A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/123380

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

University of Warwick
WMG
Experiential Engineering

UNIVERSITY OF WARWICK

NEW KNOWLEDGE AND METHODS FOR MITIGATING DRIVER DISTRACTION

ENGINEERING DOCTORATE | INNOVATION REPORT

Author   Claudia Geitner

Academic Mentors Professor Paul Jennings and Dr. Stewart Birrell

Industrial Mentor Lee Skrypchuk

Submitted February 2018

Abstract

Driver distraction is the diversion of attention to a non-driving related activity. It has been identified as major cause of accidents. Even as we move away from traditional 'driver' and towards highly-automated vehicles, distraction remains an important issue. A distracted driver could still potentially miss a handover of control message from the car, or have a reduced awareness of the traffic environment. With the increased number and complexity of new features being introduced in vehicles, it is becoming more important to understand how drivers interact with them, to understand the benefit they offer in helping the driver to focus on-road, but also to identify their limitations and risks. Thereby it is important to consider that the interaction between human and technology, e.g. driver distraction, can be described by many aspects. To learn the most about the interaction between user and technology, it is important to select a suitable measure and to utilise that measure in best practice, which can be hard to find in literature. This research project is divided into two research streams that investigate the opportunities of new in-vehicle interfaces to mitigate driver distraction and that research how to efficiently identify measures for the ergonomic evaluation of in-vehicle interfaces.

Research stream one, comprising four studies, evaluated tactile information as a new interface technology to mitigate distraction in manual and automated cars. Tactile perception requires physical contact between the driver and the device delivering the feedback. It can be decreased by clothing. In the first user trial it was evaluated, for the first time, how shoe type, gender, and age influence the driver's perception of a tactile pedal. Shoe type did not, but gender, age, and the feedback's duration and amplitude did influence the perception. In some durations and amplitudes, the feedback was recognised by all participants and was rated highly intense, both aspects a warning should have. Next, it was evaluated how fast people would react to a tactile warning compared to a traditional auditory warning and an auditory-tactile warning. The participants reacted significantly slower to the tactile warning. Following, a tactile warning might not be suitable as an in-vehicle warning. However, adding an auditory component to the tactile warning increases its efficiency and people missed less auditory-tactile compared to auditory warnings. Newly introduced interfaces, such as tactile interfaces, put an effort on drivers to adjust to them and might lead to unsafe interactions. In the third and fourth study, it was investigated how a driver's trust effects the reaction time and glance behaviour. Trust was not associated with the reaction time towards a tactile warning signal, but it influenced the glances at a voice-navigation interface that was new for the majority of the participants. The findings can be utilised to increase the trust in the interface dialogue and thereby decrease a driver's time glanced off-road.

Research stream two investigated how Human-Machine-Interface (HMI) engineers can be supported in the comparison and selection of measures (e.g. a usability score) to evaluate the ergonomics of in-vehicle devices, for example to measure driver distraction. Industry projects are often restricted by tight deadlines and limited availability of equipment. Measure selection can then become a time critical issue. In published literature, there existed no guidelines to support this task. In four rapid prototyping evaluations, an interface was developed that can aid HMI-engineers in the comparison and selection of measures for an ergonomic evaluation. The tool functions as knowledge management and foresees to inform users about the best practice to utilise a measure, tips to set-up required equipment, and templates for the measure, for example templates for the analysis or electronic versions of questionnaires.

Acknowledgement

ENGINEERING DOCTORATE MENTORS

Academic Mentors:

Professor Paul Jennings, Dr Stewart Birrell

Industrial Mentor:

Lee Skrypchuck

DECLARATION

I hereby declare that all of the work contained in this report was produced by the author and that none of the work has been previously submitted for an academic degree. All sources of quoted work have been referenced accordingly.

Claudia Geitner

Table of Contents

List of Figures

List of Tables

Abbreviations

ABS	Anti-lock Braking System
CSV	Comma-Separated Values
EngD	Engineering Doctorate
HCD	Human-Centred Design (process)
HF(E)	Human Factors (Engineering)
HMI	Human-Machine Interface
JLR	Jaguar Land Rover
MMW	Multimodal Warning
NHTSA	National Highway Traffic Safety Administration
PDT	Peripheral Detection Task
RSAP	Rapid Serial Auditory Presentation
RSP	Rapid Serial Presentation
RSTP	Rapid Serial Tactile Presentation
RSVP	Rapid Serial Visual Presentation
RT	Reaction Time
SAE	Society of Automotive Engineers
SUS	System Usability Scale
UEQ	User Experience Questionnaire
VDT	Visual Detection Task
VIS	Visual Information Seeking

1. Introduction

Drivers, like all humans, are social beings and need to be understood in their social context. Whereby their current role is that of a driver, there are many other social roles that they need to fulfil as well. Other roles might include being the employee on the way to an important meeting, a consultant, a friend for someone in need, or a parent. Driving has become safer over the years and the car is easier to manoeuver. This ease in the driving task leads to more spare capacity that drivers have. When spare capacity appears while driving it provides room to engage in those other social roles simultaneously, for example to make a business call and ensure everything is set up for the meeting, or to attend to the crying baby in the back of the car. Drivers are distracted when they engage in non-driving related activities while driving (submission 1). Besides switching between social roles this also happens while driving and listening to music, adjusting the navigation system in the car, selecting another track of music on their mp3 player, or eating and drinking.

