Staggers, 1993). In Staggers study (1993) screen density refers to total number of characters present on a screen, including pertinent and graphics characters. According to Staggers, high density is 58.5%, moderate 33-41% and low 27-32%. Morrison (1989) label screens with less than 25% density as low, 26-50% as medium, and above 50% as high.

Studies have been conducted on screen density since the beginning of the 1960s (Staggers, 1993). These studies have not yet given a clear-cut answer to which level of density is preferable. Table 3.4b show the different results of research on the area.

**Table 3.4b: Research on screen density.**

<table>
<thead>
<tr>
<th>Study</th>
<th>Measured</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ross et al. (1994)</td>
<td>Preferences for learning from computer-based instructions.</td>
<td>When the text is realistic subjects prefer medium to high density.</td>
</tr>
<tr>
<td>Cohen &amp; Ivry (1991)</td>
<td>Reaction time for detecting presence of a target</td>
<td>Subjects are faster with low density and high distribution.</td>
</tr>
<tr>
<td>Somervell et al. (2002)</td>
<td>Time for locating where a target is located on a screen while busy with a primary task on another screen</td>
<td>Subjects are as good or better with low density. (low=20 items, high=320)</td>
</tr>
<tr>
<td>Study</td>
<td>Measures</td>
<td>Findings</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Staggers (1993)</td>
<td>Task time, accuracy, and subject satisfaction for finding information targets.</td>
<td>Information targets are found faster on high density screens. Users were equally satisfied with either type of density.</td>
</tr>
<tr>
<td>Everett &amp; Byrne (2004)</td>
<td>Visual search times for varying icon spacing (distribution)</td>
<td>Low distribution lead to faster search times.</td>
</tr>
<tr>
<td>Morrison et al (1989)</td>
<td>User preferences</td>
<td>Users prefer medium to high density</td>
</tr>
</tbody>
</table>

To summarise the table, users prefer medium to high density (Morrison et al 1989; Ross et al., 1994), or are equally satisfied with either type of density (Staggers, 1993). When the users are tested they are fastest with low density in some studies (Cohen & Ivry, 1991; Somervell et al., 2002), and fastest with high density in some studies (Staggers, 1993). When it comes to distribution, some research show that high distribution give fastest search times (Cohen & Ivry, 1991; Burns, 2000), and some research show the opposite (Everett & Byrne, 2004).
4 METHOD

The following chapter describes the development process towards a screen density experiment, the methods used in each stage of the process, and the method of the screen density experiment. During the process, decisions concerning line of work and choice of method were guided by three assumptions, described in detail in 1.2 Method. In short, the assumptions are:

I. Controls should be designed before the interface.

II. In order to compare high and low density, two usable prototypes should be designed.

III. All aspects (except screen density level) of the two prototypes should be equal.

Consequently, in order to compare high with low screen density in the experiment, the controls were designed before the interface. Furthermore, the interface was designed to be usable and to enable the creation of two equal prototypes, except for screen density level.

The overall development process freely followed Shneiderman & Plaisant’s (2005) three step development plan for creating a successful user interface. In this study, the first step included benchmarking, theoretical research, and development of guidelines. During the second step, controls and interface were designed and implemented. In the third and final step of the process, the two prototypes were tested for usability.

4.1 Heuristic Evaluation

The starting point of the development process was a heuristic evaluation of in-vehicle information technology in Volvo cars and competitor brands. This benchmarking was done to explore the in-vehicle information technology already available on the market. The evaluation was carried out according to Nielsen’s (1993) method for heuristic evaluation. With this method, ten general principles for user preferences (described in detail in 3.3.1 Usability Heuristics) were used to identify good and bad aspects of user interfaces (Nielsen, 1993). In this survey, some of the principles were modified to suit in-vehicle information systems. Besides being used in this evaluation, the
principles were used when designing controls and user interface for the screen density experiment. The modified principles are listed below:

- **Feedback** – The system should through appropriate feedback within reasonable time keep the user/driver informed about what is going on.

- **Speak the user's language** – The system should speak the user’s language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Following real-world conventions and making information appear in a natural and logical order.

- **Clear marked exits** – The system should have clearly marked "emergency" exits. When a user by mistake chooses a function by mistake, the possibility to leave an unwanted state without having to go through extended dialogue should exist.

- **Consistency and standards** – The system should follow a platform of conventions. The user should not have to wonder whether or not different words, situations, or actions mean the same thing.

- **Error prevention** – The system should have a careful design that prevents a problem from occurring in the first place.

- **Recognition rather than recall** – Objects, actions, and options should be made visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

- **Shortcuts** – The system should allow users to tailor frequent actions. Shortcuts, unseen by novice user, can speed up the interaction for expert user, this way the system can cater to both inexperienced and experienced users.

- **Simple and natural dialog** – Dialogues should not contain information that is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

- **Provide good error messages** – Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
• **Help and documentation** – Although a system that is easy enough to understand without reading the documentation for help and pointers is to prefer this may still be necessary to provide. The information should be easy to search, focused on the user’s task, list concrete steps to be carried out, and not be too large.

The evaluation was carried out by the authors of this thesis. Nielsen (1993) recommends 3-5 evaluators for optimal weight in cost-benefit, but due to lack of resources, only two evaluators were available in this study. A table with the principles listed in the left most column and the car models listed in the upper row (see Appendix A, in Swedish) was used. During the evaluation, each of us separately noted comments concerning the system in the table. Pictures of each display and controls were also taken to serve as inspirations in the future design process. The following four cars were evaluated:

- Toyota RAV 4
- Saab 9-3 Sedan R4
- Mercedes Benz E-270
- Audi A8

After collecting comments about the different cars, the comments were compiled to a list of usability concerns which is presented in 5.1 **Heuristic Evaluation**.

### 4.2 Controls

The emphasis of this thesis is the screen density experiment but since the system as a whole had to be developed the controls had to be designed first according to the first assumption (see 1.2 **Method**). Before designing the controls, good and bad aspects of controls for in-vehicle systems were identified. To start with, four people\(^2\) with close connection to HMI in cars were interviewed; unfortunately it turned out that very limited research had been done on controls in vehicles. Therefore, literature about controls in other domains was studied, and benchmarking was done by looking at PDAs and mobile phones. Unfortunately, today’s PDAs have touch screens, and mobile phones have joy sticks, which neither was relevant to this study. Subsequent to this, the annual report from J.D. Power and Associates (2004)

---

\(^2\) This was Torbjörn Alm, instructor in industrial ergonomics at Linköpings University, and Azra Moric, Anders Hallén and Johannes Aagard, employees at Volvo Cars working with HMI in vehicles.
was investigated in order to identify control features that customers are satisfied with. The method for this is described in detail in 4.2.1 Report from J.D. Powers and Associates. After studying the J.D. Power and Associates report a control survey with real end-users was carried out. This was done in order to discover what users actually do when they interact with different combinations of menus and controls, and to determine user preferences for the different combinations. The method for this is described in detail in 4.2.2 Control Survey. The results from studying the report and the testing real end-users, combined with literature studies (see 3.3.2 Guidelines for Usable Controls), led to the design of the center stack prototype described and pictured in 4.2.3 Center Stack Prototype.

4.2.1 Report from J. D. Power and Associates

The 2004 Navigation Usage and Satisfaction Study conducted by J.D. Power and Associates (2004) investigates users’ experiences of navigation systems in their own cars. The report from the study consists of answers to questions concerning navigation systems in 78 different car models. For each model, answers were collected from approximately 100 subjects. In our survey, the J.D. Power and Associates report was used to identify features of good and bad in-vehicle information system controls. Answers to four of the reports questions were compared among the car models represented in the study. The four questions selected were:

- Things gone wrong with button/knob controls
- Satisfaction concerning understanding the controls
- Satisfaction concerning inputting destination
- Satisfaction concerning look and feel of controls

The subjects in the study graded their preferences on a scale from 1 to 10, except for the first question where they counted mistakes that had been made. In the report, these grades were presented both as a mean for each question and as a mean for every car on each question. This made it possible to compare a car’s grade on one specific question to the average of all cars on that question. By doing this comparison, cars that had one unit better or worse than average on the four questions presented above were selected from the total 78 cars. The trend was that cars with high scores in one of the selected questions also had high scores in the other three questions and vice versa for the cars with low scores. (In the first of these four questions, a low score was good, i.e. few things gone wrong, and was therefore equivalent to high scores in the other questions.) 16 car models with good grades and 11
cars with bad grades were selected. Pictures of the selected cars were looked up on the Internet.

According to one of the delimitations stated in 1.3 Delimitations considerations would not be taken to touch screen. Unfortunately, most of the displays in the selected cars had touch screen. Due to this, cars with nothing but touch screen were eliminated but cars with both touch screen and controls were kept. Another limitation was the fact that some of the models did not have any good detailed pictures on the Internet, which resulted in that these cars were eliminated as well. After this process, 9 cars with good grades and 5 cars with bad grades were left. The high grade models were: Acura MDX, Acura TL, Acura TSX, Mazda 3 (although many responders made mistakes on the controls in this model, they were satisfied with the look of them), Mercedes Benz S-class, Lexus GS Series, Lexus GX 470, Lexus RX 330, and Lexus SC 430. The low grade models were: BMW X3, Buick Rendezvous, Cadillac Escalade/ESV, GMC Envoy (which looks like the Envoy XUV, Yukon Denali, and Yukon XL Denali models), and Toyota Avalon.

