Towards optimized instrument panels

Carl Jörgen Normark

Luleå University of Technology

Department of Human Work Sciences
Division of Industrial Design

Luleå University of Technology 2009
Towards optimized instrument panels

Carl Jörgen Normark
At the time of writing this thesis, the worldwide automobile industry is facing the perhaps deepest crisis so far. In the aftermath of a worldwide recession, automobile sales have dropped massively, and several automobile manufacturers are fighting for survival. However, people in the future will still need to transport themselves, but perhaps in a more sustainable way than today. Hence, vehicles will still have drivers, and as long as there are drivers, there will be driver focused research. And it is needed.

This licentiate thesis within Industrial Design was performed within the OPTimized system integration for safe Interaction in Vehicles (OPTIVe) project, financed by the Swedish Intelligent Vehicle Safety Systems (IVSS) research programme.
There are a number of people that deserves some special attention as they have in one way or another helped with making this thesis come true: My main supervisor Anita Gärling for her time, her valuable comments, and for looking after the commas and periods in this thesis. My assistant supervisor Dennis Pettersson. Jan Lundberg for bringing the OPTIVe project to LTU. My fellow PhD students in room F701; Phillip Tretten, Maria Johansson, and Therese Öhrling for the support in both small and large issues. The division of Industrial Design at LTU. Volvo Car Corporation and the OPTIVe people. My cats, which even though they stepped on the keyboard from time to time, managed to not erase one single character. And finally, my lovely wife Monica for enduring long nights in front of the computer.
Abstract

Today, automobiles are becoming more and more technologically complex, with more and more built-in driver information systems. This increases the amount and range of information the driver needs to be aware of while driving. For a safe traffic environment, drivers must have their eyes on the road and hands on the wheel rather than on in-vehicle displays and controls. However, many of these in-vehicle systems are safety systems designed to prevent accidents why there is a constant trade-off between monitoring interior systems and exterior road view for possible upcoming hazards. Therefore, in-vehicle information displays must be designed to be conspicuous enough to be noticed and discrete enough not to distract the driver. In this thesis a user-centred approach is used with the aim to optimize driver information presentation by testing display design guidelines, evaluation methods, and display configurations in order to increase the drivers’ performance regarding noticing, reading, and understanding in-vehicle information displays while driving safely. In Paper I a literature review of design guidelines regarding optimal presentation in in-vehicle information displays was undertaken. The reviewed guidelines appeared to be coherent and valid for today’s automobiles and can, hence, lay the foundation for the design of an optimized display. In Paper II an experimental driving simulator study where 19 participants evaluated two display configuration designs regarding their effect on distraction, noticeability, driver stress, and driving performance during driving while an easy secondary detection task was conducted. The reviewed guidelines from Paper I ensured that results were only effected by the differences between the display configurations. One of the compared display configurations resemble a layout found in automobiles of today, and the other a layout, which according to the literature would improve noticeability, but perhaps be somewhat more distracting. The results showed that the latter display configuration improved driving performance without causing any unnecessary distraction. Papers I and II were mainly focusing on performance based usability, but since image and impression are also part of usability, the aim of Paper III was to study eye-tracking based methods’ appropriateness for evaluating vehicle information clusters, and to connect perceptions of vehicle information clusters with quantifiable measures. Twenty-three subjects participated in the study where a triangulation of eye-tracking, semantic environmental description, and interviews was made. The results indicated that gaze behaviour data adds no additional value compared to if the other assessment methods were used on their own. All together, this thesis brings up important aspects that have implications for the design of in-vehicle information and systems, and gives guidance on how to optimize instrument panels to achieve a safer traffic environment.

Keywords: In-vehicle technologies, information presentation, vision, user centred design, traffic safety, driver distraction.
List of appended papers


<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Aims and research questions</td>
<td>2</td>
</tr>
<tr>
<td>2 In-vehicle technologies</td>
<td>3</td>
</tr>
<tr>
<td>3 Theoretical framework</td>
<td>5</td>
</tr>
<tr>
<td>3.1 Design</td>
<td>5</td>
</tr>
<tr>
<td>3.1.1 Definition of design</td>
<td>5</td>
</tr>
<tr>
<td>3.1.2 The design process</td>
<td>5</td>
</tr>
<tr>
<td>3.1.3 User centred design</td>
<td>7</td>
</tr>
<tr>
<td>3.2 Cognition</td>
<td>8</td>
</tr>
<tr>
<td>3.2.1 Information processing</td>
<td>8</td>
</tr>
<tr>
<td>3.2.2 Vision in driving</td>
<td>10</td>
</tr>
<tr>
<td>3.2.3 Distraction</td>
<td>11</td>
</tr>
<tr>
<td>3.2.4 Guidelines for visual systems in driving</td>
<td>12</td>
</tr>
<tr>
<td>4 Research Approach and Methods</td>
<td>13</td>
</tr>
<tr>
<td>4.1 Literature review</td>
<td>13</td>
</tr>
<tr>
<td>4.2 Synthesis and concept generation</td>
<td>13</td>
</tr>
<tr>
<td>4.3 Quantitative methods in design research</td>
<td>14</td>
</tr>
<tr>
<td>4.4 Driving simulator experiment</td>
<td>15</td>
</tr>
<tr>
<td>4.4.1 Driving performance measures</td>
<td>15</td>
</tr>
<tr>
<td>4.4.2 Secondary task measures</td>
<td>16</td>
</tr>
<tr>
<td>4.4.3 Eye-tracking and glance behaviour</td>
<td>17</td>
</tr>
<tr>
<td>4.4.4 Physiological measures</td>
<td>18</td>
</tr>
<tr>
<td>4.4.5 Subjective ratings</td>
<td>19</td>
</tr>
<tr>
<td>4.4.6 Statistical analysis</td>
<td>19</td>
</tr>
<tr>
<td>5 Summary of appended papers</td>
<td>20</td>
</tr>
<tr>
<td>5.1 Paper I – Display guidelines and principles: A literature review.</td>
<td>20</td>
</tr>
<tr>
<td>5.2 Paper II – Do redundant head-up and head-down display configurations cause distractions?</td>
<td>21</td>
</tr>
<tr>
<td>5.3 Paper III – Evaluation of Car Instrumentation Clusters by Using Eye-tracking</td>
<td>22</td>
</tr>
<tr>
<td>6 Discussion and conclusions</td>
<td>24</td>
</tr>
<tr>
<td>6.1 Methodological considerations</td>
<td>26</td>
</tr>
<tr>
<td>6.2 Conclusions</td>
<td>28</td>
</tr>
<tr>
<td>7 References</td>
<td>29</td>
</tr>
<tr>
<td>Appended papers</td>
<td>40</td>
</tr>
</tbody>
</table>
1 Introduction

The automobile has become a safer mean of travelling the latest 50 years (Lee, 2008). Research and tests have been conducted to improve the passive safety in automobiles, and to reduce damages for drivers, passengers, and other travellers when a crash already has occurred (Blythe & Curtis, 2004). In spite of safer automobiles, traffic fatalities were in the year 2000 the worldwide ninth most common cause of death with about 1.2 million fatalities (Peden, McGee, & Krug, 2002), and the fatalities are estimated to increase with 66% the next 20 years due to the rapid growth of automobiles and drivers, especially in emerging countries such as India and China (Kopits & Cropper, 2005). Also the western world has its share of traffic fatalities. NHTSA (2009a) estimated that about 37,000 fatalities were caused by motor vehicle crashes in USA only in 2008.

Effort in traffic safety research has shifted from reducing damage in a crash to preventing the accident to happen in the first place with the help of active safety systems (Lee, 2008). There has been a rapid growth of in-vehicle systems in recent years where the automobile has gone from holding just the necessary controls for managing the driving, and perhaps a radio as the only comfort oriented system, to be packed with different in-vehicle systems aimed at increasing the driver’s and passengers’ safety, comfort, and enjoyment (Figure 1).

![System trends](image)

Figure 1. The flow of information in the vehicle has increased substantially recent years. © Robert Broström, OPTIVe, 2007.
Major causes of accidents are the human factor and usability related ones where the driver plays the foremost important role in driving safely. Perceptual errors such as in visual search, decision errors related to attention, and operating errors such as speed adjustment have been involved in nearly 90% of road accidents, and have caused the accident in nearly 60% of the cases (Fell, 1976). Hence, there has been an increase in driving safety publications in recent years, reporting that in-vehicle technology either reduces traffic accidents in the form of active safety systems, or causes distractions in the form of, for instance, mobile phones, which can lead to crashes (Lee, 2008). Green (1999) predicted there to be 2,110 injuries in USA in 2007 caused by not looking at the road, whereof 21 would be of fatal outcome. The actual statistics show that 4,704 drivers were involved in fatal crashes due to inattention (NHTSA 2009b), however, these figures include all forms of inattention and not only inattention caused by the vehicle’s own in-vehicle systems as well as all vehicles involved in the accident. These high numbers of fatalities stresses the fact that the automobile and its in-vehicle systems must be adapted to the driver (Bishop, 2005). Moray (1990) proposes the following solution to traffic safety problems:

Even as completely new and advanced systems appear, the principles of information display sampling, and acquisition will remain a key to the design of good transportation systems, be it training the operator, designing the vehicle, or configuring the environment in which transportation takes place. (p. 1212)

1.1 Aims and research questions

A user-centred approach, the Roozenburg and Eekels (1995) basic design cycle, is in this thesis used with the aim to improve methods for determining optimal designs of safe and usable displays in automobile instrument panels as well as to design and test in-vehicle information presentation concepts designed with existing guidelines for optimal information presentation. The specific objectives are:

- Review, evaluate, and compile existing design guidelines and principles for in-vehicle information presentation in order to build a base for further studies regarding optimized in-vehicle information presentation (Paper I).
- Compare and evaluate two different display configuration designs, based on the earlier work in Paper I, in a driving simulator study in the LTU Designlab driving simulator environment (Paper II).
- Investigate the suitability of an eye-tracking (gaze behaviour analysis) based method for assessing in-vehicle information clusters (Paper III).
2 In-vehicle technologies

Automobiles are becoming more and more technologically complex with more and more built-in driver information systems (Baber & Wankling, 1992; Noy, 1997; Tsimhoni & Green, 2001) (Figure 2). The amount of in-vehicle systems is also predicted to continue to increase in the future (Caird & Dewar, 2007). The automobile has transformed from only being a mean of transportation to being a mobile workplace connected to the outside world with internet access and mobile phone conversations (Caird & Dewar, 2007; Lee & Strayer, 2004). The automobile as a mobile office offers increased mobility, efficiency, and comfort for the driver, and since these systems are readily available to the driver, they are most likely to be used while driving and might, at the same time, cause inattention (Lee & Strayer, 2004).

Among modern in-vehicle systems are not only comfort and convenience systems included but also active safety systems such as lane departure warnings, blind spot indicators, and enhanced night vision (Figure 2). Although these in-vehicle safety systems are aimed at improving the driving performance, safe driving is not always improved. The driver can change behaviour and start to rely on the safety system and, for example, in icy winter conditions drive faster when the car is equipped with studded snow tires than when it is not (Rumar, Berggrund, Jernberg, & Ytterbom, 1976). In-vehicle systems, such as the Adaptive Cruise Control, may also allow the drivers to adapt their behaviour and attend less to the driving task (Hoedemaeker & Brookhuis, 1998). Some safety equipment such as seatbelts has, on the other hand, shown great safety improvements as the drivers do not show any adapted behaviour, e.g. reckless driving when wearing seat belts (Evans, Wasielewski, & von Buseck, 1982). Some in-vehicle technologies have the potential to reduce vehicle accidents, for example cruise control, adaptive cruise control (ACC), manual speed alerting systems, and intelligent speed adaptation (ISA), which all can reduce speeding, but only if they are used as they are intended to (Young & Regan, 2007). A safety system must have a good design so the driver is not mislead into false security and a belief that the safety system will prevent all accidents (Lee, 2008).
Figure 2. An example of in-vehicle systems in automobiles today. © Anders Lindgren, OPTIVe, 2007.

Not only specific in-vehicle safety systems can improve driving performance and safety. Recently information presentation technologies such as Head-Up displays (HUD) have made their way into the automobile, which allow the driver to receive information without taking the eyes far from the road. Displays placed lower on the dashboard forces the driver to look away from the road which may lead to incidents (Green, 1999). More glances away from the road generally causes the driver to get “out of the driving loop” more often (Kircher, 2007), which might have negative effects on safety. Using a HUD can also result in reduced workload, decreased response times, and increased driving comfort (Liu & Wen, 2004; Nakamura et al., 2005). The most common HUD technology of today is projected onto the windscreen, but there are also other methods where the information is displayed on a thin film in the windscreen or an opaque LCD display on top of the dashboard.
3 Theoretical framework

3.1 Design

3.1.1 Definition of design

The word design originates in the Latin words “de” and “signare”, which together has the meaning of making something, distinguishing by a sign, give significance, or designate a relation to things (Krippendorff, 1995). The concept of design has a broad and sometimes contradictory definition, but there seem to be some common views such as solving problems, finding solutions, meeting needs, and that design generally refers to a process and that this process is goal-oriented (Friedman, 2003). Design might be considered as one of the more intelligent human behaviours, and there is a strong connection between cognition and design, as design uses cognitive processes for human vision and perception, thought processes, analogical thinking, and memory (Oxman, 1996). A hands-on definition of design which is convenient to use is formulated by Roozenburg and Eekels (1995), where design is “to conceive the idea for some artefact or system and/or to express the idea in an embodiable form.” (p. 53).

3.1.2 The design process

There are several different views on how a design process should be carried out and there are several different models of the design process. However, they all strive for the same result and differences among design processes are to a large extent of a terminological nature. Roozenburg and Eekels (1995) basic design cycle can be considered as one of the most fundamental models for describing design and is a repeating iterative process that consists of the following stages (Figure 3):
Function – Regardless if it is a problem that needs a solution, which can be the case in research, or a function that needs to be realized, which can be the case in design, the function stage is where a discrepancy that needs to be removed is identified.

Analysis – This stage implies formulating the intended behaviour for the new design. Goals in the form of a design specification are formulated as concrete as possible.

Synthesis – The generation of a design proposal. Already known building blocks, components, and separate ideas are combined into one entity by the creativity of the designer.

Simulation – Testing, experimental research, or purely generalizations from past experience is carried out in order to test the expected properties of the designed product.

Evaluation – The properties from the simulation stage are compared to the desired properties formulated in the analysis to be able to make a decision.

Decision – This step decides if the proposed design is acceptably close to the desired product. If this is not the case, an iterative process with incremental improvements takes place, either from scratch where new functions or problems are formulated, or from the analysis or synthesis steps where the proposed design is improved.