A crucial finding is that attention not spent on-road increases the reaction time to sudden changes on-road, which results in a higher risk of road accidents for distracted drivers (Dingus et al., 2006; Mohebbi et al., 2009). Driver distraction is a major contributor to road accidents (Dingus et al., 2006). The growing number of non-driving related services available in the car (Press, 2014), and connection to other people that has become a custom habit for many (Hope, 2016) increase temptations for distractions. This development is of concern to road safety administrations. For example, it led to the regulation of the use of handheld mobile phones while driving (Gov, 2017). It is therefore important to understand how drivers are affected by newly introduced technology and how effects of distraction can be mitigated to minimise road accidents.

The tactile modality is a new interface technology with the potential to be less distracting as it does not require the driver to take the eyes off-road to perceive the information. However, its potential and limitation are not fully understood yet. For example, perception of a tactile stimulus requires physical contact, though, the clothing a person wears can reduces that contact. Besides potential and limitations, it is important to understand how drivers interact with a newly introduced interface. Some drivers might be reluctant towards technology. Specifically, for those it is

important to ensure that interactions with a newly introduced technology while driving do not have a negative impact on their driving behaviour. For example, sceptical drivers might feel the need to monitor a new device more often, that could result in an increased distraction from the road.

Safety administrations developed, not yet binding, design guidelines for in-vehicle devices to reduce driver distraction (JAMA, 2004; NHTSA, 2013). For automobile manufacturers those guidelines are important as they might result in a regulation in the future. Consequently, the topic of driver distraction needs to be considered when designing new in-vehicle systems. Subsequently the interaction with in-vehicle systems needs to be tested with users (drivers) against criteria that affect driver distraction. To learn the most from such a test of the in-vehicle system with users it is important to use a suitable measure and to utilise that measure most efficiently. This can be difficult for Human-Machine Interface (HMI) engineers, because the decision can involve many measures and a description of best practice can be hard to find. Further, other project constraints, such as the available equipment and available time for the user trial, can require a time-consuming literature review to make an informed decision about the most suitable measure. For industry, it would make the preparation for a user trial and its conduction more effective if the HMI engineers would be guided in the measure selection process and if best practice knowledge for utilisation would be collected.

This Engineering Doctorate (EngD) project explored interface modalities as a mitigation strategy for driver distraction and it explored a new conceptual interface design to support HMI engineers in the ergonomic evaluation of in-vehicle devices. Seven documents describing research projects were uploaded into the portfolio. All are summarised and compared to the project's research aims in this Innovation Report. The next sections define the research aims of this EngD, provide an overview of the submissions in the portfolio, and suggest a reading order for the portfolio.

1.1.   Research aims

This EngD project is conducted in the frame of Human Factors Engineering (HFE). HFE "*... is concerned with ways of designing machines, operations, and work environments so that they match human capacities and limitations*" (Chapanis, 1965). Matching a tool to human capabilities and limitations makes it more "usable" to support a user in a task. Usability, according to ISO 9241, "*... is the effectiveness, efficiency and satisfaction with which specific users can achieve a specific set of tasks in a particular environment*". In this research project, that means to improve existing solutions to mitigate effects of driver distraction by understanding the users' capabilities and limitations better. The effectiveness of a strategy can be understood by comparing how users interact with an old and with a new solution, for example, comparing a tactile information in a car to an existing visual or audio information.

The application of the Human Factors (HF) approach to the research meets the objectives of the sponsoring company of this EngD project, Jaguar Land Rover (JLR). The Chief Executive Officer of JLR, Professor Dr Ralf Speth, stated in the annual report (Jaguar Land Rover (c), 2017, pp. 6):

> "*Our customers are at the heart of everything we do. Our passion and our purpose are to meet and exceed their aspirations; to delight them with experiences they will love for life.*"

JLR's business strategy foresees to invest in innovations to keep up with the larger changes the industry is expected to undergo in the next years, such as the development of higher automated vehicles (Jaguar Land Rover (c), 2017). A part of the innovation derives from understanding how users interact with in-vehicle devices to increase a positive experience. For example, HFE methodology helps to decide about the most effective strategy to mitigate distraction. Further, it also helps to understand how users are affected by new technology. For an automobile manufacturer safety of the drivers is a selling point, but a number of customers expect and would prefer to buy a vehicle with the latest technology in-vehicle systems (Aloisio and Mrasek, 2017). Whereas some users are enthusiastic about the latest technology, it can be demanding and perhaps distracting for other, sceptical, users. Therefore, the first research stream of this EngD project investigated how new technology affects drivers and how driver distraction can be mitigated.