Norman’s usability principles, described in 3.3.2 Design Principles, were used to subjectively establish the features of each car. We separately wrote down our comments of each model. After this, comments on the good cars were compared in order to establish common features and this was done in the same way with the bad cars. The result was a set of do’s and don’ts, which can be found in 5.2.2 J.D. Power and Associates.

The evaluation made from the results in J.D. Power and Associates was based on a mixture of qualitative and quantitative data. This decision was based on the fact that we only had access to statistical data when choosing the best and worst models regarding controls, and when establishing their core feature. However, this quantitative data was based on subjective opinions (not the first of the four questions evaluated, which was qualitative). Accordingly, the first step was quantitative evaluation on qualitative data. The second step was a qualitative evaluation on the quantitatively selected cars.

![Diagram](Figure 4.2.1: Overview of the steps (arrows) in evaluating the report from J.D. Power and Associates)
4.2.2 Control survey

An experiment was conducted to see how users interact with a menu via a control. The control survey was not designed to make the outcome possible to generalize to all in-vehicle controls and is therefore only valid for this screen density experiment.

The experiment was made to evaluate two things; how subjects use the controls, and which combination of control-menu subjects prefer. The experiment had within-subjects design and two independent variables 1) menu location, and 2) control device. The conditions were horizontal and vertical in the menu location variable, and knob and four-way seesaw in the control device variable. The dependent variables were which interaction most subjects prefer and how subjects interact. The design of the control survey is illustrated in the table below.

<table>
<thead>
<tr>
<th></th>
<th>Knob</th>
<th>Four-way seesaw</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Through literature studies it was established that there are four ways to structure a menu in an in-vehicle display; via pie menu, within a text (a hyper text document), horizontally or vertically (Paap & Cooke, 1997). In order to navigate in pie menus and hyper text documents, interaction has to be done with a cursor or a touch screen. For safety reasons, a cursor was not possible in an in-vehicle system, and touch screen was ruled out based on the delimitations stated in 1.3 Delimitations. Therefore, horizontal and vertical menus were the available options. There were also several ways to interact with a menu in an in-vehicle display; direct buttons, soft keys, four-way seesaw, knob, touch screen, touch pad, and voice (Paap & Cooke, 1997). The last three controls were not possible due to the delimitation stated in 1.3 Delimitations. Neither soft keys nor direct buttons could be used. Soft keys have to be placed in direct access to the display and since the display in this study was placed 85 cm from the driver’s eyes, soft keys would be placed out of the driver’s reach. The placement of the display was a delimitation given by Volvo. Direct buttons to every function in the system was not a possibility.
since there were approximately 300 functions. The four-way seesaw and the knob remained as possible interaction devices.

18 subjects, 9 men and 9 women (age 23-65, average 35 and median 28) were asked to navigate in the horizontal menu with the knob and with the four-way seesaw and again in the vertical menu with both controls. The subjects were asked to picture themselves in a car having the menu in a display placed on top of the center stack and the controls placed beneath in the center stack to the right. The menus and controls are pictured below and the task was to go from Players to Climate in the horizontal menu, and from Navigation to Climate in the vertical menu. The starting points were highlighted in blue. The tasks were given as instructions instead of scenarios, which could have been “You want to change a setting under Climate, what do you do?”

![Figure 4.2.2a: Pictures of the menus; horizontal and vertical](image)

![Figure 4.2.2b: Picture of the controls; knob and four-way seesaw.](image)

The order of the variables was distributed over the subjects in the following order to retain the internal validity:

- horizontal seesaw, horizontal knob, vertical seesaw, vertical knob,
• vertical knob, vertical seesaw, horizontal knob, horizontal seesaw,
• horizontal knob, horizontal seesaw, vertical knob, vertical seesaw,
• vertical seesaw, vertical knob, horizontal seesaw, horizontal knob,
and so on.

The subjects repeated the procedures four times, two times with the horizontal menu and two times with the vertical menu. During the test it was noted how the subjects pushed and turned the controls and afterwards they were asked which combination of control and menu they preferred. The test instructions are attached in Appendix B (in Swedish) and the results can be seen in 5.2.2 Control Survey.

4.2.3 Center stack prototyp

Literature studies, the results from the heuristic evaluation, the evaluation of the J.D. Power and Associates’ report (2004), and the control survey led to the design of the center stack prototype. The design had to be settled at an early stage in order to have the prototype ready in time for the experiment. The prototype consisted of six direct buttons, a four-way seesaw, an OK button and a BACK button. The center stack prototype is pictured in Figure 4.2.3 and explained more in detail in 5.2.3 Center stack prototype.

![Figure 4.2.3: The center stack](image)

4.3 The Display of Screen Density

Contextual design (Preece et al., 2002) has proven to be a good methodology to use when developing a design for a display. In this study, contextual design was used in order to emphasize the user’s (i.e. the driver’s) interaction with the system. The process started with studying existing research made on screen
density, see 3.4 Screen Density, followed by an iterative brainstorming and diverging of optional solutions. This step of the process is described in 4.3.1 Definition of screen density, and resulted in several sketches of design options that needed to be divided into categories to get an overview. The design rationale technique QOC, described in 4.3.2 QOC, was chosen for this purpose. After the QOC, a card sorting test was conducted, see 4.3.3 Card Sorting Test, to structure the functions that were going to be displayed on the screen. How the definition of screen density was applied is described in section 4.3.4 Applying the Definition of Screen Density, and the design process finished with development of the graphical interface of the display described in 4.3.5 Graphical User Interface.

4.3.1 Definition of Screen Density
When related research on the area screen density was studied, it was concluded that there are as many ways to define screen density as there are experiments about it. This was also what Geraci (2002) concluded. A clear definition of screen density had to be established in order to apply the definition to a design and to define the level of density on different screens. The first step was to come up with possible definitions of screen density. Preferably, a whole team of designers should be involved in this step in order to ensure that every single option is externalized, but in this study, there were only two of us. Using earlier studies (see 3.4 Screen Density) and iterative brainstorming, five different definitions were established. These are:

- **Distributed** – The spreading of a set number of objects on a screen. With this definition, low density when the objects are wide apart, and high density is when the objects are close together.

- **Mediated** – The addition of information from other mediums to a finite amount of core information e.g., adding icons to a menu. With this definition, low density is when only the core information is displayed, and high density when the additional information is added.

- **Graphically congregated** – The addition of graphical content to a finite amount of text on one screen e.g., borders are added to menus. With this definition, low density is without graphical content and high density with graphical content.
• **Divided** – The amount of objects on a screen. With a finite amount of information, low density result in a higher number of screens than high density. With this definition, the information is divided into several screens when the density is low and on fewer screens when the density is high.

• **Descriptive** – Clarifying information is added to a text, e.g. ‘tuning’ is clarified with ‘tuning - find radio frequency’. With this definition low density is without clarifying information and high density is with clarifying information.

Lo-Fi examples of these definitions can be found in Appendix C. To select one of the above mentioned definitions the design rationale technique QOC was chosen in order to argument and decide which one to choose.

### 4.3.2 QOC

A design rationale technique is a technique where the ideas, argumentations, alternatives, and reasoning made by the designers are rationalized, i.e. information about the design that is only in the heads of the designers is externalized and structured with this technique. There are a few different design rationale techniques to choose among and the three most common ones are **Issue Based Information Systems** (IBIS), **Decision Representational Language** (DRL), and **Questions Options Criteria** (QOC) (Burge & Brown, 2000). IBIS does not take into consideration more than one design alternative which QOC does, and DRL is a very rich notation technique which was thought of as too ambitious for this argumentation. An affinity diagram facilitates to organize the alternatives but does not help when selecting one of them as affectively as the design rationale. Accordingly, the QOC technique was chosen for its ability to arrange possible design solutions. (McKerlie & MacLean, 1993; Bellotti, 1993). The QOC notation helped organise the different definitions of screen density and, most importantly, to in a rational way evaluate them. QOC represents a network of **Questions** which are design issues to be addressed, **Options** which are solutions to those issues, and **Criteria** which are reasons for and against an option. The options are positively or negatively assessed by each set of criteria by a solid line or a broken line. The solid lines indicate a positive assessment and the broken ones a negative assessment. (McKerlie & MacLean, 1993; Bellotti, 1993).

The question addressed was: **How can information be distributed to define screen density?** The criteria important to a design of screen density were: \(C_1\) viable, \(C_2\) realistic, \(C_3\) measurable density, and \(C_4\) effectiveness (i.e., whether it measures the
The criteria were selected to ensure the validity of the screen density experiment. Construct validity was the main issue and the first criterion *viable* was set to ensure it. The second criterion *realistic* was related to external validity. The criteria *measurable* and *effectiveness* were related to internal validity.