Figure 3. The basic design cycle (Adapted from Roozenburg & Eekels, 1995).
3.1.3 User centred design

The amount, and complexity, of in-vehicle systems of today require them to be usable, intuitive, self-explanatory, and not interfere with the driving task. The design of these in-vehicle systems must therefore be shifted towards a user-centred design perspective where users are both in the foci of the design process and involved in its different stages. It is the needs, wants, expectancies, requirements, and behaviour of the user that primarily need to be considered. By making a more user centred system, economic and social benefits can be found as the system helps to protect the user from health and safety risks (ISO 13407). If a system is designed according to user centred design, the system becomes easier to understand and use, the user becomes more satisfied, less stressed, and more productive, and, finally, the system might have a higher appeal to the users which yields competitive advantages (Cherri, Nodari, & Toffetti, 2004). All design should start with the goal to understand the intended users and their different requirements and attitudes towards the technology in question (Shneiderman & Pleisant, 2005). There is also a need to understand how the user might use the system in the future. Usability is a term often used when a user and a system have some form of interaction, and superior usability is often sought for when a system is designed. Usability is here defined as "effectiveness, efficiency, and satisfaction with which a specified set of users can achieve a specified set of tasks in a particular environment." (ISO 9241, §. 3.1). Another definition is that of Han, Yun, Kim, and Kwahk (2000) where usability is the outcome of a user’s actual performance with a product combined with the user’s image and impression of the product. Even though the user plays the foremost important role when driving, the user cannot be perfectly controlled but, designers, on the other hand, have extensive control over the design of the vehicle (Dewar & Olson, 2007).
3.2 Cognition

Driving an automobile is a common daily activity for millions of people but also a complex activity consisting of several tasks that require different but linked skills that rely on visual, motor, and cognitive systems (Graydon et al., 2004). In order to design in-vehicle systems that are well-suited for the user, there has to be a good understanding of how the user is functioning and what human capabilities and limitations there are that needs to be considered.

3.2.1 Information processing

Since Broadbent’s seminal work on information processing was published, it has become important and convenient to construct models of information processing (Wickens & Carswell, 2005). There exist several information processing models, however, in the most simplified model they all share the same structure where the information processing is represented as a sequential series of stages (Figure 4). A driver’s performance is assembled around his or her information processing. The driver must interact with the automobile, attend to information from both the vehicle and the traffic environment, transform this information in order to make it usable, make suitable actions according to this information, and process feedback from the effects of that action (Wickens & Carswell, 2005).

![Figure 4. A simple model of information processing.](image)

The information processing model can be further expanded and adapted to driving according to Wickens and Carswell (2005) as shown in Figure 5. The information processing does not necessarily start at the event or stimuli stage, it should rather be considered as a continuous loop. But for the comprehension of the process, this step makes the most logical starting point. It can also be noted that this framework does not necessarily have to be followed as sequential steps as it has been proven that experts in a task can take shortcuts and jump between steps for a more efficient solving of the task (Rasmussen, 1974; 1976).
Sensory processing - Information from the outside world, for example a vehicle that suddenly crosses the road, must gain access to the driver's brain and consciousness, either through visual, auditory, or tactile receptors in a sensory processing stage. If these receptors are impaired by for example looking away towards something else, or listening to loud music instead of an important message, the information might not get through.

Perception - The information from the sensory receptors must be interpreted and given meaning in the perception stage. A suddenly emerging object on the road or a sudden loud sound is often perceived as danger. These perceptions can be different whether they are inferred by lower channels of neural information, which can be the case in unfamiliar circumstances (bottom-up processing), or by past experience, which can be the case when something is highly expected or when the sensory processing is impaired (top-down processing).

Response selection - This stage is driven by conscious activities and requires some mental effort or attention to transform and compare the perceived information with previous knowledge stored in memory. A suitable action or response can then be chosen. In the case of a hinder on the road, a correct choice would involve something that avoids a crash.

Response execution - This stage implies by muscular activity execution of the action chosen in the previous stage, which in the case of avoiding a crash might be to hit the brakes or to make a manoeuvre to avoid the vehicle.

Feedback - The feedback loop is required for the driver to know if the right action was taken, or if another action needs to be taken as a consequence of the previous action. The information processing loop is therefore a continuous activity.

Figure 5. A model of human information processing. (Adapted from Wickens & Carswell, 2005).
Attentional resources – This is an essential component of this model and concern three stages that are for the most part not automated. Selective attention can influence which information is attended to, and passed on, to the next stage. Past experience can influence the response selection, and be a cause of errors if the environment and events are not carefully attended, and some events may never get to the perception stage if the visual channel, that is important for driving, is impaired.

3.2.2 Vision in driving

Driving makes intense demands on the driver’s visual perception (Dewar, Olson, & Alexander, 2007) and the driver is mainly guided by vision for accomplishing safe driving (Hills, 1980; Wierwille, 1993). Therefore, visual behaviour in traffic is of great value when designing in-vehicle systems. A visual task while driving effects the driver more than an auditory task (Recarte & Nunes, 2000; Wierwille, 1993) and impairment of the visual information processing channel is considered a main cause of accidents (Baumann, Keinath, Krems, & Bengler, 2004). There are also other aspects such as auditory and haptic factors that can improve and/or degrade driving performance. Everything that can distract drivers from the driving task with eyes-on-the-road and hands-on-the-wheel must be investigated with regard to traffic safety (Baumann et al., 2004).

According to Moray (1990), unconscious and automatic habits largely control the visual attention. The visual attention needed by an in-vehicle system should therefore be minimized so the driver can pay more attention on the road while, for example, turning on the radio (Dukic, Hanson, Holmqvist, & Wartenberg, 2005). Visual information can be lost in the quantity of information displayed to the driver but also have a negative effect on driving safely as it competes with the same perceptual and cognitive resources as the task of driving (Horberry, Anderson, Regan, Triggs, & Brown, 2006). Important information, such as warnings, can be lost in the vast quantity of information constantly displayed to the driver as the driver has limited visual resources (Wickens, 2002). The driver’s ambient vision supports lane keeping, and the focal vision is used for detecting critical events on, or along, the road (Wickens, 2002). Both these aspects of vision are used when driving and simultaneously managing a secondary task. Several studies on visual behaviour have proved that secondary tasks can affect the driving performance and increase the driver’s visual and mental workload (Alm & Nilson, 1995; Recarte & Nunes, 2003; Zwahlen, Adams, & DeBald, 1988).
3.2.3 Distraction

It has become more difficult for automobile drivers to accurately allocate attentional resources the last few years since they nowadays are exposed to a higher risk of becoming overloaded due to more traffic, more advanced in-vehicle display designs, more traffic signs, more mobile phone conversations and other distracters (Dewar et al., 2007). Distraction has emerged as one of the most important factors for traffic safety (Lee & Strayer, 2008). Up to 25% of accidents can be caused by distraction and the increased use of different technologies being placed in the automobile can cause some of these distractions (Horberry, et al., 2006), which can lead to worse driving performance and, hence, decreased safe driving. Even though there already is a vast quantity of in-vehicle systems, traffic safety depends highly on behaviour and the emerging problem of distraction requires systems that improve the drivers’ behaviour (Lee, 2008).

From a driving safety point of view, it is better to keep the eyes on the road. A short glance away from the road which might seem quite harmless can make the driver overlook safety related issues, specifically on the side of the road (Fisher et al., 2002). Even if crashes often occur because the driver fails to look at the right place at the right moment, they can still occur when the driver is watching the road. When adding a secondary task, such as controlling an in-vehicle system to the task of driving, the increased cognitive demand can cause drivers to overlook safety critical events not only on the side of the road, but even in a close perimeter on the road (McCarley et al., 2004; Strayer, Drews, & Johnston, 2003). Driving is most of the time a manageable task for most drivers, however, there can be more intensive periods where the demands of driving leaves no spare attentional capacity why even simple secondary tasks can interfere with the driving and lead to incidents. Also, when the drivers feel safe and confident in driving on less demanding roads, they often choose to perform some kind of secondary task such as dialling mobile phones or changing radio channel, even if the detection of unexpected events might be seriously slowed down during this secondary activity.

Victor, Harbluk, and Engström (2005) state that the complexity of a visual task is highly correlated to the time spent looking at the display. Increased visual clutter in the driver’s environment can also cause high cognitive workload (Horrey & Wickens, 2004), and cognitive load can also cause the driver to neglect highly salient safety related information (Lee, Lee, & Boyle 2008). Two examples of inattention that might be caused by high cognitive workload are inattentional blindness where stimuli presented in the observer’s field of view are missed, and change blindness where the observer misses to detect changes in observed stimuli (Varakin, Levin, & Fidler, 2004).
3.2.4 Guidelines for visual systems in driving

In the strive for a safer traffic environment, guidelines for how the vehicle interior and in-vehicle systems are allowed to effect the drivers visual behaviour have been formulated. Even a brief glance away from the road can infer a missed signal, and can lead to delayed braking and increased crash risk (Zhang, Smith, & Witt, 2006). Having attention on something else than the road is one major factor that can cause danger on the roads. If for example a driver’s attention leaves the road for more than 2s, the risk of having an accident is considerably increased (Zwhalen et al., 1988). According to Wierwille, Antin, Dingus, and Hulse (1988), a glance towards the vehicle interior is on average no longer than 1.6s. If glances away from the road are increased from 1s to 2s, the risk for lane departures is increased 3.6 times (Green, 1999). Horrey, Wickens, & Consalus (2005) and Wittman et al., (2006) state that if information is acquired from a display located close to the road view, the driving performance is less degraded than for a distant display. According to Summala, Lamble, and Laakso (1998), the detection of signals is easiest near the line of sight and decreases significantly for larger eccentricities, especially vertical. Since signals such as warnings are not sought after regularly by people, warnings must in some way stand out from other information and call upon the user’s attention in order to provide the required safety effect (Laughery, 2006).
4 Research Approach and Methods

Since the aim of this thesis is to initiate the work of determining optimal designs of safe and usable displays in automobile instrument panels, there are several research methodologies used.

4.1 Literature review

According to Webster and Watson (2002) a literature review is a good way of analysing previous knowledge and comparing different views and results. A literature review is a study of past research, which should cover the relevant literature and not be too confined regarding research methodology, journals, or geographical areas. A literature review represents the cutting edge of the research area, but also identifies gaps of knowledge regarding concepts and comprises a map for future research within the topic. The identified gaps of knowledge are essential when further research is planned.

4.2 Synthesis and concept generation

In order to solve problems regarding traffic safety, ideas and previous knowledge must be taken from several areas, such as drivers’ expectancies and design guidelines, and be combined into one or several testable concepts. According to Roozenburg and Eekels (1995), the word synthesis means the combining of objects, components, and ideas into something new. The synthesis phase in a development process is crucial to the entire outcome of the process as this is where the description of an idea comes to form several concepts of which one could be a possible solution to a problem. The synthesis phase involves an iterative process which leads to the elimination of eventual uncertainties of the concepts until an acceptable solution is found. According to Ulrich and Eppinger (2003), a concept is a description of how the user’s needs are to be fulfilled. Concept generation is generally an inexpensive activity where changes can be
made quickly. A concept generation process generally begins with the users’ needs and the desired specifications of the solution, and ends with a set of concepts of which one will be selected as the final solution. According to Roozenburg and Eekels (1995), the synthesis or concept generation phase is highly dependent on human creativity for a successful outcome and there are no strict methodological rules that need to be followed. Ulrich and Eppinger (2003), however, suggest some steps that can be followed: 1. Divide the problem into smaller sub problems and focus on the most critical ones. 2. Search externally for solutions, a conventional solution can be combined with a novel solution and form a superior design. 3. Use personal and team knowledge together with the existing battery of creativity methods to generate solution concepts. 4. Explore concepts systematically to find the desired solution. 5. Finally, take some time to reflect on the process and the results.

4.3 Quantitative methods in design research

The use of quantitative methods and statistical analysis seems at first sight to be a contradiction to the freedom and creativity within the design field, but according to Purpura (2003), quantitative methods where incremental improvements on practical issues are a complement to qualitative methods in design are a good way to go. Qualitative research can help to broaden the area of research with more angles of approach, while quantitative research can help to reduce the complexity of the researched problem. Quantitative research can be used to find the right audience, to understand their needs, and to test the usability of the technology so the designer can make the product more appealing to customers. Quantitative research is often used in the development and refinement stages of technical solutions, and through experiments the “optimal” solution among several possible solutions can be found. Often, this comes down to improve details in an iterative design process where incremental gains are made. Even if these incremental gains are small, they can add up to the difference between a safe ride and a crash for the automobile driver.
4.4 Driving simulator experiment

When in-vehicle systems are researched and evaluated, it is often a good idea to do so in the correct context, i.e. while driving. Even though research on in-vehicle systems is aimed towards making a safer traffic environment, conducting research in real traffic could be highly unsafe since much research is about finding limits for the driver’s comprehension. A driving simulator offers a safe method of conducting driving like studies at a low cost. Another advantage with the driving simulator lies within the controlled test environment; experiments become very predictable as all participants can be exposed to exactly the same situation, while real traffic is highly unpredictable. Researchers have full control over the situation and effects of specific conditions can be isolated, which makes a driving simulator suitable for evaluating in-vehicle systems and information (Carsten et al., 2005). However, there is an ongoing debate whether findings in a driving simulator study can be considered valid in real life traffic. A driving simulator environment is always simplified compared to a real traffic environment, regarding visual and auditory realism as well as feedback from motion. Drivers in a simulator also tend to drive more recklessly than normal as they do not experience the situation as dangerous at all (Cherri et al., 2004).

If the driver is visually or cognitively loaded, the driving performance can be affected in the form of reduced speed or decreased lane keeping ability, or in the form of reduced situation awareness as in weakened perception and judgement of the traffic situation (Östlund et al. 2004). For distractions and workload measurements, there are broadly five categories of suitable techniques: 1. Primary task or driving performance measures such as steering performance. 2. Secondary task performance such as the completion time of a secondary task while driving. 3. Glance behaviour. 4. Physiological measures. 5. Subjective ratings (Tsimhoni & Green, 1999).

4.4.1 Driving performance measures

There is a number of driving performance measures that have proved to be highly diagnostic, both in simulated driving and in real traffic. The driving performance is an indirect measure of traffic safety since great variations in lane keeping or speed can impose danger on the road. Speed is generally the most important measures in Advanced Driver Assistance System (ADAS) and In-Vehicle Information System (IVIS) studies since there is a strong relation...
between increased speed and increased accidents (Roskam et al., 2002). Speed measures are also simple to both measure and compute (Johansson et al., 2004). Some of the most used speed measures within research are:

- **Mean speed** is considered to be the most diagnostic speed measure since using an in-vehicle information system generally leads to decreased speed in order to reduce the demand from driving resources (Östlund et al., 2004).