Research questions stream 1: Contribution to mitigating effects of driver distraction

1) What is driver distraction, and when and how does driver distraction occur?
2) What strategies could an automotive company employ to mitigate driver distraction?
3) Haptic feedback has been shown to be less visually distracting for the driver, however, what variables influence the perception of haptics?
4) Can a tactile warning as such or a tactile warning enhanced by another modality initiate a faster reaction time compared to a traditional auditory warning?
5) How does a driver's trust in technology effect the visual interaction with a new in-vehicle device?

The saturated markets in Europe and the expected large changes in industry through automation put competitive pressure on automotive companies. To remain innovative and to employ innovations best, it is required to manage knowledge in the company efficiently. Such knowledge comprises gaining the most insights from the interaction between driver and an in-vehicle device in user trials. An essential part to obtain such insights from a user trial is to utilise the most suitable HF measures to observe the interaction and to apply them in best practice. It can be difficult and there is no defined procedure to select the best measures from the manifold that are available, and further a measure likely needs to suit limitations in time and resources of an industrial research project. Another obstacle is to learn about its best practice application. Typically, colleagues are asked, but are not always available due to their own workload. In the second research stream it was aimed to fill this gap and develop a novel tool that can aid HMI engineers in the comparison and selection of HF measures for their user trial and that can function as knowledge management for best practice application of measures.

Research questions stream 2: Contribution to support HMI engineers in their task to understand, compare, select and utilise HF measures for the ergonomic evaluation of in-vehicle interfaces

1) How do designers select measures for user studies?
2) Can measure selection benefit from electronic support, and, if so, how can designers be supported in their task in a usable way?

1.2.   Research project structure

Figure 1 presents an overview of the submissions to this EngD portfolio. Each box represents a document of the EngD portfolio. The arrows indicate how the documents are related to each other. The literature review is suggested to be the first document to read (Figure 1 (1)). It sets out the project frame. It defines driver distraction and discusses why research in this topic is important for academia and for automotive industry. Based on the definition of driver distraction, potential research areas are explained and evaluated with regards to their relevance to automotive industry. Two research streams emerged from the literature review, that were already introduced in Section 1.1: one focusing on practical strategies to mitigate driver distraction in in-vehicle communication (Figure 1 (2 - 5)) and the other aiding HMI engineers in the ergonomic evaluation of new in-vehicle devices to learn the most about the users (Figure 1 (6)). Dependent on the reader's interest, it is suggested to follow either research stream one or stream two first.
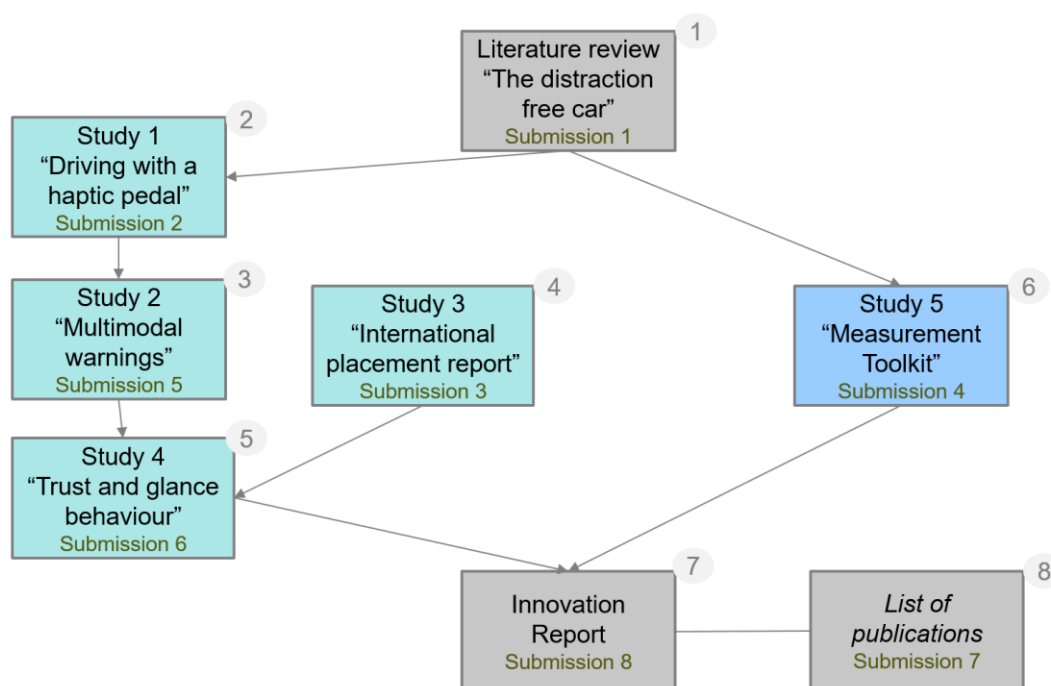
Figure 1. Overview of the documents in the EngD portfolio.