Each of the five options was given 15 min of brainstorming and 3-4 ideas were generated per option. Some of this documentation can be seen in Appendix C. This was done to externalize the concepts and see the available layouts of each option. Following, the diagram was fulfilled according to the figure below.

![QOC-diagram](image)

**Figure 4.3.2: QOC-diagram**

To be able to use the result from the QOC-diagram the criteria had to be weighed since none of the option fulfilled all of the criteria. This was done with pairwise comparison where every pair is compared with each other and given a point (Dym, C. L. et al., 2002). If A is more important than B, A gets 2 points, if they are equally important A receives 1 point, and if A is less important than B, then A gets nought. This is a subjective judgement which is helpful when criteria or features are ranked in order of importance and was therefore chosen because its effectiveness. The sum of each criterion is presented in the second rightmost column of the table below. Each criterion’s points was divided with the total sum and shown in percentage in the rightmost column. The percentage was a measure of importance and consequently ‘measurable density’ was the most important criterion in Table 4.3.2, which illustrates each criterion’s point.
Table 4.3.2: Pairwise comparison between criteria used in QOC

<table>
<thead>
<tr>
<th></th>
<th>Viable</th>
<th>Realistic</th>
<th>Measurable density</th>
<th>Effectiveness</th>
<th>Total points</th>
<th>Weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viable</td>
<td>X</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>4</td>
<td>33%</td>
</tr>
<tr>
<td>Realistic</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Measurable density</td>
<td>2</td>
<td>2</td>
<td>X</td>
<td>2</td>
<td>6</td>
<td>50%</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>X</td>
<td>2</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>100%</td>
</tr>
</tbody>
</table>

The third criterion was considered the most important, following the first, the fourth, and less important, the second. To clarify they are listed below with decreasing importance:

1. Measurable density
2. Viable
3. Effectiveness
4. Realistic

This order was motivated by the three assumptions behind the method, see 1.2 Method. Of course, being able to measure density was the highest priority, following the criteria viable since construct validity was the main issue. The internal validity was more important than being able to generalize the results, and therefore effectiveness was a priori to realistic. When applying the weights in the QOC divided received the highest points and was the best way to define screen density in this experiment.

4.3.3 Card Sorting Test

Knowing how to define screen density the next step of the process was to establish what to display, in other words which functions to display and how to arrange them. This was settled with a card sorting test. A list of functions was received from Volvo, but neither the order nor the structure. A card sorting test was done by writing each function on a piece of paper and then sorting the papers into logical groups (Nielsen & Darrell, 1994), similar to an
affinity diagram. Normally this should be done by a group of real end-users, but due to lack of resources and time before the screen density experiment would be performed, it was done by the two of us instead. The sorting into logical groups was guided by the limitations of short term memory. Due to the fact that Miller (1956) established the limitation of short term memory to be 7±2 chunks and Sternberg stated that this capacity can drop to 3±2 by interference, it was concluded that the breadth of a menu should be approximately 5. The results from the studies (Miller, in Larson & Czerwinski, 1998; Tolle et al., 1987) which found that a menu structured 8² is to prefer were not used as guidance in the card sorting test. The reasons for this was that these studies tested menu structure for primary tasks and that a menu structured 8² only can contain 64 items whereas we were creating a menu for approximately 300 items.

4.3.4 Applying the Definition of Screen Density

To define screen density is not enough, it has to be applied in prototypes as well. In order to do this, the concept divided was diverged from the sketches produced earlier and during QOC. The result, after much debating, was to add a status field in the right of the display for high screen density, see Figure 4.3.4. This status field would show information like radio presets, radio frequency or CD and currently played tune, phone reception, and temperature for driver and passenger.

![Diagram of menu with status field](image)

Figure 4.3.4: Illustration of how the definition of screen density was applied. The left display is for low screen density and the right display with a status field for high screen density.

The advantage of this design was that a finite amount of information can be displayed on different number of screens without testing menu depth and breadth. The reason for this is that the menu in which the subjects perform their tasks is identical on the high and the low density display, i.e. the subjects have to go to the same depth and breadth in the menu regardless of density level. However, if real end-users should use the low density systems freely, they would have to go to additional screens to find the information already present in the status field in the high density display. For example, if a driver
wants to know the temperature in the low density display he/she has to navigate to a screen that displays this information, but to find the same information in the high density display it is sufficient to simply look at the display. In other words, the information on the high density display is divided into several screens on the low density display.

The choice to place the status field to the right was based on a comparison between different positions. To place it in the upper or lower field of the display would force the maps used by the navigation system to be extremely rectangular. To place it to the left and consequently close to the driver would result in the menu being further away from the driver therefore it appears to be not so important, i.e. assuming the driver sits on the left side of the car.

The level of screen density was defined as a quotient of the number of character on a display divided by the number of possible character spaces on that display. To calculate the number of possible character spaces in the 8” display used in this thesis, an 8” frame was drawn in a Word document. Subsequent to this, section after section from a 10 pages long text without white spaces was pasted into this frame. The number of characters in each section that was pasted in the frame was calculated. This resulted in an average number of possible character spaces that later could be used to calculate the screen density.

The difference between the two screens had to be as high as possible to give any possible effect later in the screen density experiment. Seen in earlier studies (see 3.4 Screen Density), the difference between high and low screen density is at least 20 percent units (Staggers, 1993; Morrison, 1989). The levels of high and low screen density have differed slightly in the previous studies on screen density, but high is about 50% and low about 25%. For this survey, the level for low screen density varied from 13% to 33% and for high screen density the level varied from 32” to 53%, due to different text amounts on each alternative in the menu. In average, the levels were at respectively 20% and 40%.

4.3.5 Graphical User Interface
The graphical user interface was developed from the QOC sketches and previous iteration of ideas of how the display could look. Considerations were taken to the research presented in 3.1.3 Guidelines for Usable Interface on colours, font sizes etcetera and to the findings from the heuristic evaluation presented in 5.1 Heuristic Evaluation. The elements will be explained and argued below the figures of the graphical user interfaces.
To the right in the screen with high density was the status field and the elements within it were grouped according to the Gestalt laws (see 3.1.2 Perception). The information presented was based upon requests made by Volvo and the graphics were taken from other concepts at Volvo, using Volvo’s own font. The colours of the bars of the phone reception were tested to be clear but yet melt in with the other elements and hence not be dominant and draw attention from the menu. The darker green against a lighter blue background ought to be legible even for a person who is blue-green colour blind.

In the top of the screen a navigator bar was displayed as a title and a navigation help to the system (Molich, 2002).

The menu was placed in the largest area of the display. The menu was divided into two sections, where the top menu was placed to the left and a list of submenus to the right, which gives the user direct feedback of what the
different menus contain. This description of each option hopefully minimized errors made by users (Paap & Cooke, 1998). When a submenu was chosen, the orange square moved to the top of that submenu and then the whole right side moved to the left and a new submenu appeared to the right. The title bar then displayed the chosen top menu as well e.g., Start > Radio.

![Diagram](image)

Figure 4.3.5c: Illustration of how a new submenu appears from the right when choosing Radio

The menu was accordingly vertical, and this choice was based on the fact that a top menu with five items would not fit horizontally with the minimum font size being used.

The rounded corners and to have the menu frame shadowed were motivated by its appealing look. The blue background colour was chosen because it was soft and suitable for backgrounds (Boyle, 1997; Bradshaw, in Geraci, 2002). The darker blue and greyish sides of the menu were tested with different shades of blue and grey to colour code the different levels of the menu and these colours were also soft and anonymous. The orange square around the present menu selection was taken from an earlier thesis work at Volvo (Nilsson, 2005). All the lines were placed to define, clarify and group the different types of information.

The end state of each branch of the menu hierarchy was a list of choices. Only one of the items could be selected and the choices were therefore marked with radio buttons. The user has to move to the item in order to select it and there was consequently no need for having the users confirm their choice by an extra button press. E.g., if a list had the alternatives 1, 2, 3, 4 and the goal was to select 3, it was sufficient just go down to 3 in order to select it, the user did not have to both move down to 3 and press OK.
4.4 Screen Density Experiment

One goal with this thesis was to test the safety aspect driver distraction for low and high screen density, according to the first problem statement. The experiment was done in a simulator. The experiment was within-subjects and is described in the following part. Within-subjects design was chosen to give the opportunity to compare the difference between performance with and without a treatment. Within-subjects design was also encouraged by the designers of the simulator test (ISO proposal for LCT, 2005) since it facilitates a statistical comparison between the different conditions.

4.4.1 Variables
The dependent variable was distraction. This was dependent of the independent variable screen density with the conditions 1) low and 2) high.

4.4.2 Task Description
The subjects were asked to perform tasks while driving in a simulator. The character of the tasks was to interact with a display showing functions like radio, navigation, phone etc. There were eight tasks (e.g., “Increase the volume on the phone to 2”) to be performed during each set. A description of the tasks is attached in Appendix D (in Swedish). If the subjects hesitated on were to go or went wrong in the menu, the test instructor helped them.