- Also, **maximum speed**, which is the single highest measured speed and **standard deviation of speed** are amongst the most common driving performance measures. Increased variation in speed can be an indication of decreased driving performance, however, this measure does not take into consideration normal speed variation due to adaptations to road conditions or other traffic (Roskam et al., 2002).

- **Standard deviation of lane position** is one of the most commonly used performance measures. It has high face validity and is simple to compute. Task duration has, however, a strong influence on this measure (Johansson et al., 2004), but it is relatively independent of the speed (Godthelp, Milgram, & Blaauw, 1984).

- **Lane exceedences** are good safety indicators because of the higher face validity compared to lane variation measures such as standard deviation of lane position. However, lane exceedences may be insensitive to small differences in workload and distraction (Johansson et al., 2004). A lane exceedence is here defined as a major lane deviation where more than half of the vehicle is outside the lane. The number of exceedences is then computed and analyzed (Liu, Schreiner, & Dingus, 1999).

4.4.2 Secondary task measures

Event detection is highly correlated to traffic accidents and has strong safety relevance. Detection times can both be related to the driving task, i.e. detection of sudden events on or along the road, and to secondary tasks inside the vehicle. This measure is often used to evaluate visual and cognitive aspects of secondary tasks and is equally sensitive to both kinds of distractions (Johansson et al., 2004). Task completion time, or response time, is defined as the time from a clearly defined stimulus in the vehicle is presented to the time the driver responds correctly (Wierwille et al., 1996).
4.4.3 Eye-tracking and glance behaviour

It has been proved that eye movements have great impact on traffic safety since glance behaviour in the form of fixations and saccadic movements give valuable information about drivers’ perceptual processes (Treat et al., 1977 as cited in Shinar, 2008), attention and cognitive processes (Rayner, 1998), visual search and hazard recognition (McKnight & McKnight, 2003), and information acquisition process (Rockwell, 1972). Based on drivers’ eye data, placements and functions in in-vehicle systems and safety systems can be optimized. However, caution must be taken when analyzing eye-data; the eyes always fixate somewhere, but a person’s attention does not always have to lie on the fixated object. Many traffic incidents origin from the “looked but did not see” phenomenon, where the mind is on something completely different than the traffic environment (Stutts, Reinfurt, Staplin, & Rodgman, 2001).

Eye data usually comprises of several different measures:

- **Glance duration** is defined as the time from when a person directs his or her gaze toward a specific object to the time the gaze is moved away from it (ISO 15007). Long glance durations can indicate high workload or high visual demand for the participant (Johansson et al., 2004). According to Wierville and Tijerina (1998) and Dingus et al., (2006), increased in-vehicle glance duration is associated with an increased crash risk.

- **Glance frequency** is defined as the number of glances toward a specific object during a task or during a pre-defined time period. This measure is highly correlated to glance duration since low glance frequencies can indicate long glance durations (ISO 15007).

- **Total glance duration** is the total glance duration towards a specific object. This measure can indicate the visual demand or cognitive load caused by that object (ISO 15007). Percentage of time spent on different areas has proven to be useful in other studies (Harbluk & Noy, 2002; Victor et al., 2005).

- **Mean glance duration** is highly correlated to the difficulty of a visual task (Östlund et al., 2004; Johansson et al., 2004). Mean glance duration is defined as the mean of total glance duration towards a specific object.

- **Total time off road scene ahead** is defined as the total time the driver’s gaze is directed towards something other than the road (Johansson et al., 2004). For a safe traffic environment, drivers must have their attention on the
outside of the vehicle rather than on in-vehicle displays (Baber & Wankling, 1992).

- Event detection related to eye-tracking and glance behaviour is also measured. The same can be said about this measure as with the secondary task measure; event detection is highly correlated to traffic accidents and has strong safety relevance. Time to notice is here defined as the time from a clearly defined stimulus in the vehicle is presented to the driver’s gaze is shifted towards this stimulus.

4.4.4 Physiological measures

It is widely accepted that using several physiological measures of workload is better than any single one. Several studies have also shown that a combination of physiological measures and subjective ratings to utilize the strengths of both methods are preferred (Wilson & Eggemeier, 1991). Some of the most used physiological measures within research are:

- Galvanic skin resistance (GSR) measures the electrical resistance between two positions on the skin. GSR, or the variations of the similar technique electrodermal activity (EDA), has been extensively used in research of various effects such as, for instance, arousal but also on workload and stress (Kramer, 1991). The level of skin resistance is highly affected by stress and arousal. Skin resistance can indicate perceived risk and complexity of traffic situations and the mental effort related to task demand (Helander, 1978). However, interpersonal differences and artefacts can make skin resistance hard to analyze (Östlund, Nilsson, Törnros, & Forsman, 2007).

- Heart rate is the perhaps the most frequently used measure when it comes to mental workload. Heart rate can be considered as a good measure of information processing activities such as arousal and stress, which are good indicators of workload (Kramer, 1991; Wilson & Egggenmeier, 1991).

- Hand temperature can also be used as a measure for anxiety and stress but tends to reflect more of body relaxation. When there is tension in the major muscle groups, blood circulation decreases and the hand temperature drops (Kappes & Michaud, 1978).
4.4.5 Subjective ratings

Subjective measurements comprise an addition to performance based measurement techniques, especially applicable in multi-task situations (Eggemeier & Wilson, 1991). Subjective ratings are often used for their high face validity, ease of use, low costs, and no intrusion on the task at hand. However, there are also downsides such as that subjective and objective measures can have a low correlation and the participant might be confused and have trouble distinguishing external demands from actual efforts (Johansson et al., 2004). The results from subjective ratings in this driving simulator study are presented in Tretten, Normark, and Gärling (In press).

4.4.6 Statistical analysis

Statistical analyzes are used to reveal significant differences between different conditions and to determine how likely that these differences are not caused by chance. Normally t-tests or F-tests are used if the analyzed data are normally distributed and if not non-parametric tests such as, for instance, Mann-Whitney’s U-test are used for between subjects’ conditions, and a Wilcoxon signed-rank test for within subjects’ conditions (Greene & d’Oliveira, 1999). The non-parametric tests can better handle outliers, why these does not necessary have to be removed or transformed unless they are clearly erroneous (Howell, 2002).
5 Summary of appended papers

5.1 Paper I – Display guidelines and principles: A literature review

There exists a vast amount of guidelines and principles for how in-vehicle information should be presented optimally. However, many technologies have emerged since the formulation of many of these guidelines. The aim of this study is an attempt to search for, and gather, existing guidelines in a guideline compilation as well as to find discrepancies among guidelines and to evaluate their validity for the new automobile technologies that have emerged in recent years. A manual literature search in published handbooks, other guideline compilations, and other published material was where appropriate performed. A total of 26 handbooks and guideline compilations were consulted.

Guidelines were predominantly found in, and categorized under, the following categories: eye and head movements for visual displays, types of displays suitable for different purposes, dials, characters, colours and illumination, data density, graphics, warnings, and attention.

In general, the reviewed guidelines seem to be applicable for information presentation in the automobile of today since they can be employed regardless of the technology used in the automobile. There are, however, many high-level guidelines merely stating “the way it should be” without giving any guidance about how to get there. Most guidelines seem to be coherent and point in the same direction and only minor discrepancies were found such as, for instance, whether the use of graphical boundaries increase or decrease user performance. Another step to take from here can be to validate these guidelines in driving situations.
This paper describes a driving simulator study designed to investigate differences in two display configuration designs that could affect safe driving; information presented redundant in two displays near the driver’s line of sight, and information spread out in four different information displays in the automobile. Driving an automobile is a highly visual task (Hills, 1980), and since automobiles are becoming more and more technologically complex, with numerous built-in driver information systems, there are high demands on the driver’s visual load. For a safe driving environment, there is some proof that drivers should have their attention on the road rather than on in-vehicle information displays (Baber & Wankling, 1992). There is now a technology that could facilitate reduced workload, decreased response times to information, increased driver comfort, and, at the same time, display information necessary to the driver, namely the head-up-display (HUD) (Liu & Wen 2004). Safety critical information, such as warnings, must not be overlooked by the driver. Therefore, a redundant display of warnings in both the regular automobile instrument cluster and in the HUD could facilitate the noticeability and interpretation of warnings (Ellis, 2005). However, presenting highly salient information in the driver’s line of sight might have a distracting effect, which could deteriorate the driving performance. The aim of this study is to investigate the impact of display designs on driving performance and glance behaviour.

Nineteen participants completed a driving simulator study in the LTU Designlab driving simulator. Driving performance, glance behaviour, physiological measures, and task completion times were measured for two display configuration designs both during driving only and during driving together with a simple secondary task, which consisted of detection, and off-setting, of ten presented warnings. In the redundant display configuration, warnings were presented redundant ten times in the same two displays. In the spread out display, warning presentations altered between four displays.

The results showed that there were some differences between the two display configurations. When the task was to only drive the simulator, the redundant display configuration showed better speed maintenance than the spread out display configuration, probably an effect of the speedometer displayed in the HUD. When the task was to off-set warnings while driving, the redundant display configuration showed better driving performance and glance behaviour than the spread out display configuration. Hence, the use of a HUD does not distract the driver but enhances the driving performance. The different display
configurations did, however, not affect the physiological stress measures. The stress caused by the display configurations might have been too small in comparison with the stress caused by the driving task itself. In this setting, it seemed that the physiological measures used were not suitable for evaluating in-vehicle systems.

5.3 Paper III – Evaluation of Car Instrumentation Clusters by Using Eye-tracking

The types of driver information and safety systems that are present in automobiles of today are well known by researchers and automobile manufacturers. In recent years, many articles have been published regarding driver distraction, usability, traffic safety, and similar areas. Although, the massive research effort in these areas, there are still gaps of knowledge that need to be filled such as, for instance, the lack of knowledge regarding the driver’s perception of in-vehicle systems as well as their appearance. Usability is, according to Han et al., (2000), an outcome of a user’s actual performance with a product and the user’s image and impression of the product. Eye-tracking is a method for studying what a person’s gaze is upon, however, eye-tracking alone does not give any clues about how an object is perceived. The aim of this study is an attempt to develop a method for studying users’ perceptions of appearance based on gaze behaviour. Twenty-three participants completed a study where a triangulation of three methods were used; eye-tracking, assessments of viewed instrument clusters, and interviews. The subjects viewed a total of eight static pictures of automobile instrument clusters displayed on a computer screen and were asked to assess them according to a modified semantic environmental description scale (SMB) (Küller, 1975) and a modified product semantic analysis (PSA) (Wikström, 2002) evaluation method. After the assessment task was completed, an interview was conducted. The participants were shown their gaze patterns and gaze frequencies from the eye-tracking together with the instrument cluster’s assessment scores, and were asked to describe why specific areas of the clusters were gazed upon and why the specific assessments were made.

The results showed that there were no clear connection between how an instrument cluster is viewed and how it is assessed. There was a problem with identifying details that caught the participants’ attention as the pictures seemed to have been viewed in a more holistic way, therefore, the connection between gaze patterns and interviews were somewhat vague. The eye-tracking could, however, show that, in general, the participants viewed the instrument clusters from top-left to bottom-right, the same order as in the western reading direction.
This could give some guidance for automobile designers in where to place critical information that quickly needs to be comprehended to enhance safety. The gaze-data could also be an aid in the interviews as the participants do not have to recall exactly what was gazed upon at certain times.
6 Discussion and conclusions

The main aim of the thesis was to take a step forward towards designing optimal instrument panels in automobiles, through utilizing a user centred design cycle, where safety, driver performance, and usability regarding information presentation in the automobile were considered. The introduction states the function, or the problem, that needs a solution, namely traffic fatalities in some way caused by the vehicle interior. The literature review comprises an analysis of how this problem could be solved followed by a synthesis of some information display concepts which were evaluated based on the results of a simulation.

The literature study in Paper I showed that there exist a large amount of previously stated guidelines for information presentation and that they, mainly, seem to be coherent with each other. Since the guidelines mostly stated information presentation principles, they seem to be valid regardless of the technology used, and can, therefore, also be applied to today’s automobiles even if the information flow and quantity have increased considerably since the formulation of most of the guidelines in this review.

Based on the findings in Paper I, two information display configurations were designed and tested in the experiment in Paper II. One display configuration represented a layout that can be found in automobiles of today where information is spread out on different displays, while the other presented information redundantly in a head-up display and a head-down display. The latter was thought to have better noticeability, but to be more distracting. According to Wickens and Carswells (2005) information processing model, the sensory processing and perception stages together with the driver’s attentional resources lays a foundation for choosing and performing a correct response. Information displayed to the driver must therefore be positioned in a way that the visual sensors are able to detect the information quickly even if the driver is not currently focused on the source of information. The information must also be designed in a way that it can be perceived and comprehended the right way and at the same time not cause long glances away from the road as well as comply with the driver’s attentional resources even in busy traffic situations. The results showed that it is possible with a user centred approach to design a display configuration that does not cause unnecessary distraction during a secondary task while driving. More specific, the results show that the redundant display configuration yielded better driving performance as well as secondary task
performance without distracting the driver, which is in line with earlier research (Horrey et al., 2005; Wittman et al., 2006). The redundant display configuration yielded more gazes towards the displays compared to the spread out configuration. However, it also yielded more gazes towards the road, which should imply safer driving. Looking around in the vehicle interior for information is more time consuming than having all information presented near the driver’s field of view, which might have been the reason why the head-up display was the preferred display to consult in the redundant display configuration. This might, however, be effected by the highly salient and novel head-up technology and the drivers’ inability to disregard the information displayed in front of them. Even if this was the case, it strengthens the findings that this redundant display configuration did not distract the driver. The secondary task that was performed during the experiment was very simple, but could anyway cause differences between the two display configurations, which make it rather easy to evaluate and compare safety effects between different in-vehicle systems. Driver stress was also monitored during the experiment, but the levels of stress was indifferent for the two display configurations, which might indicate that stress, in this case, was caused by the driving situation itself and not by the different display configurations.

As stated in Paper II, further studies in this should be extended with a more difficult secondary task than the simple detection task used here. This might give more knowledge about how display configurations effect the response selection and execution stages of the information processing model and how the secondary task effects the usability of the information display configurations, and if the differences stem from the two information presentation concepts or from the task itself.