Research stream one describes research towards mitigation of driver distraction through new interface modalities and the effects that new interface technology has on drivers. The Haptic Pedal study evaluated the perception of haptic pulse feedback presented by a pedal. For the first time it was evaluated how age, gender, and shoe type influence the perception of a haptic pulse delivered by a pedal (Figure 1 (2); Section

3.1). The Warning study explored the effectiveness of tactile feedback as a warning in a distraction scenario in a self-driving car. It was evaluated how effective tactile feedback would perform as an in-vehicle warning compared to a traditional auditory warning in a highly demanding automated driving scenario with an emergency brake event (Figure 1 (3); Sections 3.2 and 3.3). The warnings were compared objectively by reaction time (RT) and subjectively by how they were perceived.

Tactile interfaces in a car were new to the majority of the participants in the Warning study. In the international placement, the Trust study was conducted which investigated factors that influence how drivers interact with an interface that was new to them (Figure 1 (4); Section 4.1). This research indicated a link between pre-exposure trust and a driver's glance behaviour. The Trust 3navi study looked into a more detailed understanding of the visual interaction between drivers and new in-vehicle interfaces in three levels of visual demand (Figure 1 (5); Section 4.3).

Research stream two describes the development of a novel conceptual interface that supports HMI engineers in the user trial process (Figure 1 (6); Chapter 5). User trials are an essential step in the development process of new in-vehicle systems. In a user trial, either the participants are asked about a task or technology, or the participants are observed whilst performing a task with a new in-vehicle interface. Only through user trials it can be understood whether a new interface is desirable for participants, whether it is usable, and whether it is not distracting while driving. There are a range of measures that can be used to evaluate driver distraction. A novel visual interface was developed that helps to collect those measures and provides HMI engineers with an easy way of comparing the measures and to access information on how to use a measure and how to analyse data obtained with the given measure according to the best practise from industry and academia.

Each of the following chapters outlines the above-mentioned submissions into the EngD portfolio and further provides a summary of the outcomes. For a detailed description about the research project please read the dedicated submission. The research outcome is summarised and compared to the project's objectives (Section 1.1) in the final Chapter 7 of this Innovation Report.

2. Driver distraction

"Alice is not a very good navigator. Often she tends to get lost in unfamiliar locations. Soon, she plans to visit her uncle who just moved to the countryside. To assist her in the journey, Alice bought a new navigation app for her smartphone that she can use while driving. The app works with voice command and has the latest interface technology. Before the journey, Alice puts the phone in the holder and starts driving. She knows the way out of the city onto the motorway. Driving on the motorway is very monotonous and feels easy, she plans to set up the navigation then. On arrival at the motorway she starts the app. Whereas it appeared easy to use in the shop it turns out the voice command can be time consuming. Alice learned that voice recognition requires training in order to be easy to use. This is not what Alice expected. As then the traffic suddenly gets congested driving requires more attention, there might be an accident. Because she feels that she needs to concentrate on the busy traffic she decides to set up the navigation and to train the voice command interaction at the next available motorway park."

This story describes exemplary how people try to integrate new technology into their lives. When the driving task is not very demanding, such as on a relatively empty monotonous motorway, opportunities emerge for drivers to engage in other activities, for example, to try a new device as helpful assistant. Every year more people use mobile devices which can be used everywhere, including while driving (Statista, 2018). People expect those devices to be helpful. At times, this leads to disappointment. Most people have an expectation about how the device works, but are rarely aware of the technical details and it can easily take more time to learn a new device than expected. Interaction with a new device can specifically become challenging while driving. Drivers are then required to share their attention between those activities and driving, becoming distracted from driving. Whereas driving can feel easy on some occasions, such as the monotonous relatively empty motorway, the demand can suddenly increase, such as when approaching a potential accident scenery. Distracted drivers are more vulnerable to miss important information on-road and react delayed to such sudden changes in driving demand (Dingus et al., 2006). Research in driver distraction is important for road safety and aims to understand how these dual task situations, and specifically interaction with newly introduced devices, affect the driving behaviour.

The arising safety concerns by traffic administrations (JAMA, 2004; NHTSA, 2013) contradict the consumers' push towards the application of new technology and integration of mobile devices in cars (Aloisio and Mrasek, 2017). The integration of the latest HMI technology in in-vehicle systems has become a major criterion for the decision to buy a certain car for consumers, independent of the car's brand (Aloisio and Mrasek, 2017; Scuro, 2017). Some in-vehicle functions help to reduce the impact of a traffic accident, for example the Anti-lock Braking System and airbags for pedestrians (Volvo, 2016; Jaguar Land Rover (a), 2014). Other new in-vehicle functions provide the driver with additional information, such as a blind spot warning system (Jaguar Land Rover (d). 2017), or inform the driver in a way that reduces glances off-road such as the tactile modality (Brown, 2005; Birrell et al., 2013) or Head-Up displays (Liu, 2003). However, those technologies involve risks to cause more distraction. For example, Head-Up displays can lead to visual tunnelling: the driver focuses on the displayed information but misses important information from the traffic scenery (Ward and Parkes, 1994). In addition, there is a growing set of non-driving related functions, such as the option to connect the smartphone to the car (Turkus, 2014), listening to music, or access internet-based services. To understand how interacting with those technologies contributes to driver distraction and what effects that has on safe driving, it first needs to be defined what driver distraction is.