4.4.3 Evaluation Document
According to the second problem statement a questionnaire was included in the experiment after the subjects had finished the driving performance task. This questionnaire was included to investigate what the subjects’ opinions were on the displays and controls. The questionnaire was used instead of a structured interview since there were two test instructors. We believed that two different test instructors could impact more on the subjects’ answers in an interview than in a questionnaire. There were no Likert scales3 where the subjects could make their statement by putting a mark, or any other scales, instead all the questions had to be answered in their own words. There was no need to ask questions about the subjects’ background since all needed information already had been collected at the beginning of the experiment. The questionnaire was short with questions like: ”What is your opinion of the

3 A Likert scale is a scale where subjects are expressing their opinion by putting a mark on a line where the end points are opposites, e.g. nice-ugly or depended-independent.
menu structure?” and the questions were designed to be specific but still cover as much as possible by starting with general ones (Preece et al., 2002). The questionnaire was divided into two sub topics which would make it easier to complete (Preece et al., 2002); display and controls. There were no yes/no questions but instead the subjects had to motivate their answers and have an opinion in most of them. The questionnaire, in Swedish, is attached in Appendix E and the subjects could look at pictures of the two displays, high and low screen density, while answering it.

4.4.4 Subjects
The subjects were 22 volunteers from the Analysis and Verifying department at Volvo Cars. All subjects were at the time working with tasks not related to HMI, interaction design, usability or computing. Because of the security at Volvo, there was no possibility to perform the LCT test on subjects not having access to the area. There were 6 females and 16 males ranging in age from 21 to 62. Mean age was 36. Median age was 32. All participants had driver licence. The experiment was conducted at Volvo during the subjects working hours ranging from 7:30 am to 6:30 pm and they got a homemade cinnamon role for participating in the experiment.

4.4.5 Equipment
A driving simulator test called Lane Change Test (LCT) was used to assess driver distraction. This LCT test is an ISO standard proposal from Volvo Car Corporation among others, but not yet established. This test is being developed because a reliable and sensitive method to estimate driver distraction (while a secondary task is performed) is missing in the motor vehicle industry. The LCT is ”a simple laboratory dual-task method that quantitatively measures performance degradation due to distraction on a primary driving-like task while a secondary task (e.g., operating an in-vehicle system) is being performed” (outline from the ISO proposal for LCT, 2005, page v Introduction). The equipment for this test is a PC with a screen where the lanes are displayed, a steering wheel similar to a computer game device and two pedals: accelerator and brake. The test could also be implemented in a real car where you display the lanes on a big screen in front of the car and where you have a real accelerator, steering wheel and other devices. However, since this experiment had a secondary task, which was to navigate in a menu system not yet implemented in a Volvo car, the experiment had to take place in front of a computer.

The LCT is a motoric spatial task and it has nothing to do with actual driving, it just happens to be like driving a car (ISO proposal for LCT, 2005). There
are ten different tracks and every track is 3 km long. The test involves driving at a constant speed of 60 km/h, in other words one track takes 3 minutes. The tracks consist of a straight road with three lanes and every 60 meters a white sign tells the driver which lane to change to. A picture of the graphical interface is shown in Figure 4.4.5a. The driver should start to change lane immediately when the symbols on sign is displayed. If possible, the change should be completed before the sign is passed. There are 40 meters from where the driver sees the symbols until the sign is passed.

![Figure 4.4.5a: GUI of LCT](image)

The task device, consisting of a PC computer with a screen and a prototype of a center stack, was placed to the right of the simulator (see Figure 4.4.5b). On the screen, a prototype of a display on a center stack was seen in the middle. This was made in Macromedia Director and made in to a runnable .exe file. Since it was a prototype only the end state of the tasks was programmed.

![Figure 4.4.5b: The LCT-test environment](image)
4.4.6 Procedure
An invitation to participate in the experiment was sent out one week ahead of the experiment start. The subjects were given a list from which they could choose the time best suitable for them which resulted in three and a half days of testing. The experiment took approximately 45 minutes and consisted of doing the simulator test and answering a questionnaire. The participants were introduced to the simulator and then given time to get more familiar with it before the start of the test. They were also given the chance to navigate through the display prototype using the buttons and all eight tasks were explored before the real test begun. The subjects were informed that the experiment concerned how a different amount of information on a display distracts driving. The test instructions, in Swedish, are attached in Appendix F.

The test started with the subject driving one track without any secondary task. To avoid training effects, the subjects drove different tracks every time. The order of the tracks was randomly chosen for every subject. This was followed by driving a track while simultaneously performing a secondary task, in this case to navigate through the display prototype. This was done five times according to the figure below. High/low indicates which condition of the independent variable that round has. Every second subject started with high and the others with low. If the subject had high on the second track he/she had low on the fourth track, and vice versa. This was to avoid carry over effects. Another measure to avoid carry over effects was that the 8 tasks were performed in different order in the two conditions.

<table>
<thead>
<tr>
<th>Round no</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>Without</td>
<td>High/low</td>
<td>Without</td>
<td>Low/ High</td>
<td>Without</td>
</tr>
</tbody>
</table>

When the tasks were performed, in other words when the subjects interacted with the display, the start and end point of each task operation were recorded by the test instructor by pressing M and N on the keyboard. This recording was logged in a file which was produced for every track. The program created a computer file where all information about the performance was logged. Therefore when a subject had finished the LCT-test five different files had been made with data from all five different tracks. The statistical calculations were made using data from these files.
To compare the deviation between driving with and without a secondary task, the deviation of no secondary task, called basic line, was first calculated from a normative line. This is pictured in Figure 4.4.6a, where the normative line is in green and the basic line is in blue.

![Figure 4.4.6a: Basic line driving (blue line) in relation to a normative green line](image)

The deviation is consequently the area between these lines, indicated in red in Figure 4.4.6b. This area is sensitive to lane keeping, missed signs, reaction and manoeuvre.

![Figure 4.4.6b: Deviation from a normative line indicated by the red area](image)

The deviation between the secondary task run and the basic line was measured the same way. The deviation of the secondary tasks was measured only during the time the secondary tasks were being performed, not for the whole track. Since every start and end point of each task was recorded, the exact time when a secondary task was being performed can easily be found and compared with the average performance during the first, third and fifth run.

After the LCT-test the subjects were asked to answer the questionnaire.
4.4.7 Redesign of Graphical User Interface
Comments made by the subjects in the questionnaire led to a redesign of the graphical user interface. This modification was made in order to give Volvo a usable prototype in the end. The user preferences of the menu were taken into considerations as it was noted during the test that the subject had difficulties with some of the tasks. Although this redesign was not validated, it is presumably more usable than the original version since the redesign was based on the subjects’ comments on how they wanted to interact with the system, as well as how they interacted and what errors they made during this interaction. The redesign is presented in 5.3.3 Redesign of Graphical User Interface.
5 RESULTS

The following chapter presents the results from the surveys described in 4 Method. It includes the heuristic evaluation, the control survey, and the screen density experiment. Results from the development of the display of screen density can be seen in the method (see 4.3 The Display of Screen Density). The card sorting test was made from a list of functions in a future Volvo and due to Volvo’s confidentiality concerning new cars the outcome of this test cannot be presented in this thesis.

5.1 Heuristic Evaluation

Findings from the heuristic evaluation of in-vehicle information technology in Volvo cars and competitor brands are presented below (the heuristics are described in detail in 3.3.1 Usability Heuristics). The findings are identified usability concerns, specific for in-vehicle information systems, and serve as guidance in the design of the display measuring screen density.

- **Feedback** – It is good to use a marker that indicates when the system is working, in other words the status of the system. This can be in shape of a circle being filled or a string saying that something is being loaded. The feedback should be clear so that it is obvious to the driver what the system is doing. It is also important that the feedback does not annoy the user. For example, a beep every time a button is pressed is acceptable in the beginning, but can get rather irritating after a while.

- **Speak the user’s language** – Abbreviations should be avoided as much as possible. Obviously, this can be difficult to accomplish in an in-vehicle information system since the available space to present the information on is limited. If abbreviations cannot be avoided, they should be explained. Another way to speak the user’s language is to allow the possibility for the user to choose the language in the system.

- **Clear marked exits** – Clear marked exits are important if the user wants to undo a choice. The majority of the systems in the evaluation made this possible through a BACK or RETURN button.
• **Consistency and standards** – The ON button should not stand for both ON and OFF. If a button represents both functions, the button should rather be called something else. There should be a consistency of what the user can do with the num pad. If the num pad has letters on the digit buttons as it has on a cell phone, the user should be able to enter a text string in the navigation system with it as well as dial a phone number or select a pre-set on the radio. The system should be designed to be consistent with the users’ goals and not merely with the underlying functionality of the system. If a user wants to write a text, the way to input text should be consistent whether the text is a street name, a SMS, or a song title.

• **Error prevention** – The buttons MENU and BACK should stand for menu as in top menu and back as in go back one step in the menu hierarchy. Systems that do not use this convention are conceived as inconsistent and hard to learn.

• **Recognition rather than recall** – It is important to show the user where in the system he/she is, not just in what menu but also how deep down in that menu. This can be done by a navigation path. A clear and obvious menu structure also helps the user navigate in the system. If a user himself/herself has to remember the position in the system it consumes cognitive resources that are better used for the driving task.