Paper I and II were more focused on the “hard” technically oriented design guidelines for information presentation in automobiles and driving performance, even though the user was still in focus, while Paper III focused more on the “soft” side of usability and the users’ perceptions of in-vehicle information presentation. In Paper III, an eye-tracking based method of evaluating in-vehicle information clusters was developed and tested. This was an attempt to couple users gaze behaviour to their perceptions of the viewed product but the results did not show any clear correlations between these aspects. There were, however, some significant differences for some measurements regarding in which order the clusters were viewed. However, since some viewed clusters had widely spread assessments and gaze behaviours, more participants would have been needed to ensure valid results. It would have been of great value to designers if the presented method had proven to fill its purpose, but the results can only give an indication of the uncertainty of the presented method. It was obvious that eye-tracking alone cannot give information about how an item is perceived. The addition of other assessment methods such as semantic environmental descriptions
and interviews did not increase the value of the combined methods compared to if the methods had been used on their own. Therefore it is suggested that the perception and the perceived usability of in-vehicle systems should only be collected by interviews, which also was done in Tretten, Gärling, and Pettersson (2009).

6.1 Methodological considerations

The literature review was mainly based on human factors handbooks and other similar publications containing design guidelines. Even though this probably was the most legitimate source of guidelines, a more comprehensive view of the research field would have included published journal articles since results of the recent few years of research have probably not made its way into handbooks yet.

The concept generation prior to the driving simulator experiment would probably have had benefits from more iterations of the concept generation and also the utilization of more creative methods. Even though experimental control was maximized and complexity reduced since of the concepts in this study were merely put together by the design guidelines, some additional improvements in the pursuit of the “optimal” instrument panel would perhaps have been made with the use of creative methods. There were some user involvement in the concept generation phase by participants in pilot tests, where many valuable reflections were attended to in the next step of the iterative design process. However, participants should perhaps have been even more involved in the concept generation phase to get more input about their needs and requirements.

A driving simulator study generally offers high internal validity through high experimental control, but since a driving simulator generally lacks ecological validity, conclusions drawn in simulator studies are not necessary applicable to real life situations. Effects from the researched conditions in this study should have parallels to effects on driving performance even in real traffic. The intention of the simulator study was, to make the performed tasks, the driving situations, and the traffic environments as realistic and natural as possible. Even though this might have increased the ecological validity, this was also a trade-off where experimental complexity increases and experimental control decreases. With a simpler experimental setting the internal validity should have been higher and effects caused by specific aspects in the conditions could have been better isolated.
One major drawback of the simulator study is that the vehicle handling of the simulator is different compared to a real automobile. Thus, it takes some time to get used to driving the simulator. The participants in this study only had a few minutes of practice, or until they felt confident in driving the simulator, so more practice time would have been needed to get better and smoother driving performance results. More practice time would also be needed to minimize the differences between novice participants, and participants that were used to applications similar to a driving simulator, e.g. computer games, in order to gain a higher external validity.

Since a driving simulator was used the ethical considerations that arise in real traffic were avoided. However, it has been shown that driving a simulator has some kind of simulator sickness related effect on up to 61% of the participants (Kolasinski, 1995). Driving a simulator is substantially different from driving an automobile on the roads, especially as the simulator used in this study was a fixed base simulator. Since all participants in the study had a valid driver’s licence, they were to some extent experienced with driving and were familiar with how it feels to drive an automobile. Since the simulator had a rather narrow angle of forward road view and was of a fixed-based type without motion cornering accelerating and braking, the participants’ knowledge of how driving should feel and the experience of the simulator were conflicting with each other which might have lead to the experienced simulator sickness. The study should have begun with a screening process for simulator sickness, which could have spared some participants from discomfort. It can also be noted that in USA, participants in driving simulator studies are not allowed to drive their cars until one hour after a completed simulator study. The participants in this study had, however, no restrictions in driving after completing the study.

As in many studies it would also have been good to have more participants to ensure reliability and generalizability. The development of an eye-tracking based method for evaluating perceptions also had a small and rather homogenous test group; all participants were students in either a design or a psychology programme at Luleå University of Technology, and the number of participants was quite few why generalizability and external validity is quite low. However, the results in this study would probably have remained unchanged even with a higher number of, and more diverse, participants as the results showed no indications of a correlation between image or impression, and gaze behaviour. The statistical analyzes would have had higher statistical power if parametric tests would have been applicable. Unfortunately, the data was not normally distributed, which required non-parametric tests with repeated separate analyzes for the within-subject factor as there are to the author’s knowledge no corresponding non-parametric test that could reveal any interaction effects between different conditions.
6.2 Conclusions

To conclude, research in traffic safety have to continue since there still are issues to deal with, especially in the driver’s nearest environment due to that the rapid technological development pushes more and more systems into the automobile. The results from these studies, give some indications of what might be possible to design with respect to:

- Driver information layout that works well with the driver by utilizing existing design guidelines.
- How to place driver information to make it noticeable without being distracting.

First, the reviewed design guidelines seem to be applicable for information presentation in the automobile of today since they are mainly coherent with each other and can be used regardless of technology used in the automobile. Secondly, it was shown that a centrally placed redundant presentation of information in a head-up and head-down layout yielded better driving performance compared to a layout spread out on four different displays. The redundant display layout did not distract the driver even though it had a highly salient placement. Thirdly, it was shown difficult to couple the image/impression dimension of usability to gaze behaviour, hence, this aspect might be more suitable to study with the help of subjective ratings. Finally, it can be said that even though this work can add some knowledge to field of traffic safety, more research always needs to be undertaken to improve the lives of all travellers on the roads.

For future considerations regarding the possibilities of the driver-automobile interface, the automobile with all of its safety and comfort systems, is more and more transforming into something that can be compared to a personal computer, displaying whatever the user desire for the moment. What we do not know, however, is what drivers’ really want to make use of when they drive their automobile, and if the vehicle’s systems can be adapted to the individual user’s needs. Neither do we know how the interaction of adapting the vehicle will take place, and what the safety consequences will be, but this thesis can be seen as a starting point for such an endeavour.
7 References


ISO 13407:1999 Human-centred design processes for interactive systems.

ISO 15007-1:2002 Road vehicles -- Measurement of driver visual behaviour with respect to transport information and control systems -- Part 1: Definitions and parameters.


Display Guidelines and Principles:

A literature review

Carl Jörgen Normark, Division of Industrial Design/LTU DesignLab, Department of Human Work Sciences, Luleå University of Technology, 971 87 Luleå, Sweden
Anita Gärling, Division of Engineering Psychology, Department of Human Work Sciences, Luleå University of Technology, 971 87 Luleå, Sweden

As more and more technology is added to the automobile interior, it needs to be designed in a usable and efficient way to facilitate safe driving. This paper reviews guidelines and visual design principles for automotive instrumentation and for visual information presentation in general. Guidelines are compiled, categorized, and analyzed in order to determine whether they still are valid and usable for today’s design of information presentation in automobiles. By doing this, contradictory guidelines and gaps of knowledge regarding automobile information presentation were identified.

Keywords: automotive design, interdisciplinarity, specification, interface design, literature review
As more and more technology is added to the automobile and the driver environment, the question arise whether these driver help and infotainment systems can be used without causing unsafe driving. Driver distraction has shown to be a large cause of accidents; in 25 % of all crashes distraction have been the main reason (Horberry, Anderson, Regan, Triggs, and Brown, 2006). However, in today’s automobiles more and more advanced systems are implemented which might take away attention capacity from the main focus of the driver; e.g. to drive as safe as possible. Efforts must be made to ensure that in-car systems do not have negative effects on the driving task. One way to do this might be to make sure that all new systems are presented and visualized in the best way possible.

Despite the large number of existing guidelines and principles for how information output should be designed and presented, there are to the authors’ knowledge no wide-ranging compilations of these guidelines or any comparisons and analyzes of these guidelines. The aim of this literature review is to analyze existing guidelines and principles for visual displays and information presentations in general and for automotive instrumentation presentation as well as for automotive ergonomics regarding visibility especially. Many of the present guidelines originate from a time when there was a considerably lesser amount of information presented inside the automobile. Are these guidelines for designing information displays and instrumentation still valid for today’s automobiles? The areas of interest are placements and technological restrictions for information displays, limits of human perception, and issues regarding understandability and usability.

The following research questions will be answered:

- Are the existing automotive design guidelines relevant for today’s automobiles?
- Are there contradictory results among guidelines?
- Are there areas that require further research?

1 Method

A literature search was conducted in June 2007 using the key words display, design, and guideline in the Compendex, Inspec, and Referec databases which returned 936 records. A refined search with the same keywords in the ISI Web of knowledge in the areas of ergonomics, psychology, applied psychology, and computer science resulted in 288 records. To narrow down the amount of articles a manual search in published handbooks, guideline compilations, and of other cited material was where appropriate performed. In all, 26 handbooks and guideline compilations were consulted. The search included only publications in the English language.
2 Results

2.1 Eye and head movements for visual displays

The area of eye and head movements for visual displays is well examined and limitations of head movements are first considered since they largely constrict where information in the driver's compartment can be placed. Diffrient, Tilley, and Bardagjy (1974) state that an effortless head movement can be done within 45° in the horizontal plane and 30° within the vertical plane. According to SAE (1994), an easy vertical rotation of the neck is done within 45° while the movement should not exceed 60°. According to Chaffin, Anderson, and Martin (1999) the limit should be within 20-30°. The movement of the head narrows the space for information presentation. Limitations of eye movements give further restrictions for where frequently needed, and important, information should be placed. An effortless rotation of the eye should, according to Chengalur, Rodgers, and Bernard (2004) be within 30° around the sightline. This value agrees with rotations in the vertical plane stated by SAE (1994), and an easy eye rotation downwards stated by Diffrient et al. (1974). The maximum downward rotation according to Diffrient et al. (1974) is merely 35°, while SAE (1994) suggest 60° down and 45° up (Figure 1). All previously mentioned authors agree that the eye easier could be rotated downwards than upwards.

![Figure 1 Head rotation and eye rotation](image)

Both head and eye movements constitute the line of sight. The normal sightline when seated in an automobile is according to Diffrient et al. (1974) 15° down while Grandjean (1983) reports a range of 4-14°. According to Diffrient et al. (1974), this is also the minimum limit for the visual area above the steering wheel. When all limitations of head and eye movements are considered, guidelines for limits of where different information should be placed can be put together. Diffrient, Tilley, and Harman (1981) state that primary displays should be located within 15° around the normal sightline. Van Cott and Kincade (1972) add that this area is a little oval and more flatter above the line of sight than below and that the displays should be placed perpendicular to the line of sight. Chengalur et al. (2004) also suggest this perpendicular placement for avoiding parallax errors. Diffrient et al. (1981) conclude that the limits for secondary displays are 30° around the sightline and that the limits for detecting objects lies within 60° to the sides, 50° up, and 70° down (Figure 2). Alliance of Automobile Manufacturers [AAM] (2002) state no absolute limits, but suggest that
displays that carry information relevant to the driving task should be positioned as close as possible to the driver’s line of sight. Flashing signals is according to Diffrient et al. (1981) detected within 15-50° up from the sightline and 45-70° down from the sightline.

The state and the condition of the eye itself also confine the distances for where information is legible. For a normal reading distance Diffrient et al. (1981) suggest a value between 40.6 and 50.8 cm, but for reading of displays 33 and 71.1 cm with an absolute maximum of 76.2 cm. Van Cott and Kincade (1972) agree with the maximum value of 71 cm with the argument that the distance is at a reachable arm length. Leibowitz and Owens (1975) found that the resting state of the eyes, and hence, a good reading distance, has a mean value about 59 cm which lies somewhere in between previously stated limits. However, Chengalur et al. (2004) state that humans automatically adjust to shorten the viewing distance to improve focusing when it is difficult to read. According to Boff, Kaufmann, and Thomas (1986), the size of the visual field where it is possible to attain a sharp image is 1° around the fixation. According to Woodson, Tillman, and Tillman (1992) (as cited in Chengalur et al., 2004) objects 30° outside of this area is required to be ten times larger to be seen clearly.

2.2 Types of displays

In visual displays, Chengalur et al. (2004) state that for qualitative and check readings a moving pointer is preferred, while for quantitative readings without increase or decrease, a digital readout is preferred. Furthermore, unless only precise readings are needed, digitals should be redundant with analogues. For operating instructions in a panel with many other functions, an annunciator light should be used, and for a status reading a conventional light should be used. Woodson (1981) agrees with the use of an annunciator light. According to Ivergaard (1999), a moving pointer and a fixed scale should be used unless the range is very large. According to Diffrient et al. (1981), values in a display should increase clockwise, left to right, or from bottom to top. According to Heglin (1973), circular or semi-circular scales are preferred, but if numerical increase is related to some other natural representation such as more or less or up and down, it is easier to interpret a straight line or thermometer scale with a moving pointer. Information is not always suitable for visual presentation, instead sometimes an auditory presentation can be more suitable. Deatherage (1972) suggests some
circumstances where visual presentation is to be preferred such as when the message is complex, long, will be referred to later, does not call for immediate action, the user’s auditory system is overburdened, or when the user’s task allows to remain in one position. Auditory presentation, on the other hand, should be used where the opposite conditions are valid and when the receiving location is too dark or too bright. Chengalur et al. (2004) state that simultaneous presentation of aural and visual signals increase the probability of detecting the signal. Heglin (1973) agrees with the later, if the warning is urgent. According to Green, Levison, Paelke, and Serafin (1993), voice messages should not be used, except for the most urgent and unusual warnings. However, they also state that if the information is limited to a few chunks of noun-verb pairs, or prepositional phrases, it might be of value to provide both auditory and visual navigation information.

2.3 Dials

Circular dials are a common type of displays traditionally used for presenting information in the automobile. According to Diffrient et al. (1981) a dial diameter should optimally be between 7-7.6 cm with a pointer length of 2.9 cm. For a precision readout, the dial should be enlarged to 10-15 cm, while for a check reading it is sufficient with 4.4 cm. There are also guidelines specifically for pointers appearance. Diffrient et al. (1981) and Grether and Baker (1972) suggest that the pointer should reach the major scale markers but should not overlap smaller scale markers. According to Grether and Baker (1972), the pointer should also lie close to the dial surface to avoid parallax errors, be pointed with a tip angle of about 20°, have a uniform colour from tip to pivot, and have the remaining part of the pointer as short as possible (Figure 3).

While a horizontal scale is used, the pointer should be on top. If a vertical scale is used, the pointer should be on the right hand side (Figure 3). According to Chengalur et al. (2004) and Sanders and McCormick (1993), the placement of zero in a continuous scale should be at twelve or nine o’clock. If the scale does not fill the perimeter, zero should be placed at the lower part of the dial, six or twelve o’clock (Figure 4). Diffrient et al. (1981) state that vertical characters and few indices should be used for speedy reading, and the location of numbers should be beyond the indices to eliminate pointer overlap.
For guidelines regarding dial scales, Woodson (1981) reports a formula for calculating index sizes: dimension at D inches equals dimension at 71 cm times D inches divided by 28, where the dimensions at 71 cm can be seen in Figure 5. Chengalur et al. (2004) suggest that the maximum number of markings between numbers should be nine. Chengalur et al. (2004) and Ivergaard (1999) suggest that the progression of numbers should be in steps of one, two, or five.