The next sections summarise findings from the literature review, answering research question 1 from research stream 1 (Section 1.1). At the beginning, driver distraction is defined (Section 2.1), the importance of driver distraction is discussed in general and in specific for automobile manufacturers (Section 2.2), and, finally, the impact of driver distraction on vehicle design is described (Section 2.3). The Chapter ends with areas for potential research that is valuable for automobile manufacturers.

2.1.   What is driver distraction and how can it be measured

For over 60 years automobile manufacturers offered appliances for leisure activities besides driving. For example, Chrysler introduced the Highway HiFi, an in-vehicle record player, in 1956. Research on driver distraction followed soon after, it therefore has a history equally as long. John Senders, one of the pioneers in this area, evaluated the amount of visual attention required for driving on a highway in the 1960s. With the emergence of mobile technology in the 1990s, driver distraction became a

concerning regulatory topic. The first regulations banning the use of hand-held mobile phones while driving emerged in 2001. In 2013, the National Highway Traffic Safety Administration formulated the well-known guideline for minimising distracting in-vehicle interfaces (NHTSA, 2013).

While the HMI of the car changes, the main purpose of the driving task does not change and that is travelling safely from point A to point B. In order to do so drivers are required to comply with the Highway Code in the United Kingdom, pay attention to the road and avoid distraction (Highway Code, 2017, rule 148). Hence, drivers need to control their vehicle, be aware of the traffic situation, and detect changes in the traffic situation. Based on the detected changes, future changes and potential hazards need to be estimated. According to that estimate, the driving behaviour must be adjusted to avoid potential hazards. Thereby, the level of attention on the road needs to be sufficient to sample necessary information about the traffic situation, e.g. potential hazards, and interpret it.

The required level of attention on-road to drive safely varies dependent on weather, road, and traffic conditions (The et al., 2014). When drivers engage in another task in parallel to driving, attention needs to be divided between both tasks. Multitasking costs mental resources, which are naturally limited (Wickens et al., 2002). While switching between tasks, drivers might be unaware of the information they miss on-road. Indeed, research indicates that drivers might not be able to self-assess their level of distraction from the driving task correctly (Horrey et al., 2008). Gaps in attention on-road can lead to accidents when the demand of the traffic situation suddenly increases and the driver is unable to react timely (Greenberg et al., 2003).
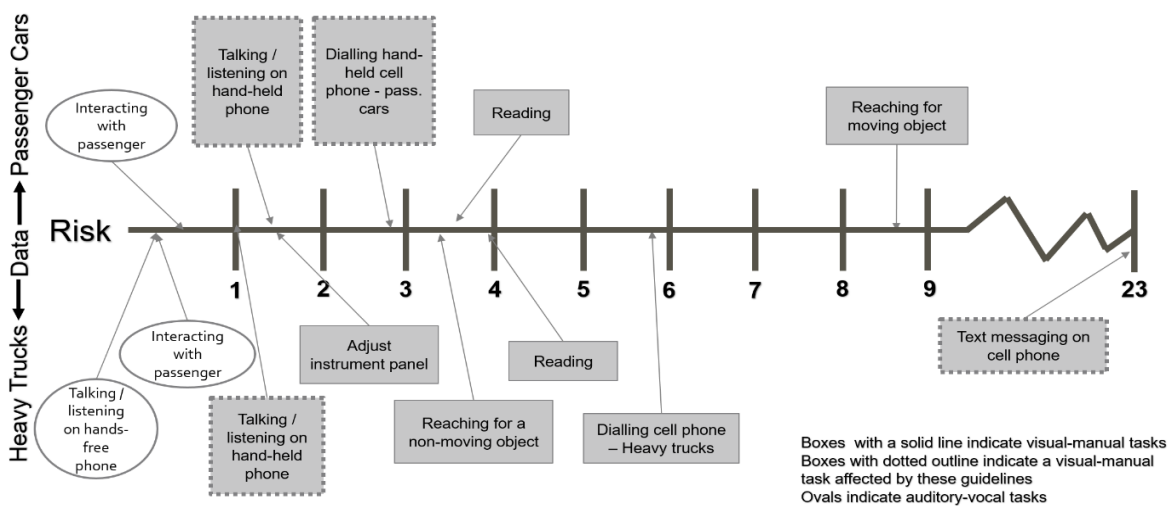
Figure 2. Accident risk of non-driving related tasks conducted while driving (NHTSA, 2013, Figure 1).

Indeed, research identified driver distraction as one of the major causes of accidents, specifically when drivers glance at locations off-road (Dingus et al., 2006; Klauer et al., 2006; Wilson and Stimpson, 2010). The activities in which drivers engage are manifold, comprising smoking, phone conversations, listening to music, eating, drinking, reaching for objects in the car, and talking to a passenger (Klauer et al., 2006; Stutts et al., 2001). The risk of accident varies between the activities (Figure 2). An activity with a very high risk to cause an accident is texting (NHTSA, 2013). Therefore, drivers are prohibited to use hand-held mobile phones while driving in many countries. The accident risk while interacting with new interface designs is an unknown component. That is why it is important to assess how newly introduced interfaces or functions in the car influence the driver's behaviour, beginning with an assessment of driver distraction in the design process of new in-vehicle interfaces.