• **Shortcuts** – The most frequently used functions should be made easy to reach. This can be done by placing them high in the menu hierarchy, or by providing direct access buttons to them.

• **Simple and natural dialogue** – Keep the sentences as short as possible but avoid using abbreviations. The designer of an in-vehicle interface has to be aware that the vast majority of the users do not know all the technical car terms. If the technical terms cannot be avoided there should be simple ways for the users to read an explanation of the term.

• **Provide good error messages** – There were not many cases where an in-vehicle information system has error messages probably because the functionality of the system often is limited and the system prevents the user from doing wrong. Furthermore, there is neither the time nor the place to show a long understandably error message to the driver.
• **Help and documentation** – If the system communicates with abbreviations and/or a technical language, it is of great importance that it also provides help and explanations.

### 5.2 Controls

The results presented in this section are from two different investigations; the 2004 *Navigation Usage and Satisfaction study* (J.D. Power and Associates, 2004), and the control survey. The results are presented in the form of design recommendations.

#### 5.2.1 J. D. Power and Associates

The result from the evaluation of the selected cars from the 2004 *Navigation Usage and Satisfaction study* (J.D. Power and Associates, 2004) is a collection of do’s and don’ts for controls in vehicles. There was clear trend amongst the cars evaluated – cars with good grades had similar properties and cars with bad grades had similar properties. The do’s and don’ts served as guidance in the design of the controls used in the screen density experiment, and are the following:

**Do’s**
- Provide explicit cues if a knob can be turned *and* pressed
- Group controls logically
- Provide direct buttons to frequently used functions
- Make the buttons clear and explicit
- Provide some kind of tactile response which extinguishes different controls
- Use gestalt laws to group of controls
- Provide space between buttons
- Place, if possible, controls close to respective function. E.g. stereo buttons close to the CD-slot.

**Don’ts**
- Knobs with a turnable outer ring and a pressable button in the centre
- Double buttons with two functions
- Short distance between controls
• Many different sizes of controls
• Small controls
• Unclear symbols and abbreviations
• Clumsy controls
• Plastic ridges between buttons
• Buttons in the same level as the surface
• Small colour differences between the background and the text colour

Accordingly, the do’s are guidelines to follow and the don’ts things to avoid in the design of in-vehicle controls.

5.2.2 Control Survey
The control survey investigated two issues, first how the subjects use the controls, and secondly what combinations of control-menu they prefer.

All of the subjects used the four-way seesaw in the same manor when they performed the tasks. They pressed the right arrow to go right in the horizontal menu and the down arrow to go down in the vertical menu. However, differences occur when subjects used the knob. In the horizontal menu, all the subjects turned the knob to the right (i.e., clockwise) to go to the right. In the vertical menu, 13 subjects turned the knob to the right and 5 subjects to the left, to go down.

The test showed that 8 out of 18 real end-users preferred knob with horizontal menu, 1 user preferred knob with vertical menu, 2 users preferred four-way seesaw with horizontal menu, 3 users preferred four-way seesaw with vertical menu and 4 users preferred four-way seesaw with any of the two menu-types.

Table 5.2.2: User preferences for different type of control-menu interaction

<table>
<thead>
<tr>
<th>Knob - horizontal</th>
<th>Knob - vertical</th>
<th>Four-way seesaw – horizontal</th>
<th>Four-way seesaw – vertical</th>
<th>Four-way seesaw – any menu</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

Comments from the subjects during the test were noted and are attached in Appendix G but are not translated from Swedish.
To summarise, the four-way seesaw gives the least delimitations when designing the graphical interface of the menu. The reason for that is, with the four-way seesaw it does not matter if the menu is designed horizontally or vertically, and the interaction with the four-way seesaw would presumably give the least incorrect button presses. However, this result is only valid for a menu structure with relatively few alternatives, 6 to be precise. Volvo has found\(^4\) that a knob is to prefer when you have a long list of items, e.g. a play list of artists in a MP3 player.

### 5.2.3 Center Stack Prototype

The results from the above mentioned evaluations and survey led to a design of a center stack prototype. This was specially made for the screen density experiment. It consisted of five direct access buttons representing the top level menu: Radio, Media, Navigation, Phone and My Car, see Figure 4.2.3 in \(4.2.3\) Center Stack Prototype. These menus could be reached from everywhere in the system which was a feature found usable in the heuristic evaluation. The sixth button was not used in the experiment. There was also a four-way seesaw with an OK button in the middle. The four-way seesaw enabled navigation up and down within a chosen menu, choice of function (right arrow), and back one step (left arrow). The OK button had the same functionality as the right arrow, accordingly to go down in the hierarchy. A BACK button was placed beneath the four-way seesaw and it had the same functionality as the left arrow button. This was based on the results from the control survey.

The additional sixth direct button and the redundancy of OK and right arrow button, and BACK and left arrow button, was a result of having to decide a design of the center stack prototype at an early stage of the process. The reason for this was that the center stack had to be built before the experiment was conducted and at this stage in the process, the design of the graphical user interface of the display had not started.

---

\(^4\) Staffan Davidsson, Human Factors Engineering and Ergonomics, Volvo Cars Corporation
5.3 Screen Density Experiment

The results from the LCT and the questionnaire are presented in this section. The aim of the LCT was to investigate whether there was a difference in driving performance between interacting with a low screen density display while driving, and interacting with a high screen density display while driving. In this survey driving performance is called safety and is defined and measured by the distraction on the primary task. The aim of the questionnaire was to establish if users preferred high or low screen density.

5.3.1 Safety

When performing secondary tasks on a low density screen the distraction on the primary task was 0.426 m and when performing secondary tasks on a high density screen the distraction on the primary task was 0.450 m (see Table 5.3.1).

<table>
<thead>
<tr>
<th>Level of density</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of distraction (m)</td>
<td>0.426</td>
<td>0.450</td>
</tr>
</tbody>
</table>

The results from the LCT-test was analyzed with a one way analysis of variance (ANOVA) on the mean deviation in meters between the two conditions of the dependent variables, low and high screen density. The ANOVA showed no significant difference between the two conditions low density and high density on distraction.

The distraction of the secondary tasks (both high and low density) was below the recommended value of how much a secondary task may distract driving. This level is set at 0.5 meter but is not an empirically tested value in the LCT proposal (ISO proposal for LCT, 2005).

5.3.2 User Preferences

The subjects finished the screen density experiment with a questionnaire. The outcome was the following. 19 of 22 subjects preferred the screen with the extra field to the right (i.e., high density screen), 1 subject preferred the screen without the extra field (i.e., low screen density), and for 2 subjects it did not matter which, see Table 5.3.2. The comments on why they chose the high density screen were of the same character; the extra information was not
disturbing when navigating in the menu and it was better to have the
information present than to find it in the menu structure.

Table 5.3.2: User preferences for low or high screen density

<table>
<thead>
<tr>
<th>Low density</th>
<th>High density</th>
<th>Either way</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19</td>
<td>2</td>
</tr>
</tbody>
</table>

The subjects’ answers about the menu structure as well as their proposals and
comments are summarised below. It was obvious that a feature in the system
could be liked by one subject and disliked by another. The citations below are
the ones that differed from the most common comment, i.e. that the menu
structure was logically compiled. The comments are ordered with the negative
ones first, followed by the positive ones (translated from Swedish to English
by the authors).

"why so many button presses? one press per function is optimal"

"hard to find with so many sub menus"

"too much information to search, it is too much text"

"sometimes the menu structure is not logically because it is ambiguous where
some items can be found"

"avoid too many menu levels"

"icons would facilitate"

"good top menu, but then complicated"

"natural"

"reminds me of a mobile phone or a TV"

"easy to use, good with optional ways of navigating"

"good with the head menu to the left and the sub menus to the right"

"clear menu structure"

"good with relatively few choices"
The subjects also answered questions about the feedback, i.e. the graphical feedback of an action. Unfortunately, the answers were poorly developed, most of the subjects simply wrote “OK”. However, a few interesting points were made. Two subjects thought it was good that an action did not have to be confirmed with ENTER or OK, but two other subjects wrote the opposite, that they wanted to confirm every choice. Two other comments were that the feedback was clear and that more colours in the system would be preferable.

The last part to evaluate in the questionnaire was the controls. It was good with arrow and direct access buttons. Some proposals of improvements were made by the subjects.

- The BACK button should be placed where left arrow was located.
- It was easy to accidentally press OK when aiming for the up or down arrow.

5.3.3 Redesign of Graphical User Interface

All these proposals and comments were taken into account when redesigning the user interface. The new prototype has the following improvements:

- TRAFFIC INFORMATION in RADIO menu was moved to SETTINGS, one menu level lower down, see Figure 5.3.3a

![Figure 5.3.3a: Redesign of menu structure](image-url)
• VOICE MAIL in PHONE menu was moved to PHONEBOOK, one menu level lower down, see Figure 5.3.3b.

• SETTINGS in PHONE menu was named SOUNDS & ALERTS with RING VOLUME, RING SIGNAL and MESSAGE ALERT, not SELECT PHONE and AUTOMATIC ANSWER, see Figure 5.3.3b.