Figure 4 Recommended placement of zero in dials

2.4 Characters

The characters used in a display are important for providing fast and safe readouts and many guidelines have been formulated regarding size, case, and fonts. Wickens and Hollands (1999) recommend the use of four character sizes with larger sizes for attracting attention. Diffrient et al. (1981) recommend at a reading distance of 71 cm character sizes of 1.3 mm for a routine marking, 2.5 mm for a critical marking in good lighting conditions, and 3.8 mm for a critical marking in low illumination. At 91 cm the corresponding sizes are 1.5 mm, 3.3 mm, and 4.8 mm. According to Grandjean (1987) and Sanders and McCormick (1993), the preferred character height of a capital letter should subtend about 16-24(25) minutes of arc in the visual field. At a distance of 70 cm this would suggest 4.3 mm. For other dimensions of letters, Grandjean (1987) reports that the width of a capital letter should be 75% of its height, the stroke width 20%, and the distance between characters 25%. Smith (1979) has stated the “James Bond Rule” which state that the visual angle of a character (its height divided by the viewing distance) should be at least .007 radians for 100% legible reading. Hence, at 70 cm viewing distance the characters should be 4.9 mm. For .0046 radians, characters should be legible for 98% of people, and for .0028 radians 90% of people. For electronic displays specifically, Chengalur et al. (2004) suggest that the height to width ratio of digits should be 0.6-0.8, the distance between digits 1.1-1.4 times the stroke width, and
the dot spacing in a dot matrix display should be 0.4-0.6 mm. The case of the character is another issue regarding the understandability of text. According to Diffrient et al. (1981), capital letters could be read further away because of increased size, but lowercase letters can be read 13.4% faster because of the up and down cues. Poulton (1967) states that a combination of upper- and lower-case letters generally provide more information than upper case alone. Green et al. (1993) add that mixed case should be used for messages longer than two or three words while Brown, Brown, Burkleo, Mangelsdorf, Olsen, and Perkins (1983, as cited in Smith and Mosier, 1986) argue that for alphabetic codes, all letters should be displayed in either upper- or lowercase letters, and if an arbitrary code is used it should not be longer than four to five characters. There are some guidelines regarding fonts in the area of automotive instrumentation. Woodson (1981) states that simple block styles are preferred for instrument labelling. According to Wickens and Hollands (1999), up to three fonts could be used at the same time, while Chengalar et al. (2004) state that vertical numbers should be used rather than slanted and that the use of segmented alphanumerics should be minimized (Figure 6).

![Figure 6 Avoid slanted numbers](image)

### 2.5 Colours and illumination

All guidelines in this study up to this point have not considered the effects of colour and illumination, which of course add other dimensions to the information design. Sanders and McCormick (1993) point out that the extremes of the colour spectrum (reds and blues) should be avoided, however, they also mention that there is contradictory evidence to this. Both Sanders and McCormick (1993) and Shneiderman (1997) suggest that on a dark background, colour pairs of saturated red and blue, and to a lesser extent red and green, or blue and green should be avoided. Yellow on purple and magenta on green are also hard to read. According to Reynolds, White, and Hilgendorf (1972), different colours ability to attract attention is from highest to descending: red, green, white, and yellow. They also found that stimulus colour, background colour, and ambient illumination must be considered in order to choose the most effective signal. There are a limited number of colours suitable to use when maximizing human performance. According to Jones (1962), nine colours can always be discriminated by a normal user. Smith and Thomas (1964) state that absolute discrimination for colours falls somewhere in the range of 5 to 12 alternatives depending on viewing conditions. According to Wickens and Hollands (1999), four standard colours should be used with additional colours reserved for occasional use. Also, two levels of intensity with limited use of high intensity are suitable. According to Shneiderman (1997), the number of colours for codes in an alphanumeric display should be four in a single display, and seven in an entire sequence of displays. Sanders and McCormick (1993) are a little bit more cautious with colours and suggest the use of as few colours as possible to avoid distraction. Not only the number of colours and how they are perceived limits the use of colours. There are also physiological limits regarding where colours can be seen.
According to Diffrient et al. (1974), the limit of colour discrimination upwards is 35°, and downwards 40° while Ivergaard (1999) states that colours should not be placed more than 15 minutes of arc away from the line of sight, due to that they are not colour coded to the same extent if they are placed further away, and that a coloured light should subtend at least 20 minutes of arc. If the light is situated in a larger display, at least 1° of visual angle should be subtended for increased colour discriminability. Chengalur et al. (2004) state that dim light is best spotted outside the fovea. Different colours can be used to represent different conditions and situations, though, it is important that the coding of each colour is meaningful to the user. According to Poulton (1975), colours could be used for size coding, where the order from largest to smallest are represented by red, orange, yellow, green, blue, violet, and white. According to Brown et al. (1983, as cited in Smith and Mosier, 1986) gradual colour changes could be used to show the relative values of a single variable. Colour coding should always be redundant with some other feature such as symbology. The authors also conclude that each colour represents only one category of displayed data.

Diffrient et al. (1981), Green et al. (1993), and Woodson (1981) suggest that the colour red should be used for critical failures and stops, yellow in cautionary and less critical incidents, and white for general status or non-critical events. Diffrient et al. (1981) and Woodson (1981) suggest that green should be used for normal and safe conditions. Diffrient et al. (1981) also suggest amber for check purposes and that blue should be avoided. According to Brown et al. (1983, as cited in Smith and Mosier, 1986) blue should only be used for backgrounds. There are also a few guidelines regarding the relationship between fore- and background colours. Not only colour matters when increasing readability, contrast also has a large impact. According to Diffrient et al. (1981) text should be displayed in black on a light background under ample light conditions. For dark adaptation, text should be light on a dark background (Figure 7). Sanders and McCormick (1993) suggest that the colour contrast between text and background should be maximized. Changes of illumination in surroundings introduce changes in appearance of colours (cromacities), which may cause confusion in identifying objects on a colour basis. There are a number of guidelines that deals with this issue in order to get as much as possible out of colours. Weber’s law of just noticeable difference (\(\Delta I = kI\)) (Salvendy, 2005), describes \(\Delta I\) as the difference between stimulus and background intensity, \(k\) as a proportionality constant, and \(I\) as the intensity of the background field. Smith and Thomas (1964) suggest that luminance is not a reliable cue in colour perception because dark hues are difficult to recognize. Sanders and McCormick (1993) state that higher levels of illumination may reduce demands on users’ information processing systems and that illumination recommendations are continually increasing.
A way of increasing the eye-catching abilities of colours is to utilize blinking symbols or messages. Both Wickens and Hollands (1999) and Sanders and McCormick (1993) state that the change of colour should be used with great care and in limited areas, even in warnings. However, Wickens and Hollands (1999) state that if a display must blink, it should blink with a rate of two to four Hz. Smith and Mosier (1986) report a similar rate, two to five Hz, with a minimum on period of 50 %. According to both Woodson and Conover (1964) and Heglin (1973), a suitable flash rate for attracting attention is three to ten times per second with a duration of least 0.05 s four times per second and with equal light and dark intervals. According to Markowitz (1971) a flash rate of 60 to 120 times per minute is suitable for highway driving. According to Fisher and Tan (1989), flashing text is inefficient when it comes to reading a message. According to Spoto and Babu (1989) inverted colours are more effective than increased brightness.

2.6 Data density

The density of the data presented is important because the size of the space for presenting is rather restricted inside an automobile. The information must, however, be presented in an orderly manner to the driver for increased understandability and usability.

There exist some different views on how large percentage of a display screen that should be utilized. Helander, Landauer, and Prabhu (1997) suggest that the optimum amount of information to be presented is what is necessary for the task at hand – no more and no less. Danchak (1976) suggests that 15 % is a good percentage of active screen area usage and it should not exceed 25 %. NASA (1980, as cited in Helander et al., 1997), has another view and suggests that the percentage of active screen area usage should not exceed 60 %. According to Shneiderman (1997), local, and overall, densities make it easier to read displays. Local density is here considered to be a 5° visual angle around each character and at normal viewing distances this means approximately 15 characters wide and seven characters high. Colour could also be used in graphical displays to increase the information density. Another way of increasing the usability of a display and shortening the search time is to group information in meaningful chunks.

The search time for objects is also a critical aspect to consider when designing driver information. The Society of Automotive Engineers (SAE) (as cited in Young, Regan, and Hammer, 2003) state that if a secondary task could be completed within 15s when the vehicle is stopped, it is suitable in a driving situation as well. Tijerina et al. (2000, as cited in Young et al. 2003) have, however, found contradictory results to this in the literature. AAM (2002) claims that systems with visual displays should be designed so the driver can complete the desired task with sequential glances that are brief enough not to adversely affect driving. Sanders and McCormick (1993), Helander et al. (1997), Tullis (1983), and Tullis (1986, as cited in Shneiderman, 1997) all agree that the number of groups on the screen should be minimized by making each group as close to 5° as feasible in a structured display. The number of groups also has an impact on search time, even when overall signal rate is constant due to load stress. Groups smaller than 5° can be fixated once with the necessary information still extracted. For larger than 5°, the size of the group determines the number of fixations. Sanders and McCormick (1993) conclude that if a small amount of information is spread all over the screen,
search time can be long. By packing the information more densely, search time can be improved, but eventually, the information becomes too crowded and the performance deteriorates. Search times are generally faster for items arranged in columns than as a horizontal list of running text. It is nearly impossible to focus attention on one visual source and ignore another within 1° of visual angle from each other. Shneiderman (1997) suggests that similar colours should be used to group objects. Green et al. (1993) suggest that a common space may be used to display messages from any source.

A third way of improving the use of a display’s area is to make use of layering and separation. According to Tufte (1990, as cited in Salvendy, 2005), it is important to visually stratify different categories of information and weigh importance of different information. It is also important to eliminate irrelevant graphical structures such as containers, decorations, and gridlines. Galitz (2007) states that everything must have meaning to users as well as serve a purpose in performing a task. Easterby (1967, 1970) claims that contrast boundaries, such as solid shapes are preferable to line boundaries. Also closed figures enhance the perceptual process. According to Thacker (1986, as cited in Helander et al., 1997) graphical boundaries are preferred by users, but do not seem to reduce search time. According to Green et al. (1993) all lines and gaps between lines should be at least 0.6 mm wide. To be readily discriminated, the minimum line width, and gaps between lines or other graphic elements and all critical details, should be 1/10 of the minimum height of characters. According to Wickens and Hollands (1999) discrimination should be constituted by underlining, enclosing in a box, pointing to with an arrow, or by an indicator such as an asterisk, bullet, dash, plus, or a x.

2.7 Graphics

There are certain circumstances where information is best presented in the form of text and other where it should be presented as graphics. Galitz (2007) takes a stand for graphics and suggests that a test for good design is to check if all screen elements can be identified by cues other than by reading the words that make them up. Stewart (1980) and Foley and van Dam (1982) suggest that graphics should be considered rather than a text explanation, or a tabulation, for displaying time-varying phenomena or variations in trends. Smith and Mosier (1986) and Cleveland (1985) suggest the use of graphics when users quickly must scan and compare sets of data. According to Baber and Wankling (1992) symbols need to be redundant with additional text information in order to increase performance. Woodson (1981) states that words should be used for increased understandability and where everything not can be pictorialized. Pictorial symbols should have sufficient details so they are recognizable, a border so they do not blend with the surroundings, and they should not be placed on reorientable controls.
2.8 Warnings

An important form of information is warnings, which could be presented as text, symbols, or as a combination of both. Chengalur et al. (2004) recommend for warning presentation that vague, ambiguous, or ill-defined terms, highly technical terms or phrases, double negatives, complex grammar, and sentences with more than 12 words should be avoided. A warning should be larger and more separated from the information in its surrounding for easier detection. An icon representing the consequences that might occur attracts the user’s attention easier. According to Green et al. (1993), warnings should rely upon text messages supplemented by standard international symbols. Chengalur et al. (2004) suggest that in the western world, warnings should be placed towards the top-left for easier detection due to the reading direction and also placed as near as the hazard as possible. Heglin (1973) states that there should ordinarily be one warning light, if more are required, a master warning with a word panel indicating specific conditions should be used. The warning light intensity should be at least twice as bright as its background. A warning light should be placed within 30° of the user’s normal line of sight, and at least be 1½ the size of other indicators. Flashing lights should only be used for extreme emergencies since they are distracting. Woodson (1981) also agrees with the use of a master warning light.

2.9 Attention

Sanders and McCormick (1993, pp 71-76) have compiled various lists of guidelines with respect to attention. For selective attention tasks they recommended the following:

- Where multiple channels must be scanned for signals, use as few channels as possible, even if it means increasing the signal rate per channel.
- Provide information to the person as to the relative importance of the various channels so that attention can be directed more effectively.
- Provide the person with some preview information as to where the signals will likely occur in the future.
- If multiple visual channels are to be scanned, put them close together to reduce scanning requirements.
- Where possible, stimuli that require individual responses should be separated temporarily and presented at such a rate that they can be responded to individually. Extremely short intervals (say, less than 0.5 or 0.25 s) should be avoided. Where possible, the user should be permitted to control the rate of stimulus input.

For focused-attention tasks:

- Make the competing channels as distinct as possible from the channel to which the person is to attend.
- Separate, in physical space, the competing channels from the channel of interest.
- Reduce the number of competing channels.
- Make the channel of interest larger, brighter, louder, or more centrally located than the competing channels.
For divided-attention tasks:

- Where possible, the number of potential sources of information should be minimized.
- Where time-sharing is likely to stress a person’s capacity, the person should be provided with information about the relative priorities of the tasks so that an optimum strategy of dividing attention can be formulated.
- The tasks should be made as dissimilar as possible in terms of demands on processing stages, input and output modalities, and memory codes.

For sustained attention tasks:

- Increase the conspicuity of the signal. Make it larger, more intense, of longer duration, and more distinctive.
- Reduce uncertainty as to when or where the signal will likely occur.
- Improve motivation by emphasizing the importance of the task being performed.
- Reduce the rate at which stimuli (which may or may not be signals) are presented if the rate is high.

Chengalur et al. (2004) offer some other features that increase the probability that a signal is detected:

- Too many or too few signals to be detected and responded to should be avoided.
- The signal should be presented redundantly.
- The signal should be amplified more and differentially from noise.
- The signal should be made dynamic.
- Knowledge of the results to expect should be provided.
- A refresher of the standard of discrimination to be found should be provided.
- Bringing in a secondary display task should be avoided to increase the probability to detect a signal.

3 Discussion

In this literature review, the body of literature on design guidelines and principles for information presentation, focusing on in-vehicle information presentation has been studied. In general, guidelines found are applicable for today’s automobiles. There are, however, some cases where the guidelines are vague and merely state “the way it should be” without giving any guidance on how to get there.