To compare driver distraction across multiple studies and to facilitate a general understanding of driver distraction and its implications on safety, it is necessary to define driver distraction. A definition that is widely used and therefore is adopted in this thesis can be found in Pettit et al. (2005). Pettit et al. (2005) describe driver distraction as: "Driver distraction occurs when:

- A driver is delayed in the recognition of information necessary to safely maintain the lateral and longitudinal control of the vehicle (the driving task) (impact)
- Due to some event, activity, object or person, within or outside the vehicle (agent)
- That compels or tends to induce the driver's shifting attention away from fundamental driving tasks (mechanism)
- By compromising the driver's auditory, biomechanical, cognitive or visual faculties, or combinations thereof (type)"

This definition describes driver distraction ranging from a source, to effects on the driver, and further to effects on the driving task. According to the definition, driver distraction delays information processing required for safe driving which results in a deviation from safe driving. Safe driving includes vehicle control, but also observation of the traffic situation. The impact, according to the definition, can be visible or invisible on the level of information processing (observing the traffic environment). A visible impact would be if a driver fails to recognise a speed sign and, in consequence,

fails to adjust the vehicle's speed. A non-visible impact would be if a driver fails to observe all streets at an intersection. Non-visible impacts of driver distraction are difficult to measure in real traffic situations. It might require hindsight to judge if the driver's recognition of on-road information in a safety critical situation was sufficient or not. A subjective rating of driver distraction, comparable to a rating of workload, does not appear to be feasible. Horrey et al. (2008) reported that subjective distraction ratings do not correlate with the actual decreased driving performance.

Pettit et al. (2005) suggest four types of distraction: visual, biomechanical, auditory, and cognitive driver distraction. However, often distraction occurs in a combination of types, such as adjusting a navigation system which involves visual-biomechanical distraction and a phone-conversation which involves auditory-cognitive distraction. The definition of types is useful to determine measurements. A match of those types of distraction to sensory channels indicates potential measures for driver distraction. The easiest to measure is visual distraction with an analysis of the driver's glance patterns, e.g. in Liang, 2009. Biomechanical distraction occurs when the driver takes a hand away from the steering wheel and can be measured by observing the driver. Biomechanical distraction typically occurs in combination with visual distraction for coordination. Auditory distraction can effect the perception of a warning sound from the vehicle. It can be measured indirectly by measuring the noticeability of a signal from the car. Otherwise, auditory distraction can occur in combination with cognitive distraction and can be measured with those effects. Measurement of cognitive distraction is more difficult as it occurs in the driver's head and is only indirectly measurable. For example, cognitive distraction can affect the glance pattern, resulting in a decreased radius of eye glances mainly to the road centre (Recarte and Nunes, 2003; Engström et al., 2005). Further a high cognitive load leads to a decreased variability of the lane change position (Reimer et al., 2009), it can lead to a reduction in speed (Reimer et al., 2009), and it can reduce the monitoring of the traffic environment (Liang, 2009).

The definition of driver distraction from Pettit et al. (2005) applies principally to a manual driving context. Distraction in a manual driving scenario as the standard scenario in current cars remains important for research within the development as long as the driver has the ability to drive manually in the car. However, considering the increasing amount of assistant systems in the car, it is important to understand

that the driving task, and hence distraction, changes in a highly automated car. With increasing electronic assistance and automation, the driving task becomes more passive, where drivers might only need to observe the system for any failures. A task previously defined as a 'distraction' can then even become the primary task, with occasional glances to the automation system. In addition, the range of distraction can be wider, for example, it can comprise more visual tasks which are difficult to conduct while manually driving a car, and the engagement in a distractor task can be increased compared to a manual driving.

2.2. Why is driver distraction important for automobile manufacturers

Safety administrations are concerned about the effects of driver distraction. The National Highway Traffic Safety Administration (NHTSA) (2013) proposed a guideline to design less distracting in-vehicle systems for automobile manufacturers, the "Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices". Currently it is a voluntary guideline. However, automobile manufacturers adhere to it, as it might become law in the future.

The NHTSA guideline includes proposals for in-vehicle design, for example, in-vehicle systems should avoid e-books, scrolling text and videos (NHTSA, 2013). Another section of the guideline deals with recommendations for measurements of driver distraction. The recommendations focus on the detection of visual distraction, because it is directly observable, and safety critical (with their eyes off-road drivers are not able to recognise events on-road), and relatively (compared to cognitive distraction) easy to measure. NHTSA's recommendations can be summarised in a rule: 12/2. According to the NHTSA guideline, an in-vehicle device is minimally distracting when drivers can complete their tasks with it without taking their eyes off-road for more than 12 seconds (s) over the whole task, and no longer than 2 s at a single glance off-road.