• HANDSFREE and SELECT PHONE were moved up one level from SETTINGS, see Figure 5.3.3b

![Diagram of menu structure]

Figure 5.3.3b: Redesign of menu structure
• In MYCAR, PERSONALIZE was named SETTINGS instead, see Figure 5.3.3c.

• A radio button choice has to be confirmed with OK or the right arrow button

• If the controls were to be redesigned the redundancy and unused buttons should be avoided
6 DISCUSSION

There has not been one specific methodology throughout this project, but rather a collection of methods to suit the various sub-goals. The methodology by Shneiderman & Plaisant (2005), described in 4 Method, has been a guiding help designing the graphical user interface and the interaction. These two latter parts have been sub goals, whereas the screen density experiment has been the primary goal of this thesis. If the thesis had been primarily in the field of interaction design, the emphasis would have been on the methodology instead of the experiment.

The methods used, the results from the different surveys, and above all the validity of each part of the project, will be discussed in this chapter. The chapter is divided into several sections, each discussing a different survey, starting with the heuristic evaluation, followed by the development of the means of interaction and graphical user interface, and finally the screen density experiment with the subjective evaluation questionnaire.

6.1 Contextual Research with Heuristic Evaluation

The heuristic evaluation was conducted to discover usability problems for in-vehicle displays in order to attend these as part of the design process. The fact that there were only two evaluators had the effect that only approximately half of the usability problems were captured (Preece et al., 2002). However, since the evaluation only was used for benchmarking and not as means for capturing error in the design process, it was decided that 50% was adequate for this thesis. Nonetheless, the reader should note that more evaluators and/or complementary evaluations based on other heuristics or usability principles would probably have led to identification of additional usability problems.

The results from the heuristic evaluation are relevant for the design of the controls and the interface used in the screen density experiment. However, due to the low number of evaluators, generalization of the results to all in-vehicle information systems is not suggested. The external and statistical validity are not high enough for such a conclusion. Instead, it is recommended
to see the results as indicators of what type of usability problems can be found.

6.2 Development of means of interaction

The aim of this part of the project was to design a center stack prototype on basis of literature studies, an evaluation of the results from J.D. Power and Associates’ report (2004), and a investigation of users’ behaviour with different types of navigating devices combined with different menu structures. The discussion starts with the evaluation of the report, followed by the control survey, and ends with a short discussion of the final center stack design.

6.2.1 J. D. Power and Associates

J.D. Power and Associates is well reputed in car industry. Their annual report on Navigations Usage and Satisfaction Management is of many car manufacturer thought of as a valuable source and indicator of which models the buyers prefer. Most of the car models evaluated in the study had over 100 responders each, and the internal validity can therefore be considered high.

Due to the fact that some cars were eliminated (see 4.2.1 Report from J. D. Power and Associates for reasons behind elimination), the selection that was evaluated is not complete. Nevertheless, there was a clear trend amongst the cars evaluated – cars with good grades had similar properties and cars with bad grades had similar properties. Due to this clear trend, the control features attained can most likely be generalized.

It is interesting to question what the original survey by J.D. Power and Associates (2004) investigates. The four questions, listed in 4.2.1 Report from J. D. Power and Associate, are all related to user satisfaction of controls, but it can be questioned what satisfaction really means. That is, whether all subjects have the same understanding of the word. However, due to the high sample of more than 100 subjects for most of the evaluated models this issue can be neglected.

Another matter with the study is that subjects might not have the same common ground regarding expectation. That is, subjects might expect different qualities from their cars. The study is investigating satisfaction concerning new models. Consequently, every subject is an owner of a new car, and is evaluating this new car. However, the difference between buying a SEK 400 000 car or a SEK 800 000 car is quite large. If a subject pays twice as
much for a car, the expectations are higher and the frame of satisfaction is therefore much more narrow, than if the subject would have bought the car at half the cost. The purchase itself can also mean different things to different subjects. To buy a car can be a big event connected to high expectations for one subject, whereas another subject buys the same car simply as one in the line of new cars, and hence without equally high expectations. Accordingly, it is not possible to conclude that a high grade on one model would be equal to a high grade given to another model. This is a drawback with using a between-group design, as in the survey by J.D. Power and Associates. Nevertheless, the high sample ought to eliminate most of the risks of this character.

The survey by J. D. Power and Associates was based on opinions of American buyers, whereas the control survey in this thesis was done on Swedes. It can be argued that there is a difference between Americans and Swedes regarding opinions, expectations, and other relevant values. Based on these cultural differences it could be questioned whether the external validity is high enough to generalize the results to other populations beyond the nationality they represent. However, the cultural differences, in this case preferences for controls, between Americans and Swedes are most likely not to interfere with the results of this study.

J. D. Power and Associates investigated things gone wrong, as well as satisfaction of understanding, using, and looks and feels of the controls while the evaluation of our results only concern the looks of them. This will influence the construct validity.

When evaluating the good and the bad controls there is no certainty that all features are captured, or that the features established are the most relevant ones. The selection of evaluated models is relatively low, 9 good and 5 bad from the 78 models presented in the study. Due to this low sample, it can be questioned whether the result can be generalized to all controls in vehicles. An argument for generalization is the clear trend discussed in previous section. The cars with good grades, and the cars with bad grades, respectively had very similar features.

Detailed heuristic evaluation was not used because there were no access to the different cars in physical terms but only in pictures and this neither gave us the opportunity to touch and feel the controls by ourselves nor to navigate through the systems. These were also the reasons why we did not use qualities in use in evaluation process.
6.2.2 Control Survey

The survey investigated two aspects of controls; how do users use different interaction devices in combination with different menus, and what combination of control-menu the users prefer. The subjects’ tasks given to them in the form of instructions (e.g., ”Go to Players”) and not of scenario character (e.g., “You want to change a setting under Players, what do you do?”). Furthermore, the interfaces were plain and not complete. However, these potential distraction variables will most likely not interfere with the interaction. The plain interface also facilitates the users to only make comments concerning the menu structure, and not the graphical design. It can therefore be argued, despite of the distraction variables, that the outcome can be generalized to all in-vehicle information systems.

Something that can interfere with the outcome is the way the horizontal and vertical menus are designed. The results might have been different if the menu would have been spread out as a fan or as a half circle (see Figure 6.2.2), instead of a straight horizontal or vertical line, as they were in the experiment. The connection between the knob and a fan or half circle menu could be much tighter and the subjects could therefore potentially have chosen this interaction instead of the four-way seesaw.

![Figure 6.2.2: Sketches of how the menu could look if shaped as a fan](image)

The control survey investigates users’ behaviour and preferences for controls and menus when the interaction is a primary task. It can therefore be discussed whether the results can be generalized to situations where the interaction is a secondary task. The external validity is therefore debatable. There is also the question of whether this influences the screen density experiment. Presumably not much since the comments in the questionnaire
were overall positive and the subjects in the experiment seemed to like the four-way seesaw.

6.2.3 Center Stack Prototype
The center stack prototype was designed before the graphical user interface and as mentioned in 5.2.3 Center Stack Prototype all aspects of the center stack design was therefore not adjusted properly to the interface. It had redundant controls and a spare direct access button. It can certainly be questioned whether this reduced the usability in the system. Even though some controls were redundant and another unnecessary, the fact that the interaction was consequent hopefully kept the usability relatively high. With additional resources the redesign of the system would not only concern the interface but also the center stack.

6.3 Development of Graphical User Interface
The goal with the development of graphical user interface was to design an interface with high ecological validity. This is achieved when the interface is perceived as usable, and when it is similar to other in-vehicle information systems. The starting point of this development was establishing the design of screen density.

6.3.1 Screen Density
Before choosing a particular design in a design process, a certainty has to exist that all possible design solutions available have been externalized. If this certainty does not exist, the design team cannot know for certain that the design they chose is the best of all possible solutions – it could just as well be the least bad one of the externalized options. A design team can achieve the needed certainty by numbers, i.e. more designers generate more solutions. The probability that all solutions are externalized increases with the number of designers. In the design process of this thesis, the team only consisted of two designers and the certainty could therefore not be achieved by numbers. Instead it was achieved through literature studies, benchmarking, and iterative brainstorming with divergence, all steps with the purpose of externalizing all possible design solutions and making the design space as wide as possible. However, even though a certainty existed that all these measures resulted in the best design solution available, it can always be questioned whether the design actually is the best one or simply the least bad one.
Moving on to the next step in the process – to choose one design from the ones externalized. Choosing design rationale for selecting a design is seen as the most valid technique. It can be argued that the result from the selection process is weak since there were only two designers involved in the process. Nevertheless, the other available techniques were ruled out. Participatory design was excluded because the real end users did not have the knowledge to participate in the process towards a design that could test screen density. Furthermore, usability engineering was ruled out because quantitative goals were not relevant in this project. They are more often used when the design approach is to make changes to an earlier version of a system based on well-documented results of usability test.

6.3.2 Card Sorting Test
The validity of the Card Sorting Test can be questioned. If a design team wants to create a system that is as usable as possible, the card sorting test should be done with real end-users and not by the team itself, which was the case in this study. Furthermore, the functional structure should have been tested on real end-users before the screen density experiment. This was also indicated in the subjective evaluation. It is essential that real end-users find the menu structure logical and usable. A workshop or a cognitive walk-through could have been made for this purpose, but was excluded for the same reason as the exclusion of using real end-users; lack of time before the screen density experiment should be conducted.