3.1 Eye and head movements

Most of the guidelines in this category seem reasonable except the viewing distances. Even the maximum distance stated, 76.2 cm, is rather short since much information in automobiles is placed further away today. This will, however, probably not affect the readability since the larger distances is compensated by larger information sources. The need for a guideline which states the ratio between distance and information size would, therefore, be helpful. There are also some disagreements on how large an effortless head movement is. However, this could easily be examined by physiological measurements. It is probably reasonable to assume that the whole interval 30°–60° could be considered due to different drivers’ attitudes towards what an effortless head movement is.
3.2 Types of displays, dials, and characters

The guidelines stating that values should increase clockwise, left to right, or bottom to top, complies best with the western world due to the reading direction. The sizes of the dials in the presented guidelines seem rather small since most of today’s automobiles have larger dials then the suggested optimal value of 7-7.6 cm, though, at a larger distance. The need for a distance/size ratio is present here as well. The distances for guidelines regarding character size should be extended to include distances larger than 70-71 cm because much information is placed further away today. The guidelines for character size ranges from 1.3-4.3 mm. The reading performance of letters is, however, influenced by many factors other than reading distance, but in most cases this could probably be considered a good range to keep letter sizes within.

3.3 Colours and illumination

It is hard to draw any conclusions of exactly how many colours that are suitable in different situations, but a range of four to twelve colours seems reasonable. There is a gap of knowledge regarding graphical displays and how the use of numerous and gradient colours will effect driver performance. The tests for a number of discriminable flash rates are due to technical restrictions limited to rates under 9Hz (Mortimer and Kupec, 1983). Nowadays, LEDs have much faster response times and can generate proper flash rates over 9Hz. More research is needed regarding whether the extremes of the colour spectrum should be used or not, since Sanders and McCormick (1993) point out that there are contradictory evidence in this area. There is also a gap in the knowledge about animations of graphics inside vehicles which today’s technology in automobiles is capable of displaying.

3.4 Data density and graphics

There seems to be some contradictory results regarding a good percentage of active screen area usage. The guidelines range from 15-60 %. This range is too large to be of any use as a guideline in designing displays. However, Tullis (1983) discusses that the value of 60 % is only valid in certain situations and, hence, that a good active screen area usage should not exceed 15-25 %. The guideline that a task completion time of 15s for a stationary automobile also makes the task suitable when driving (the 15s rule) seems understated or ill-defined. However, 15s with eyes completely away from the road seems to be appropriate only in very specific traffic and road situations. There should be a more precise definition of how these 15 seconds are to be distributed since many sources are questioning the appropriateness of the 15s rule (Young et al. 2003). Regarding graphical boundaries, there are some uncertainties whether they should be used or not. The boundaries are preferred by users and do not reduce search time, but do they distract? There seem to be some lack of knowledge whether information in automobiles is best presented in text or as graphics. A combination of both could be usable, but could, at the same time, distract the driver due to the amount of information displayed. There is a lack of knowledge regarding which form of presentation that requires the least time to be understood for specific messages or warnings.
3.5 Conclusions

It has not been possible to identify all relevant articles and, hence, not all relevant guidelines and principles. A deeper look into journal publications of original research would have been preferred. The aim of this literature review was, however, a first step to examine if older, widely accepted guidelines still are valid for today’s automobiles. Generally, the presented guidelines are relevant in today’s cars, as they primarily consist of information presentation principles that can be used regardless of technology used in the automobile. Most guidelines also point to the same direction, there are only some results that differ from the main opinion, mainly in the areas of data density and search time. Some areas need, however, some further research:

- There is a lack of directives regarding suitable ratios between distance and size of information presented in the instrument panel. Since automobiles of today make use of areas further away from the driver, and the technology has developed since many of these guidelines were stated, there is a need to expand the limits of the recommended distances and state an appropriate ratio between distance and size.
- There is a gap of knowledge regarding graphical displays and how the use of numerous and gradient colours effect driver performance. Also animations in automobiles could be an area for further studies.
- The use of graphical boundaries is another area that needs more attention, because it is not clear whether they increase, or decrease, user performance.
- More knowledge is also needed concerning whether information should be presented as text or graphics in different situations.

The next step in this should be to test the validity of these guidelines in the design of automobile instrumentation concepts to be used in driving simulator studies for increased usability and safety in traffic.
4 References


Baber, C and Wankling, J (1992) An Experimental Comparison of Text and Symbols for In-Car Reconfigurable Displays Applied Ergonomics Vol 1 No 4 pp 255-262


Danchak, M M (1976) CRT displays for power plants Instrumentation Technology Vol 10 No 23 pp 29-36


Easterby, R (1967) Perceptual organization in static displays for man/machine systems Ergonomics Vol 10 No 2 pp 195-205

Easterby, R (1970) The perception of symbols for machine displays Ergonomics Vol 13 No 1 pp 149-158


Heglin, H J (1973) NAVSHIPS display illumination design guide: II Human factors (NELC/TD223) Naval Electronics Laboratory Center, San Diego


Leibowitz, H W and D A Owens (1975) Anomalous myopia and the intermediate dark focus of accommodation Science Vol 189 No 4203 pp 646-648

Markowitz, J (1971) Optimal flash rate and duty cycle for flashing visual indicators Human Factors Vol 13 No 5 pp 427-433


Poulton, E C (1975) Colours and sizes-a recommended ergonomic colour code Applied Ergonomics Vol 6 No 4 pp 231-235

SAE (1994) Describing and measuring the driver’s field of view The Society of Automotive Engineers SAE J1050


Smith, S L and Thomas, D W (1964) Color versus shape coding in information displays Journal of Applied Psychology Vol 48 No 1 pp 137-146


Woodson, W E, and Conover, D W (1964) Human engineering guide for equipment designers, (2nd ed.) University of California Press, Berkeley, CA

DO REDUNDANT HEAD-UP AND HEAD-DOWN DISPLAY CONFIGURATIONS CAUSE DISTRACTIONS?

Carl Jörgen Normark, Phillip Tretten, Anita Gärling
Department of Human Work Sciences
Luleå University of Technology
Luleå, Sweden
E-mail: jorgen.normark@ltu.se

Summary: This study was designed to investigate effects of different display configuration designs. Nineteen drivers completed a driving simulator study designed to resemble normal driving. Driving performance, glance behaviour, physiological measures, and task completion times was measured for two display configuration designs both during driving only and during driving with a simple secondary task, which consisted of detection, and off-setting of presented warnings. The display configuration design with more centrally placed information, e. g. the HUD and HDD, had less detrimental effects on driving performance and glance behaviour. The physiological measures showed, however, no significant differences between display configuration designs.

INTRODUCTION

Automobiles are becoming more and more technologically complex (Baber & Wankling, 1992; Noy, 1997) with more and more built-in driver information systems (Tsimhoni & Green, 2001). This also increases the amount, and range, of information presented to the driver, which leads to the use of dynamic displays showing many types of information in one place (Baber & Wankling, 1992). The driver is mainly guided by vision for driving, and visual information can be lost in the quantity of information displayed to the driver and this can have a negative effect on driving safely since the visual faculties competes with the same perceptual and cognitive resources as the task of driving (Horberry, Anderson, Regan, Triggs, & Brown, 2006). For a safe traffic environment, drivers must have their attention on the outside of the car rather than on in-vehicle displays (Baber & Wankling, 1992). Luoma & Rämä (2002) even claim that an in-vehicle information system should only provide information when it’s really needed. Having attention on something other than driving, a secondary task, increases the in-vehicle glance durations and is associated with decreased driving performance and increased crash risk (Horrey & Wickens, 2007; Tsimhoni & Green, 2001). Horrey, Wickens, & Consalus (2005) and Wittman et al. (2006) declare that if information is acquired from a display located close to the road view, the driving performance is less degraded than for a distant display. Using a Head-Up Display (HUD) can also result in reduced workload, decreased response times, and increased driving comfort (Liu & Wen 2004). According to Wittmann et al., (2006), the detection of signals is easiest near the line of sight, and it decreases significantly for larger eccentricities, especially the vertical. According to Gish & Staplin (1995), a HUD can increase the time the eyes are kept on the road, but also deteriorate the reaction time for events on the road by increased visual clutter in the driver’s line of sight. According to Ellis (2005), redundant elements may aid in the interpretation of information, but other forms of redundancy may also deteriorate performance (Seagull, Wickens & Loeb, 2001). Physiological measures have been used in
earlier traffic safety studies and might be useful for studying workload in in-vehicle technology (Johansson et al., 2004).

Research aim and research questions. The aim of this study is to investigate the influence of two display configurations on driving performance and glance behaviour both while driving and while performing a simple secondary task while driving. One display configuration had a four display design similar to those found in high-end automobiles of today (e.g. BMW 7-series) while the other presented information centrally and redundant in two different displays; a LCD display in a HUD position 15° from the drivers’ normal line of sight, that did not obstruct the drivers view of the road, and a regular instrument cluster head-down display (HDD). The following specific questions are addressed: Does redundantly displayed information placed in the driver’s line of sight differ with respect to driving performance, added distraction, and time spent looking away from the road? Is the driver’s stress level affected by having information displayed in the line of sight? Are warnings detected and distinguished from normal in-vehicle information while driving?

METHOD

Participants and equipment

Nineteen drivers (10 males and 9 females aged 20 to 58 years with a mean age of 37.6 years) conducted the study. All participants had a valid driver’s licence and either normal or corrected to normal vision. None of the participants’ private vehicles were equipped with any type of HUD display. The experiment took place in Luleå University of Technology Designlab’s driving simulator, consisting of a fixed base Volvo XC90 cockpit where four LCD displays replaced the original instrumentation (Figure 1) and it’s handling was configured to simulate a front-wheel drive SUV (e.g., Volvo XC90). The road view was projected by a NEC NP-1000 projector on a 1.8m high by 2.4m wide screen in front of the driver which subtends about 33.4° of the driver’s forward view. Eye movements were monitored by Seeing Machine’s FaceLab system (version 4.5). The minimum duration for a glance was in this study set to 100ms (Horrey & Wickens, 2007). Eye data calculations were based on fixations towards areas of interest (AOIs). Physiological measures were collected with Mind Media’s Nexus-10 hardware and BioTrace (version 1.20) software. EKG sensors in a Lead II chest position were used for measuring heart rate (HR). Galvanic skin resistance (GSR) sensors were mounted on the index and ring finger of the left hand. A temperature sensor was mounted on the middle finger on the left hand.

Driving environment

The driving environment was designed to simulate a realistic route with traffic, surroundings, and events that might occur in a realistic driving situation. Carsten et al. (2005) suggest that a rural road generally gives the largest effect sizes for a driving simulator study. In this study, an approximately 15 km long road with two lanes through rural areas and with a short four lane segment through a city environment was used. There were segments with 50 and 70 km/h speed limits. Throughout the study there was some oncoming traffic in the opposite lane and some in the same lane as the driver to simulate realistic driving. To keep the driver focused on the driving, some cars had to be overtaken, some cars made unexpected manoeuvres by abruptly braking, and at one time a cyclist entered and crossed the road from behind a parked truck.
Experimental design

The experiment was a 2 (driving condition) x 2 (display configuration) factorial design with repeated measures on the first factor. The driving conditions were “driving only” and “driving with a task”. For the latter warnings were presented to the driver in the form of common automobile warnings consisting of a 15×15mm icon accompanied by a text such as for instance “Low washer fluid”. There were ten generic warnings of similar length and similar complicity displayed. The participants were instructed to offset the warnings as soon as they noticed them. The two display configurations were “Redundant HUD” where vehicle speed and warnings was presented to the driver redundant in the HUD and HDD, and “Spread out” where vehicle speed was presented in the HDD display and warnings appeared in one of the four display positions (Figure 1). The order of displays the warnings appeared in was: Infotainment display (IF), Centerstack (CS), HDD, IF, HUD, HDD, IF, HUD, HDD, and CS.

Figure 1. Display configurations. HUD – Head-up, HDD – Head-down, IF – Infotainment, CS – Centerstack

Procedure

The experimental session started with the participants being introduced to the simulator and given a five minute practice run to get familiar with handling the simulator. There were then two driving blocks of 15 minute each; a “driving only” block and a “driving and task” block. The order of these was balanced and both were made on the same road segment but in reverse directions in order to prevent the driver from getting too familiar with the road segment. The “driving only” block consisted only of driving through the road segment. The “driving and task” block consisted of driving with the addition of a simple secondary visual detection task, where the drivers were asked to, while driving, reset a total of ten warnings as soon as they had been discovered. The participants were asked to drive as they normally do with their own vehicles and to obey presented speed limits. Custom software was used to synchronize and reduce all data regarding the dependent measures (Table 2) to 10Hz, to analyze gaze data, driving data, and physiological data. Mann-Whitney’s U-test with significance levels set to .05 was used for between subject analyzes, and Wilcoxon Signed Ranks Test for between subjects analyzes.
Table 1. Dependent measures and their definitions

<table>
<thead>
<tr>
<th>Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean speed</td>
<td>Vehicle’s mean speed</td>
</tr>
<tr>
<td>Standard deviation of speed</td>
<td>How much the vehicle’s speed deviates</td>
</tr>
<tr>
<td>Maximum speed</td>
<td>Maximum speed reached</td>
</tr>
<tr>
<td>Standard deviation from speed limit</td>
<td>How much the vehicle’s speed deviates from the posted speed limit</td>
</tr>
<tr>
<td>Standard deviation of lane position</td>
<td>How much the driver’s lateral control of the vehicle deviates</td>
</tr>
<tr>
<td>Number of lane exceedences</td>
<td>The number of times more than half of the vehicle exceeds the lane</td>
</tr>
<tr>
<td>Mean GSR</td>
<td>Difference in Galvanic Skin Resistance while relaxed and while driving</td>
</tr>
<tr>
<td>Mean temp</td>
<td>Difference in mean hand temperature while relaxed and while driving</td>
</tr>
<tr>
<td>Mean HR</td>
<td>Difference in drivers mean Heart Rate while relaxed and while driving</td>
</tr>
<tr>
<td>Time to notice</td>
<td>Time from a warning appears until gaze is directed towards display</td>
</tr>
<tr>
<td>Glance frequency</td>
<td>Number of glances to the warning while it is displayed</td>
</tr>
<tr>
<td>Total glance duration</td>
<td>Total time the warning display is gazed upon</td>
</tr>
<tr>
<td>Mean glance duration</td>
<td>Mean time the warning display is gazed upon</td>
</tr>
<tr>
<td>Gaze duration off road scene ahead</td>
<td>Time the gaze is not directed towards the road scene ahead</td>
</tr>
<tr>
<td>HUD duration</td>
<td>Total time spent viewing the Head-up display</td>
</tr>
<tr>
<td>HDD duration</td>
<td>Total time spent viewing the Head-down display</td>
</tr>
<tr>
<td>Task completion time</td>
<td>The time from a warning appears until it is reset</td>
</tr>
</tbody>
</table>

RESULTS

Differences between driving tasks for “Redundant HUD”. Adding a simple secondary visual detection task to the driving task significantly increased the mean speed, \((z=-2.701, p=0.007)\), while the standard deviation of speed was significantly lower during “driving with task”, \((z=-2.191, p=0.028)\). No significant differences were found regarding the physiological measures.