The sponsoring company integrated the NHTSA guideline in its design process. Prototypes of new systems are validated against the NHTSA design criteria, in the driving simulator or earlier on. Additionally, automobile manufacturers investigate in new forms of interfaces (Table 1), with potential to minimise the distraction to the driving task, such as Head-Up displays (Jaguar Land Rover (b), 2014), voice command (BMW, 2015) and gesture control (Jaguar Land Rover (a), 2017) for in-vehicle devices.

Automobile manufacturers invest a large number of resources into the development of cars with partly automated aspects of the driving task (Jaguar Land Rover (b), 2017). Automation involves challenges as well, for example how to communicate its limitations and boundaries to a non-technical audience (Norman, 1990). However, as long as the driver needs to respond to the vehicle, distraction remains an important design criterion that impacts the time a driver needs to react and the accuracy of the response. Higher levels of automation enable drivers to use resources previously needed for driving and hence they can engage in other tasks even more. This might require the design of a more efficient warning system, which is able to capture the drivers' attention even while they are deeply engaged in another task. A more efficient warning might be able to reengage drivers faster in the traffic situation compared to a traditional warning.

Table 1. Research areas (published) of automobile manufacturers

Company	Self-driving car	Driver Assistant / Safety Systems	New interface technology
JLR	(Jaguar Land Rover (b), 2017)	Airbags for pedestrians (Jaguar Land Rover (a), 2014)	Head-Up displays (Jaguar Land Rover (b), 2014) Gesture control (Jaguar Land Rover (a), 2017)
Volvo	(Volvo, 2017)	Airbags for pedestrians (Volvo, 2016) Connected services (Volvo, 2017) Collision avoidance system (Volvo, 2017) Steering assistance (Volvo, 2017)	Head-Up displays (Volvo, 2017)
VW	(Hawkins, 2017)	Blind sport warning Lane-change assist Emergency Braking assist (VW, 2017) Modular warning systems that work together and synchronise their feedback (VW, 2017)	New design of the instrument cluster Multimodal warning systems (visual, auditory, and haptic components) (VW, 2017)
Toyota	(Toyota, 2014)	Cooperative adaptive-cruise control (Toyota, 2014) Lane Trace Control (Toyota, 2014) Automated Highway Driving Assist (Toyota, 2014)	3D Head-Up display (Toyota, 2014)

2.3. Mitigation strategies for driver distraction

This section discusses research opportunities combining mitigating driver distraction as a major cause of fatal road accidents and interests of automobile manufacturers in that topic. The broad topic of driver distraction comprises research possibilities in measurement, analysis and mitigation of driver distraction. Figure 3 shows an overview of potential research topics (as presented in the literature review, submission 1). Not all research topics are of interest for automobile manufacturers. Regulation plays a large role in driver distraction as well and determines the use or prohibition of certain devices while driving (such as hand-held mobile phones), and

the definition of the formal process to obtain a driver license are managed by traffic agencies. The potential areas were each evaluated for their relevance to automotive industry and the most relevant were selected for research projects in this EngD.
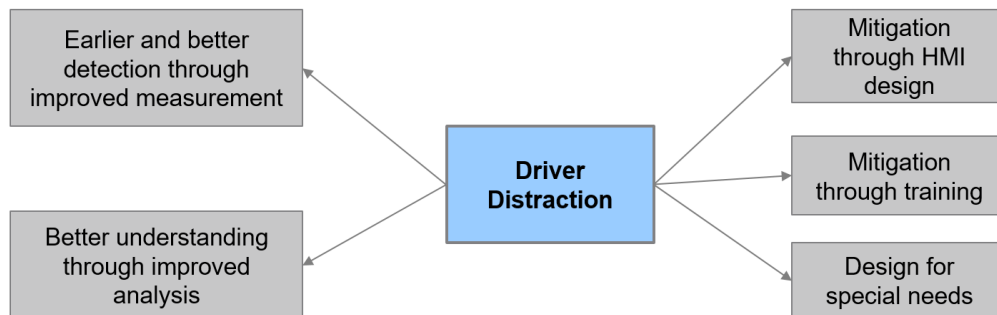
Figure 3. Mitigation strategies for driver distraction.

- *Earlier and better detection through improved measurement:* Driver distraction can be measured with many measures from literature. This research concerns the improvement of measurements for distraction, enabling an earlier detection of a critical status of the driver. Obtained knowledge can influence research to better understand the cognitive processes that shape the multitasking and cognitive capabilities. This topic is of interest for automobile manufacturers as well. New insights can change the way an in-vehicle device is evaluated in the design process. For example, the NHTSA guideline (NHTSA, 2013) proposes an evaluation criteria for visual driver distraction. This suggested process focuses on the observation of glance behaviour that requires effort in set-up and analysis (much more effort compared to a questionnaire). The development of more efficient measures for driver distraction can keep the guideline up to date, reduce resources needed for in-vehicle evaluation and increase insights for in-vehicle design by identifying variables that make a system less demanding to use.
- *Better understanding through improved analysis:* Understanding the driving task with its sequence of actions, requirements of the sub-steps and most critical steps including consequences of failures, helps to improve the detection of failures earlier on to minimise negative consequences of failures. Applied to driving it means to understand the requirements of the driving task and then to analyse consequences if failures happen at certain points to find a way to mitigate. Similar to the topic of improved measures it is of interest for

automobile manufacturers, because it could include new insights into the in-vehicle design.