6.3.3 Graphical User Interface
The design of the graphical user interface was not a primary goal of this project. Nevertheless, it could have been tested on real end-users before it was used in the screen density experiment. However, instead of basing the interface design on user opinions, the design was based on extensive literature studies and guidelines.

6.4 Laboratory Experiment
The validity of the experiment will be discussed first and foremost in this section. In addition to this, the LCT and the research area of screen density will be discussed.

One big concern is whether screen density has been measured in this study. This is a question of construct validity. What differs between the two prototypes is the field with additional information to the right. This extra field
raised the density but it did not result in higher distraction. An explanation for this could be that the additional information was grouped in a logical way and that grouping is dominant over screen density. That is, if a display is well grouped, additional information does not distract the driver. If this is the case, it is an interesting finding, though on the other hand it affects the construct validity. However, this thesis is built upon the three assumptions stated in 1.2 Method, and the second one states that two usable prototypes should be designed. Logical grouping is an important part of usable interfaces and the interface has been designed according to the assumption.

To increase or decrease the density level in a display, the information can be distributed in several ways (see 4.3.1 Definition of screen density). One way to lower the level is to divide information into several screens, i.e., raising the number of screens in breadth or in depth without changing the total amount of information. Another way is to decrease the font size which leads to a higher maximal available character space. In the first case, the numbers of screens is not constant when the level is changed. If this is used in an experiment, there would be an uncertainty whether it is the density level or the number of screens that actually is examined. In other words, it could be a study of depth versus breadth, instead of screen density. In the second case, it can be a problem to read small font size. If two prototypes display the same amount of information – one with small font size (low density) and one with large font size (high density) – it is more likely that legibility is measured instead of screen density. Furthermore, if the information available would increase in the small font size display, then the screen density would be the constant because screen density is measured as a percentage of available character space. The chosen screen density design is optimal for the experiment in this thesis but the result could depend on other factors as well, which leads to internal validity.

6.4.1 LCT
The means to measure distraction, according to the LCT-test, has been widely discussed during the course of this project and the following is an extract from these discussions. The first problem with the LCT is that the steering wheel is smaller than a normal steering wheel, so the tiniest wheel movement results in more movement on the road than a normal wheel would give. This is a threat to the ecological validity. The second issue is the screen of the LCT test being small and placed in front of the steering wheel. This will have the effect that the middle of a lane is in front of the driver, and not to the right which is the case in reality (presumed it is a car where the driver sits to the
left). The consequence is that when the driver looks at the display the eyes move more than they would have done in a regular car. This is a threat to the ecological validity but it also affects more directly the way the subjects performed because the distance the eyes have to move is longer than normally. Probably, this small difference from a normal car does not have any significant effect on the experiment. Although, since the test is similar to driving a car but the driving experience differs, the subjects are distracted by this fact. When they accidentally steer the car to the grass, they think badly of themselves as drivers. It could therefore be a mistake to have the LCT test looking like driving a car when the purpose of it is to measure the distraction of performing a secondary task while handling a more motoric character task (ISO proposal for LCT, 2005).

The LCT test is not the only method for testing the amount of distraction a secondary task imposes on the primary driving task. There are at least two other methods; the 15-seconds rule and the occlusion method (see 3.2.3 Measuring Driver Distraction). The 15-seconds rule measures the time it takes to complete the whole secondary task and the Occlusion method measures the time it takes to complete chunks of the secondary task. The assumption behind these two methods is that distraction can be measured by the time the eyes are off the road. The assumption behind LCT is instead that distraction is measured by the difference in driving performance between only driving and driving while performing a secondary task. Driver distraction can be seen as being more than having the visual attention off the road, it can also be connected to the cognitive attention being off the road (Harbluk, 2002). Due to this way of looking at distraction, the LCT was chosen since it not only measures visual distraction, but the cognitive aspect of it as well.

Having the display placed on top of the center stack will cause unnecessary change of eye focus. The placement of the display on top of the center stack is motivated by Volvo as the smallest possible change in eye focus between road and display. However, with this placement no consideration is taken to the change of eye focus that happens when the driver looks at the control. Consideration has to be paid to the system in its entirety when motivating the placement of one part of the system. With the display placed on top of the center stack and the controls placed in the middle of the center stack the driver has to change eye focus from the road, to the top of the center stack, to the middle of the center stack. If the display as well as the controls were placed in the middle of the center stack the change of eye focus would be from the road to the middle of the center stack (see Figure 6.4.1). Due to the limitations stated in 1.3 Delimitations this study did not investigate the best
placement of display and controls. However, if the placement causes unnecessary change of eye focus it would result in the prototypes used in this thesis not being usable from a safety point of view and would therefore not be designed according to the second assumption (see 1.2 Method).

![Diagram showing eye focus changes](image)

**Figure 6.4.1: Change of eye focus depending on different display positions**

### 6.4.2 Inference of subjects

Another factor that could interfere with result and internal validity is that there were two different test instructors. Although the instructions were written down, a difference between the groups was noted. In average, it took about 5 minutes longer for one of the instructor’s subjects to perform the entire test. The results from the two groups were compared and no significant differences were found. Consequently, there were no differences in test results between the subjects of the two instructors.

This leads to the question whether the time of the day when the test was performed, or other environmental factors, had any effect of the performance. The latter issue is easily answered since the environment was equal to all. The first issue is probably balanced out since the subjects were equally spread out during four days. One thing that was noted was the fact that many of the subjects were excited when they were told that information on an in-vehicle display was tested. Some of them were of the opinion that today’s cars have too many functions and that the information systems are not easily understood. This could have interfered with their performance if they wanted to direct the study according to their opinion by performing poorly. Although the risk of this type of behavior from the subjects is not high, they knew neither difference between the displays nor exactly what was being tested. Consequently, they could not have known which display to perform worse on, assuming they wanted to direct the study.

A fact that can contribute to low ecological validity is that the subjects’ cognitive processes are of different character in the LCT compared to real driving situations. In the LCT they listened to instructions and received help
to find functions in the menu, whereas in a real driving situation they would formulate a goal in their mind and find the functions because they are used to the system. Even though the cognitive processes differ, the mental workload can be the comparable and therefore effect the driving in the same way. The workload of a novice user getting help to find functions can be comparable to a user who knows the system well. They find the function equally fast; one due to help and the other due to experience.

Volvo has questioned the fact that the subjects’ driving experience was not noted. The argument against this is the within-subject design of the LCT-test. A comparison is only made between the subjects own result, and therefore it is not of any interest to show statistics of driving experience. Another argument is that LCT is not a driving simulation test but a dual-task simulator which happens to be similar to driving a car while doing a distracting secondary task (ISO proposal for LCT, 2005). It would therefore be more interesting to note the subjects’ computer game experience or having them to estimate their simultaneous capacity.

6.4.3 Levels of Screen Density
A further concern with the experiment is that the difference between low and high screen density might have been too small impact of the results. A level difference of 20 percentage units could have been too small. A third level at about 60% could have made the experiment more interesting and accordingly calculate the distraction levels and difference between low (20%), high (40%) and higher (60%) screen density.

Due to the flaws of the LCT-test and the absence of pilot tests before the real experiment, the results might not be strong enough for generalization. It would nevertheless be correct to say that the conclusions made in the next chapter consider all factors discussed here and that the result therefore gives a hint that high density can be used in an in-vehicle display if the information is well grouped and features like legibility and a usable structure of functions are applied. It is important to mention that the screen density level result can be generalized to secondary tasks screens, nothing else.

6.4.4 Ethical Issues
The ethical issue comprises four different parts; information, consent, confidentiality, and use. (Breakwell et al., 2000). The subjects were informed that the results could not be associated to them personally, and what the results would be used for. They were not informed that they could abort their
participation, how the participation would affect them physically or mentally, that all the data would be kept confidential or the Principle of informed consent, in other words that it was screen density that was examined. It was a mistake not to inform the subjects that they could abort their participation and that the test would not affect them, but not to inform them about exactly what was being tested was a deliberate choice because this could interfere with how they performed on the LCT-test and above all their answers in the questionnaire.

6.4.5 Qualitative Subjective Evaluation
The qualitative subjective evaluation was answered immediately after finishing the experiment. The ethical issues are therefore mostly the same as mentioned above

It was found that some of the questions did not generate the expected answers. This can be explained either by the evaluation document not having a satisfactory design, or that the subjects were not used to giving qualitative answers. The fact that the instructor sat next to the subject could also have interfered with the results if the subjects did not want to give negative comments about the system when the designer is sitting next to them. Another negative aspect of this part is the fact that the subjects were quite bored after the experiment and some of them did not give any thoughtful feedback in the questionnaire but rushed through it. Unfortunately, the design of the LCT method clearly states that the subjects have to drive five tracks and the number racks could therefore not be reduced. As the case was with the card sorting test, the graphical user interface and the screen density experiment, the evaluation document could have been pilot tested before using it.