Differences between driving tasks for “Spread out”. The “driving only” condition showed lower means for HDD duration, \((z=-2.31, p=0.021)\), compared to “driving with task”. No significant differences were found regarding the driving performance or physiological measures.

Differences between display configurations for “driving only”. When comparing “driving only” data for the two display configurations, the “Redundant HUD” showed a significantly lower mean for standard deviation of lane position, \((U=102.000, p=0.023)\). “Redundant HUD” also showed a lower mean for HDD duration, \((U=54.500, p=0.000)\), and a higher mean for HUD duration, \((U=94.000, p=0.012)\), caused by the HUD-speedometer. There were no significant differences between the display configurations and gaze duration off road scene ahead or any of the physiological measures.

Differences between display configurations for “driving and task”. Data was analyzed for every separate warning occurrence from the time the warning appeared until 15s after it was reset. This time interval was chosen because the driving performance and the physiological measures were hypothesized to also be affected a short while after the completion of the secondary task. When analyzing all warning occurrences together, the “Redundant HUD” showed significantly lower means for: standard deviation of lane position, \((U=17.000, p=0.022)\), and number of lane exceedences, \((U=11.000, p=0.005)\). In total 45 lane exceedences occurred for the “Redundant HUD” and 68 for the “Spread out”. The glance measures showed lower means for “Redundant HUD” for: time to notice, \((U=21.000, p=0.050)\), glance frequency, \((U=7.500, p=0.002)\), total glance duration, \((U=19.000, p=0.034)\), Gaze duration off road scene ahead, \((U=11.000, p=0.006)\), and HDD duration, \((U=6.500, p=0.002)\).
“Redundant HUD” showed a higher mean for HUD duration, \((U=14,500, p=0.002)\). The secondary task was performed faster for the “Redundant HUD” than for the “Spread out”, \((U=16,000, p=0.018)\). Detection times were analysed for each separate warning in order to study if warnings are distinguished from normal in-vehicle information (Table 4). One warning occurrence did stand out for the “Spread out”; namely warning 10 which was displayed in the centerstack position during a high workload traffic environment (Table 4). Significant lower detection times was found for the “Redundant HUD” for warning 2, \((U=15,000, p=0.026)\), and warning 10, \((U=12,000, p=0.007)\). For “Redundant HUD”, the participants had the opportunity to choose which one of the HUD and HDD displays to consult to recognize a warning. Seven out of ten participants chose to exclusively consult the HUD, one checked the HUD for nine of the ten warning occurrences, and two preferred the HDD prior to the HUD for eight of the warnings.

### Table 2. Mean detection times for each warning (in seconds)

<table>
<thead>
<tr>
<th>Warning number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to notice “Redundant HUD”</td>
<td>1.17</td>
<td>2.24</td>
<td>2.22</td>
<td>1.00</td>
<td>1.32</td>
<td>3.04</td>
<td>1.56</td>
<td>4.89</td>
<td>5.43</td>
<td>0.95</td>
</tr>
<tr>
<td>Time to notice “Spread out”</td>
<td>4.28</td>
<td>3.68</td>
<td>1.98</td>
<td>0.90</td>
<td>4.20</td>
<td>3.13</td>
<td>2.54</td>
<td>3.27</td>
<td>3.34</td>
<td>18.16</td>
</tr>
</tbody>
</table>

**DISCUSSION**

The results in this study did show some differences between the two display configurations, which are in line with earlier research, namely that the display configuration with more centralized information resulted in better driving- and task performance than the spread out display configuration (Horrey, Wickens, & Consalus, 2005; Wittman et al., 2006). According to Tsimhoni and Green (2001), the addition of a secondary task while driving significantly impairs driving performance, which could not be confirmed in this study. Probably due to the simplicity of the secondary task used which was not sufficient to cause extra workload to yield any differences in lane keeping or stress. Some differences were, however, found regarding speed maintenance in the “Redundant HUD”, where the speed varied even less when the task was performed. This could be explained by the HUD-speedometer; at the same time the warning was gazed upon, the vehicle’s speed was also monitored. More time was spent looking at the HDD when a task was added for “Spread out”. However, this result needs further investigation to be explained. Both driving tasks rendered similar results in favour of the “Redundant HUD”. Even though much time is spent looking at the HUD for “Redundant HUD”, the driving performance is not negatively effected, but instead improved. This is probably an effect of the increased time the road is viewed compared to the “Spread out”. Looking down at the HDD and the other display positions is more time consuming than looking at the HUD. The results also imply that despite the highly salient position of “Redundant HUD” close to the driver’s field of view, this does not attract unnecessary attention as the glance duration towards the “Redundant HUD” was lower. Since it is desirable to keep the eyes on the road, the fewer glances towards the vehicle interior should be the better. “Spread out” had more glances away from the road, which causes the driver to get “out of the driving loop” more often (Kircher, 2007) and, thus, could have a negative effect on safety. The lesser lane exceedences for “Redundant HUD” than for “Spread out” during “driving and task” also imply that looking too much at in-vehicle displays has a detrimental effect on driving safety. There was no problem for the drivers to notice the warnings when they were displayed, except from for the CS position. Hence, this position should clearly be avoided when presenting emergency information to the driver, as can be seen in Table 4. For warning occurrences 2 and 10, where the warning was displayed in the
CS position for “Spread out”, the detection time was significantly higher. It is noticeable that the HUD was the preferred display to observe in “Redundant HUD”. None of the participants had a HUD in their personal vehicle so the HUD was perhaps not naturally the primary choice in the search for information. Although, the highly salient position of the very novel HUD technology would ensure that presented information easily catch the driver’s eye, but not so much that it would cause distraction.

The physiological measures did, however, not imply any differences between the two display configurations or between the driving tasks. The stress caused by the display configuration and/or the task was probably too small to be detected compared to the overall stress caused by driving the simulator or the physiological measures used might not be suitable for evaluating in-vehicle systems. Nevertheless, stress could be of interest to study if the stress level caused by the driving situation can be filtered out from the stress caused by in-vehicle systems. If the trials had lasted longer, the participants’ stress and vigilance levels might have been lower, which could have made it easier to detect any differences. Although the ten warnings presented to the driver were generic, and the task was to off-set the warnings as soon as they were noticed, there was a possibility that the warnings were perceived as different from each other. However, this should not have any major effect on the results since the participants were told to take action as soon as a warning was presented. The participants seemed to remain highly vigilant during the driving blocks and in some cases repeatedly scanned the vehicle interior for new warnings, which, on the other hand, could have effected the time it took to notice a warning, but, if so, it should have affected both display configurations equally.

This study showed that even though the task of detecting and off-setting a warning is quite simple, it can still be used to reveal decreased driving- or task performance caused by in-vehicle technologies or vehicle interior design. This study is to be extended with a more demanding task in order to find out if the nature of the task causes different impact on driver performance and driving safety.

ACKNOWLEDGMENTS

This study was sponsored by the OPTIVe (OPTimized system integration for safe Interaction in Vehicles) project within the Swedish IVSS research foundation.

REFERENCES


EVALUATION OF CAR INSTRUMENTATION CLUSTERS BY USING EYE-TRACKING

Carl Jörgen Normark, Monica Kappfjell, Phillip Tretten, Jan Lundberg, Anita Gärling

Luleå University of Technology
SE-971 87 Luleå, Sweden
jorgen.normark@ltu.se

Keywords: Evaluation, instrumentation cluster, eye-tracking

ABSTRACT
The importance of the ambience presented by instrumentation in the vehicle of today is of great significance for the automotive industry. The rapid technological development of electronic equipment has enabled the large amount of information devices in the drivers’ environment. The types of driver information and their appearances are well identified by car manufacturers, but there is a lack of knowledge regarding drivers’ perceptions of instrumentation clusters as well as their appearance. To increase this knowledge a study was conducted in order to measure these perceptions. Twenty-three subjects were studied using a method consisting of eye-tracking, assessments and interviews. Each subject was to view eight clusters and to assess each cluster according to six different semantic descriptions. The interviews were based on the subject’s individual eye-tracking data. The subject was monitored by an eye-tracking equipment during the assessment task. The results showed that there were some significant correlations between time spent viewing a cluster and its assessment. Moreover, the results also give a view on how clusters are assessed and why they are assessed in particular ways.

INTRODUCTION
The rapid technological development of electronic equipment has enabled a large amount of information devices in the drivers’ environment. This information is designed by car manufacturers to be usable and safe for the driver. However there is a lack of knowledge when it comes to how the visual appearances of the information devices are perceived and how they should be measured (Kukkonen, 2005). Eye-tracking is a tool that could be useful when determining the visual perception of appearance.

Eye-tracking is a tool that collects data on eye position, gaze path, time spent looking at a stimulus or fixations at objects along with numerous other variables (Duchowski, 2003). The amount of data collected with the aid of eye-tracking would not be economically feasible or tolerably time consuming with any other method (Renshaw, 2004). Eye-tracking is used today in the automotive industry for measuring response times for subjects in given tasks while driving (Duchowski, 2003).

Human beings observe stimuli in different ways depending on why the stimuli are observed and what there is in the person’s mind while observing the object (Yarbus, 1967). The triggers to move the eyes are different according to the specificity of the task (Rayner, 1995). The pattern of the
gaze-path and fixations differs with regard to whether a person is supposed to just observe a stimulus with her/his own preferences in mind or whether the person is requested to find objects in the stimulus, or rate the stimulus according to a specific task (Koivunen, Kukkonen, Lahtinen, Rantala and Sharmin. 2004; Kukkonen, 2005). The viewing behaviour changes as the task changes, but a way of addressing this issue is to look for similar scene context effects on fixation time across tasks (Duchowski, 2003). To see how large differences there are among different tasks and in order to get meaningful eye-tracking data on the perceptions of the design in the present study, several statements to assess were to be presented to the subjects. To find suitable statements to present, the semantic evaluation methods SMB (Küller, 1975) (semantic environmental description) and PSA (Wikström, 2002) (product semantic analysis) were looked upon.

SMB is a tool for making evaluations of architecture, i.e. how a person experiences an environment. SMB has proved to function when evaluating other types of designs, such as for instance evaluation of car interiors (Karlsson, Aronsson, Svensson, 2003; Laike, 1999). This has been of importance when choosing SMB as inspiration to formulate the questions used in the present study, which is mainly aimed at car instrumentation cluster design.

PSA is a method for valuing and giving a measure of a product’s semantic functions. In PSA, descriptive words for characteristics are formulated with several different types of products in mind. PSA includes a study for evaluation of car design.

This study is a pilot study and an attempt to study if it is possible, or desirable, to carry out more research in the area of design evaluation with eye-tracking. The aim of this particular study is to use eye-tracking for evaluating visual perceptions of design in car instrument clusters, e.g. to study if the use of eye-tracking can make evaluation of design concepts efficient and accurate and generate quantifiable data out of an area of subjective evaluation.

The main aims of this study are:

- To study if perceptions of design can be evaluated with the help of eye-tracking.
- To study if subjects view instrument clusters that they find more attractive longer than less attractive clusters.
- To compare and analyze possible relations between the time spent looking at a cluster, and the assessments on the cluster.
- To identify objects in the clusters that draw a considerable amount of attention.

METHOD

Material
The present study uses three approaches to investigating perceptions of design, see Figure 1. The assessment task is done to get an assessment of the clusters, eye-tracking to see where the subject is looking when making these assessments, and an interview to get a deeper understanding on why these assessments were made.
The assessment task utilized the two semantic evaluation methods SMB and PSA. SMB uses descriptive words for characteristics. There are a total of 36 words of characteristics divided in eight dimensions. These dimensions are: "pleasantness", "complexity", "unity", "enclosedness", "potency", "social status", "affection" and "originality". Six of the dimensions of SMB were utilized when formulating statements in the present study. The category "enclosedness" was not utilized since it is difficult to apply to an instrument cluster because the subject is not located inside the cluster. In PSA, descriptive words of characteristics are formulated with several different types of products in mind. It is only these words that are utilized from PSA in the present study. PSA describes several different studies and one of the studies used to assemble the PSA method is a study to describe car interiors. In the present study the category of "potency" from SMB was not utilized since the focus groups formulating the words for the car design evaluation study within PSA, did not think it was important for car interiors. The other words used in the car interior evaluation in PSA were used in the present study since they are believed to be important in car interiors. Statements were formulated to represent the remaining six SMB dimensions with words believed to be important in car interiors. The statements were to be answered according to a seven-grade Likert scale. The first and the last statement the subject comes across are the same, e.g. a control factor.

Eye-tracking is used to determine where in a viewed cluster the subjects are looking. This is believed to give data on what in the cluster is being assessed when the assessment task is performed.

The interviews are believed to give a deeper understanding of why the subjects assessed the clusters in a particular way when looking at a particular place in the cluster.

Subjects
Twenty-three subjects participated in this study. The subjects were gender balanced; 11 females and 12 males, average age 25 years old. The subjects were students and available at Luleå University of Technology at the time of the study and they were willing to participate voluntarily. This group of subjects are however suitable for this study since they are likely to be a good target group in the future for the automotive manufacturers.

The subjects were split up into four groups (Figure 2). Half of the subjects were to view eight instrument clusters, clusters 1-8. The other half of the subjects were to view eight different clusters, clusters 9-16, to get a larger diversity of clusters (Figure 3). Two clusters appeared in both these groupings as reference clusters when looking for differences between the groupings. Half of the subjects that viewed clusters 1-8 were presented the clusters in a specific order, and the other half

Figure 1 – Eye-tracking needs to be combined with other methods to evaluate appearance.
were presented the clusters in a reversed order. The same reversal of clusters was done for the subjects that viewed clusters 9-16. The reversal of the clusters was made in the present study to examine if the order and placement would impact on how the clusters are viewed and assessed.

![Figure 2 – Grouping of subjects.](image)

**Procedure**

All participants went through the same procedure one by one. They had the same instructions and assessed the clusters under the same environmental conditions.

The test leader informed the subjects that they would view instrument clusters and the subject was asked to assess the visual design according to several statements. The subjects were to sit in front of a 17” computer screen monitored by a Tobii x50 eye-tracking equipment. The subjects were instructed to sit 60 cm away from the screen and the height was adjusted so the subjects’ eyes were in line with the centre of the screen (Tobii Technology, 2006). On the desk in front of the subject, a pile of papers with statements were placed. There was one statement on each sheet of paper and the clusters were to be assessed on a seven grade Likert scale, ranging from one to seven where one corresponds to “not at all” and seven to “very”.

![Figure 3 – All clusters used in the study. Groups 1&3 viewed clusters to the left, groups 2&4 viewed clusters to the right. Two clusters are reference clusters and used in both groupings.](image)
The test leader informed the subjects that they would first view four different instrument clusters to be assessed according to the statement printed on the first sheet of paper. The clusters would first appear on the screen one by one. For the following statements, all four clusters were presented at the same time. The reasons for displaying the clusters one by one was both to enlarge the clusters so that small details that point out would be easier to find, and to compare if the different presentations would affect the assessments. The reason for displaying all the four clusters at the same time is that it was time saving. When displayed one by one, the first cluster was called “A”, the second “B”, the third “C” and the fourth “D”. The test leader informed the subjects that the letter corresponding to the assessed cluster would be written next to the scale on the sheet of paper. The subjects were then asked to tell the test leader when the assessments of one cluster had been made, so the test leader could change to the next cluster.

The subjects were informed that when all four clusters had been assessed one by one according to the first statement, then all clusters would all appear at the same time on the screen for the rest of the following statements. Here cluster “A” was in the top-left corner, cluster “B” in the top-right corner, cluster “C” in the lower-left corner and cluster “D” in the lower-right corner. The subjects were informed that all four clusters on the screen would be assessed according to six different statements. The subjects were then asked to tell the test leader when the assessments according to one statement had been made, so that the test leader could change to the next question. For the first statement the clusters were presented one by one on the screen, and after that all four clusters were presented on the same screen. This was to compare if there were any differences in the way in which the statement “Attractive/good looking” would be assessed in the two different presentations.

The assessment task started with the following statement that was given when the clusters appeared one by one:

“Attractive/good looking”

The following statements were given in the following order when the clusters appeared four at the same time:

“Easy to understand and read out information”
“Unconventional”
“Clean design”
“Exclusive/fashionable”
“Pleasurable to look at”
“Attractive/good looking”

When these questions were completed for the first four clusters, the test leader informed the subjects that a set of four other clusters should be assessed in the exactly same way. This was done to get a larger diversity of clusters.

After the subjects had viewed all eight clusters, an interview was conducted to first gather some background data of the subjects i.e. gender, age, education, if the subjects has a driver’s licence, if the subject drives a car, the subjects’ preferences concerning what a car instrumentation should include and how important the appearance of the instrumentation is. The rest of the interview was to get a deeper understanding on why the subjects had assessed the clusters in a particular way. Gaze duration heat-maps of where the subjects had gazed when considering the different statements were compiled and shown to the subjects. The subjects could study the heat-maps while the test leader asked for each cluster, why the subjects had gazed in the places shown by the heat-maps and why the particular assessments had been made. The questions asked were “Why did you look at this specific area?” and “Why did you give this cluster this specific assessment?”.
RESULTS AND DISCUSSION

Reading direction
All of the participants tended to look at the clusters in a Western reading direction. Kukkonen’s studies (2005) came to the conclusion that when viewing two pictures side by side, the left was looked upon first in 78% of the cases. In this study the average total viewing time spent on the left side was also 2796ms compared with 2283ms for the right side. When looking at four objects together on one screen as in the current study, another pattern arose – all subjects started to follow the Western reading direction when looking at the clusters. All subjects started to look at the centre of the screen when no clusters were displayed. When the clusters appeared on the screen, their gazes went up to the top-left corner of the screen, followed by the top-right corner, the bottom-left and the bottom-right (Figure 4). However, as the subjects had already had a look at the clusters before they were all put together on the screen, this may have affected how long the clusters were viewed. A certain cluster may have aroused the subject’s like or dislike, for which reason the cluster would have been given further examination and thereby longer viewing time.

![Figure 4](image)

Figure 4 – One subject’s viewing order. The viewing is processed in the same order as reading a text in the western world; left to right, up to down.

Cluster order
The results were statistically analysed using independent samples t-tests. The clusters were analyzed in two groups, clusters 1-8 which were viewed by the subjects in groups 1 & 3, and clusters 9-16 which were viewed by the subjects in groups 2 & 4. The clusters were viewed in both forward and reversed direction. Significant differences between the orders are displayed in Table 1.
Table 1 – Significant differences between cluster orders.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Statement</th>
<th>Measured quantity</th>
<th>Mean, forward order(^a)</th>
<th>Mean, reversed order(^a)</th>
<th>Std. dev. forward order</th>
<th>Std. dev. reversed order</th>
<th>t</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 8(^a)</td>
<td>C Attractive/good looking(^1)</td>
<td>assessment</td>
<td>3.4</td>
<td>5.2</td>
<td>.89</td>
<td>.75</td>
<td>-3.50</td>
<td>.01</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>A Unconventional</td>
<td>viewing time</td>
<td>2741</td>
<td>15330</td>
<td>544</td>
<td>13533</td>
<td>-2.28</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>A Exclusive/fashionable</td>
<td>viewing time</td>
<td>3120</td>
<td>10005</td>
<td>1877</td>
<td>5479</td>
<td>-2.91</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>B Clean design</td>
<td>assessment</td>
<td>6.0</td>
<td>4.2</td>
<td>.89</td>
<td>.98</td>
<td>3.38</td>
<td>.01</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>B Exclusive/fashionable</td>
<td>viewing time</td>
<td>2108</td>
<td>8272</td>
<td>1239</td>
<td>6148</td>
<td>-2.41</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>B Attractive/good looking(^2)</td>
<td>viewing time</td>
<td>1797</td>
<td>7901</td>
<td>955</td>
<td>3472</td>
<td>-4.15</td>
<td>.001</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>D Attractive/good looking(^1)</td>
<td>assessment</td>
<td>3.2</td>
<td>4.7</td>
<td>.41</td>
<td>1.03</td>
<td>-3.31</td>
<td>.01</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>E Attractive/good looking(^1)</td>
<td>assessment</td>
<td>3.7</td>
<td>5.2</td>
<td>1.21</td>
<td>.41</td>
<td>-2.88</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>E Attractive/good looking(^2)</td>
<td>assessment</td>
<td>3.0</td>
<td>5.2</td>
<td>1.27</td>
<td>.98</td>
<td>-3.31</td>
<td>.01</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>E Exclusive/fashionable</td>
<td>assessment</td>
<td>3.0</td>
<td>5.3</td>
<td>1.67</td>
<td>1.63</td>
<td>-2.45</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>E Attractive/good looking(^2)</td>
<td>assessment</td>
<td>3.3</td>
<td>5.5</td>
<td>1.63</td>
<td>1.23</td>
<td>-2.60</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>F Easy to understand…</td>
<td>viewing time</td>
<td>2400</td>
<td>4671</td>
<td>1034</td>
<td>1766</td>
<td>-2.72</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>G Attractive/good looking(^1)</td>
<td>viewing time</td>
<td>2030</td>
<td>4717</td>
<td>902</td>
<td>2738</td>
<td>-2.28</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>H Easy to understand…</td>
<td>viewing time</td>
<td>4581</td>
<td>6978</td>
<td>1589</td>
<td>1968</td>
<td>-2.32</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>H Unconventional</td>
<td>viewing time</td>
<td>3944</td>
<td>9584</td>
<td>2183</td>
<td>5112</td>
<td>-2.49</td>
<td>.05</td>
</tr>
<tr>
<td>9 to 16(^b)</td>
<td>H Pleasurable to look at</td>
<td>viewing time</td>
<td>2269</td>
<td>4849</td>
<td>1244</td>
<td>2203</td>
<td>-2.50</td>
<td>.05</td>
</tr>
</tbody>
</table>

\(^a\) N=11  
\(^b\) N=12  
\(^1\) The first time the statement appears to the subject 
\(^2\) The second time the statement appears to the subject  
\(^c\) The dimension of viewing time is [ms], the dimension of assessment is [grades]

For cluster B (clusters 9-16) four significant differences for viewing times and assessments between the cluster orders. This was not the case for exactly the same cluster B in clusters 1-8. A possible explanation is that in the interviews, this particular cluster received the most widespread opinions and that some subjects had the opposite opinions from other subjects. The small number of subjects might have caused the subjects that viewed clusters 9-16 to have a different opinion from the subjects that viewed clusters 1-8. If more subjects had participated, it is possible that the subjects with the most extreme assessments would have been distributed more equally between the groups of different cluster orders.

For clusters E and H also several significant differences were found. It can be concluded that E and H are of similar design, but they need to be examined further in order to find a possible explanation of the differences.
Correlations in “Attractive/good looking” between both times it was displayed Independent sample t-tests did not show any significant differences between the assessments of the reference pictures B and F appearing in both clusters 1-8 and clusters 9-16 concerning the statement “Attractive/good looking” the first time and the second time they appeared. Correlation analysis with Spearman’s rho show significant correlations between the two statements, rs(22)= .50 p=.05 and rs(22)= .57, p=.001, respectively. This implies that “Attractive/good looking” the second time it appeared can be used alone as a measure of attractiveness. “Attractive/good looking” the first time it appears can then be removed from further analysis since it does not follow exactly the same procedure as the other statements; hence in the following analysis all the four clusters are viewed on the same screen.

Correlation between viewing time and assessments of the clusters
The results were statistically analyzed with bivariate correlations with Spearman’s rho. The seven-grade scale in the assessment task was transformed here into a three-grade scale in order to identify correlations more easily (Hair, Money, Samouel, Page, 2007). The three-grade scale has the value “low” for the former values 1, 2 and 3, the value “medium” for the former value 4, and “high” for the former values 5, 6 and 7. Two outliers, subjects 7 and 15, with excessive viewing time were removed from the group that viewed clusters 9-16. The question “Attractive/good looking” is only analysed for the second time it was shown to the subjects due to the results in the cluster order.

Table 2 – Significant correlations between viewing time and assessment.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Cluster</th>
<th>Statement</th>
<th>Spearman’s rho</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 8</td>
<td>D</td>
<td>Unconventional</td>
<td>.645</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>Exclusive/fashionable</td>
<td>.775</td>
<td>.01</td>
</tr>
<tr>
<td>9 to 16</td>
<td>E</td>
<td>Pleasurable to look at</td>
<td>-.640</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>Pleasurable to look at</td>
<td>.775</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>Attractive/good looking</td>
<td>-.643</td>
<td>.05</td>
</tr>
<tr>
<td>1 to 16</td>
<td>B</td>
<td>Pleasurable to look at</td>
<td>.438</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td>Pleasurable to look at</td>
<td>.534</td>
<td>.05</td>
</tr>
</tbody>
</table>

The results in the present study indicate that there is no clear correlation between how long a cluster is viewed and its assessments. Only a few significant correlations can be found when the assessments and viewing time are compared statement by statement. Some correlations are even negative (Table 2). Since there are few correlations and some of them are negative, there seems to be no general correlation between viewing time and assessments. But since the number of subjects is rather small in the present study, this could affect the correlation results.
Identification of details
There was an assumption that details which caught the viewers’ attention could easily be located by the help of eye-tracking (Hammer & Lengyel, 1991). In the present study the eye-tracking equipment can clearly distinguish whether the subjects gazed at the speedometer or tachometer or other key features of the clusters. However it is hard to say if the small details such as fonts, graphical details, or symbols the gauges consist of were looked at. Only a few specific details that caught particular attention could be identified for a few subjects. Instead the participants viewed the clusters in a more holistic way. Especially when the statement “Attractive/good looking” was considered, the gaze duration heat-maps for all subjects indicated that short fixations were spread all over the clusters. This in line with the view offered by the Gestalt psychologists; the small elements that a picture is built of are viewed in parallel as one entity (Monö, 1997; Duchowski, 2003). This is a possible explanation of why specific details in the present study were never identified. However, eye-tracking scan paths strongly suggest that the viewing and identification of objects are accomplished in a serial way where smaller objects are viewed instead of the whole picture (Duchowski, 2003). There are several competing theories attempting to explain these serial and parallel viewing behaviours (Duchowski, 2003), but in the present study it is difficult to say which of them represents the subject’s viewing behaviour. Another factor that could affect this result is that the clusters appeared downscaled on the screen. Hence the different parts of the clusters could have been too small to draw attention to them.

Interviews
The interviews with the subjects based on the gaze duration heat maps of the eye-tracking data provided a great deal of information but seemed to lack some precision. Since the interviews were carried out after the actual viewing of the clusters, the subjects had to recall how they viewed them. This did not always seem to be very accurate and the participants could afterwards easily be led to make up a deceptive reason for giving a particular assessment. This could give a completely different reason than the one they were actually thinking of when giving the assessment. This requires demands that the interviewer take great efforts not to ask leading questions or in other ways mislead the subjects. An example of how a subject responded in the interview: Person 26 for Cluster C and the statement “Easy to understand and read out information” on the question “Why did you look at this specific area?”: “I tried to understand what was written in the gauge”. On the question “Why did you give this cluster this specific assessment?”: “I was not used to it which makes it hard to find my way in the gauge”.

Limitations of the study
The subjects used were mainly from the same department of Luleå University of Technology. This is a limitation since it is uncertain if this relatively homogenous group will affect the results of this study.

The number of subjects in the study is rather small and they were to view many clusters each. More subjects are needed to validate the results from the present study.

All clusters in the present study were displayed as static pictures. This differs from a real driving situation where information in the cluster is constantly changing. This will probably affect the way subjects view the clusters, but since the study seeks to evaluate visual perception of appearance and not usability, this should not be an issue.

Future research
In the present study, the statements in the assessment task are printed on sheets of paper. Further research could examine if repeatedly taking the eyes away from the screen may have any influence
on the results. The quality of the eye-tracking data should not be corrupted in the present study since the eye-tracking equipment can resume the eye-tracking as soon as the subject’s eyes are back on the screen. Further research could be made to examine if answers from an on-screen assessment solution differ from answers made on paper.

CONCLUSIONS

The findings in this study indicate that:

- No conclusions can be drawn concerning the cluster presentation order.
- The study indicates that there is no major correlation between the time spent viewing a cluster and its assessment.
- The interviews give an understanding of why the clusters were given a certain assessment.
- It is difficult to identify specific small details in the clusters that draw a particular amount of attention to them. The eye-tracking equipment can however distinguish when different parts in the cluster, e.g. speedometer and tachometer are viewed.
- The interviews would probably be more difficult to carry out without the guidance of the eye-tracking.

REFERENCES


Küller, R. (1975), Semantisk miljöbeskrivning (SMB), Psykologiförlaget, Stockholm.

Laike, T. 1999, ”Att mäta upplevelse av bilinteriörer”, Lund, Sweden, Lund University, Department of EnvironmentalPsychology.