To sum up the discussion, an in-vehicle information system is an interesting environment to apply the concept screen density on. The amount of information to display in vehicles is a well debated subject and therefore this survey could bring some light to the car industry’s widely criticised approach to put more functions and devises into the in-vehicle environment.
7 CONCLUSIONS AND RECOMMENDATIONS

Conclusions and recommendations for further work will be presented here.

7.1 Conclusions

The problem statements, stated in 1.1 Purpose, are:

1) What level of screen density, high or low, can be recommended for a safe in-vehicle display?
2) What level of screen density, high or low, do users prefer?

Answers to the problem statements are:

1) There is no significant difference in distraction on the primary task between low and high screen density, hence both levels of density are equally safe. Furthermore, both levels are imposing a distraction on the primary task that is below the 0.5 meter limit\(^5\) for allowed distraction. Both levels can therefore be recommended for a safe in-vehicle display.

2) Users prefer high screen density over low screen density.

Conclusions and reflections from this thesis are:

- The best means for interaction is highly depended on placement of controls in relation to the display, and on mapping between controls and the item being managed on the display.

- There is a clear trend regarding good and bad features of controls for in-vehicle information systems, and these can easily be applied or avoided.

- Grouping is a central issue when designing controls and user interfaces. Particular for in-vehicle information systems where the interaction in most cases take place in expense of the driving task.

- The need to look at the hand while managing or finding controls in the center stack is a big source of distraction.

\(^5\) Anders Hallén, Human factors Engineering and Ergonomics, Volvo Cars Corporation
• A difference of 20 percentage units between the density levels is probably too small to receive significant differences in distraction.

• The LCT-test has advantages as well as disadvantages. It is important that the test instructor is familiar with these in order to use the test properly and generalize the results in a correct way.

• The term screen density is not well known and the area has not been researched to any greater extent. Earlier research has not compared screen density levels on a display for a secondary task by measuring performance on a primary task. It was therefore interesting to investigate this aspect of screen density. It is also relevant when designing in-vehicle systems. However, it can be questioned how relevant density is compared to grouping and legibility.

7.2 Further work
These recommendations include both proposals for further work at Volvo and suggestions for academic research in this area.

• It would be relevant to continue the research on in-vehicle screen density by performing tests on different density levels in the DIM (Driver Information Module, the area behind the steering wheel).

• Since no significant difference was found in driver distraction between high and low screen density it would be of interest to investigate what happens when density level is increased further. This could be done through an experiment that compares the distraction between 20%, 40% and 60% density.

• During this project, the importance of logical grouping has been noted. It would therefore be of significance to investigate the effect different grouping has on driver distraction.

• The LCT-test showed that distraction levels of both low and high screen density were below 0.5 meter. This implies that both levels can be used in in-vehicle displays without distracting the driver in a way that would reduce safety. The 0.5 meter limit is only a suggestion (ISO proposal for LCT, 2005). However, since many subjects commented on the distraction level being prohibitive for using the display in real driving situations, it is strongly recommended that this limit is empirically tested further before the LCT becomes an ISO standard.
8 REFERENCES


Ek, C. & Myhrman, E. (2004). *Presentation of driver information. User centred redesign of the instrument cluster and the trip computer*. Göteborg: Department of Mechanical Engineering, Institution for Product- and Production Development, Division for Human Factors Engineering, Chalmers University of Technology,


# Appendix A: Heuristic Evaluation Form

<table>
<thead>
<tr>
<th>Heuristic</th>
<th>Bilmodell</th>
<th>Toyota 2,0/MT-5</th>
<th>Rav 4</th>
<th>Saab 9-5 Sedan R4 2,2</th>
<th>MB E-270 kombi</th>
<th>Audi A8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Återkoppling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tala användarens språk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tydligt markerade utgångar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Överensstämmelse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Förebygga fel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimera användarens minnesbelastning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genvägar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enkel och naturlig dialog</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bra felmeddelanden</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hjälp och dokumentation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Test instructions for control survey

Formulär för kvalitativ undersökning av reglagen vippa eller vred

Tala om för fp att denne sitter i en bil och kör och att man har ett reglage för att manövrera displayen som sitter i centerstacken (konsolen mellan föraren och passageraren).

Utförande

Visa först den horisontella menydisplayen med vippan. Markera Players och be fp flytta sig till Climate. Byt till vredet och be fp göra om proceduren, alltså jag markerar Players och fp flyttar sig till Climate mha vredet. Reglagen ska vara placerade till höger under menyerna.


Fyll i formuläret med H, V, U eller N.
H = höger på vippan och medsol på vredet
V = vänster på vippan och motsol på vredet
U = upp på vippan
N = ner på vippan
Fyll också gärna i ålder och kön på fp.

Hvi Hvr Vvi Vvr
Vvr Vvi Hvr Hvi
Hvr Hvi Vvr Vvi
Vvi Vvr Hvi Hvr
Hvi Hvr Vvi Vvr
Vvr Vvi Hvr Hvi
Appendix C: Lo-Fi examples of screen density definitions

Distributed

Low

NAV

RADIO

MY CAR

PLAYERS

PHONE

High

NAV

RADIO

MY CAR

PLAYERS

PHONE

Low

NAVIGATION

Radio

MY CAR

PLAYERS

PHONE

High

NAVIGATION

Radio

MY CAR

PLAYERS

PHONE
Mediated
Graphical Congregated
Descriptive
Appendix D: Description of tasks

Uppgift A Ställ in att radion ska brytas när det sänds nyheter
> Gå in i Radio
> Gå in i Settings
> Gå in i News
> Välj ON

Uppgift B Ställ in så att radion sätts på när det sänds trafiknyheter
> Gå in i Radio
> Gå in i Settings
> Gå in i Traffic Information
> Välj ON

Uppgift C Stäng av random under Media
> Gå in i Media
> Gå in i Random
> Välj Off

Uppgift D Sätt på TV:n i baksätet
> Gå in i Media
> Gå in i Rear seat entertainment
> Gå in i TV
> Välj On

Uppgift E Sätt på telefonens SMS-pip
> Gå in i Phone
> Gå in i Sounds & Alerts
> Gå in i Message alert
> Välj ON

Uppgift F Höj telefonens ringvolym till 2.
> Gå in i Phone
> Gå in i Sounds & Alerts
> Gå in i Ring Volume
> Välj 2
Uppgift G Ändra Warning Distance till normal
> Gå in i MyCar
> Gå in i Warnings
> Gå in i Distance
> Välj Normal

Uppgift H Ändra antalet nycklar till 2
> Gå in i MyCar
> Gå in i About car
> Gå in i No. Of keys
> Välj 2
Appendix E: Questionnaire

ENKÄT 17/5 – 20/5 2005

Besvara varje fråga så utförligt som möjligt. Tänk dock på att detta var en prototyp så den taktila känslan eller detaljer i utseendet inte är viktigt. Vi ser gärna att du inte pratar om detta försök med någon som också ska delta. Fråga gärna om du undrar något!

Skärmen

Vilken av de två prototyperna föredrar du? Varför?

Hur kändes det med fältet till höger när du navigerade i menyn i prototyp B?

Kändes menystrukturen logisk? Varför?

Vad var bra respektive dåligt med menystrukturen?

Hur var feedbacken/återkopplingen i systemet?

Hur var det att avläsa fältet till höger i sista uppgiften?

Reglagen

Hur var det att använda reglagen?

Hur upplevde du placeringen av reglagen?

Övrigt

Övriga kommentarer:
Appendix F: Test instructions for experiment

JAG VILL BÖRJA MED ATT TACKA FÖR ATT DU VILL STÄLLA UPP I VÅRT TEST!

Innan


Provkör en eller flera gånger tills fp känner sig säker. Efter en stund: håll gasen i botten hela tiden, det skall motsvara 60km/h.

Under


Uppgifterna

Nu ska du få öva på uppgifterna. Uppgifterna handlar om att navigera runt i ett menyssystem som är en prototyp, så endast de uppgifter vi har går att komma till.

Pröva att bläddra igenom de befintliga menyerna. Låt fp leka och bläddra i menyerna.

Nu ska vi gå igenom de uppgifter som du ska utföra i två av seten.

Gå igenom de uppgifter som fp ska utföra.

Appendix G: Comments during control survey with users

"Huvudsaken är att hela menyn syns hela tiden så man inte måste komma ihåg de alternativen som är utanför skärmen. Så är det i min saab och det är jobbigt" (man 29)

"Om det ska va vippa måste man kunna trycka in den så att den åker förbi flera alternativ. Man ska inte behöva klicka förbi varje. När man använder vred kan man bara vrida till om man vill iväg långt" (man 24)

"Det känns inte naturligt att klicka enter med ett vred. Enterknappen borde vara bredvid" (Kvinna 28)

"Men pilar så vet man bara, de är bättre än vred. Säkrare med pilar då man kör. Men om menyn är vertikal måste pilarna sitta till vänster och om menyn är horisontal måste de sitta under."

"Horisontal meny med vred är bäst. Så var det förr så det är jag van vid. Känns naturligast" (Kvinna 53)

"Vippen är bäst, spelar ingen roll vilka menyer" (kvinna 27)