- *Mitigation through HMI design:* Emerging technology has a potential to improve deficiencies of current interfaces, e.g. tactile interfaces do not require visual attention. Research has shown that glances off-road are a major contributor to fatal accidents (Klauer et al., 2006), a reduced time looking off-road helps to reduce fatal accidents. This topic is specifically interesting for automobile manufacturers, because it overlaps with their research interest in alternative HMI, the increased development of driver assistance systems and compliance to guidelines such as the NHTSA (2013) in the design process of new in-vehicle interfaces.

- *Mitigation through training*: Research found that drivers have difficulties to detect how distracted they are. Training that shows the consequences of driver distraction and that raises awareness might influence the behaviour of drivers positively. For example, a Google cardboard (GoogleVR, 2018), or a similar device, and a mobile phone offers a cheap possibility to create a virtual experience. Such an immersive experience might have greater effects compared to paper based training. However, automobile manufacturers can influence this topic only to a minimal extent, it is more responsibility of regulation and driving schools as part of the driver education.

- *Design for special needs*: This research considers challenges for groups of drivers with limited abilities, for example drivers with disabilities. What is distracting for them? How could the situation be improved? This research is less relevant for automotive industry as it is a different, smaller, target market.

From the areas of potential research interest and a potential interest to automotive industry, two research streams were pursued in this EngD project.

Research stream one investigated the mitigation of driver distraction through HMI design by exploring how drivers engage with new interface technology. Tactile feedback is evaluated as a less distracting alternative of in-vehicle communication. First, previously unevaluated factors influencing the perception of tactile feedback were investigated (Haptic Pedal study, Section 3.1). Then, the effectiveness of the tactile modality as warning was investigated (Warning study, Sections 3.2 and 3.3). A tactile warning was compared to a traditional auditory warning and to a multimodal,

auditory-tactile, warning (MMW). The following three studies focused on an understanding how drivers interact with technology that is new to them (Trust study, Section 4.1; Trust Brake study, Section 4.2; Trust 3navi study, Section 4.3). Specifically, it was focused to gain an understanding of aspects that influence long glances off-road (>2 s).

Research stream two is related to the mitigation of driver distraction through "Earlier and better detection through improved measurement". However, no research into a new measure was pursued rather efficient knowledge management about existing measures was investigated. The research focused on the process of planning and conducting a study with users to evaluate an in-vehicle device (Toolkit study, Chapter 5). In the course of the project, an interface concept was developed to organise Human Factors (HF) related measures, including measures for driver distraction, to ease comparison of measures, make it easier for HMI engineers to select measures for their user trial, and to share best practices of measure application. The interface is called HF toolkit. The HF toolkit project has immediate impact on the automobile company and is of organisational relevance, helping to integrate new employees and sharing knowledge between HMI engineers. Knowledge is a valuable asset of a company (McMahon et al., 2014) and sharing it facilitates good quality efficient work, especially since the number of Jaguar Land Rover employees doubled over the last four years (Jaguar Land Rover (c), 2014). This research project is related to the area of measures for driver distraction, but instead of investigating a new measure, it aims to provide a fast overview of the existing measures and description how to utilise them in best practise.

3. Tactile as alternative in-vehicle interface modality

Driving is a mainly visual and cognitive task (Hancock et al., 2002). Glances off-road, specifically such longer than 2 seconds (s), are considered safety critical (Klauer et al., 2006; NHTSA, 2013). Presenting the driver with non-visual in-vehicle information helps to reduce the time glanced off-road, for example, by auditory or tactile feedback.

In tactile interfaces the information is communicated directly to the driver without informing others in the car and it is hard to ignore (Van Erp, 2002). However, tactile feedback is only perceived if the driver is in physical contact with the device submitting it (Van Erp, 2002). The physical contact can change when the driver shifts his position in the seat, or with clothes, or other factors. The Haptic Pedal study explores factors that can affect the perceived intensity and comfort of a tactile pulse (Section 3.1).

Drivers can engage in many distracting tasks varying in demand and modality. According to the Multiple Resource Theory, the perception of a warning can decrease if task and warning need to be perceived in the same sensory channel (Wickens, 2002). Tactile warnings might perform better than auditory or visual warnings as they can be perceived directly. However, an MMW might still lead to better reaction times over multiple distracting conditions. An MMW consists of at least two modalities, and can so better utilise a sensory channel that is not utilised by a distractor task. In the Warning study (Section 3.3) the performance of three warnings (auditory, tactile, and auditory-tactile) was compared in three distracting conditions (visual, auditory, and tactile). Settings for this study were determined in three pilot studies that are described in Section 3.2.

Table 2. Overview of the studies presented in this Chapter.

Study	Aim	Method	Participants
Haptic Pedal study (question 3)	This study aimed to assess the influence of shoe type, age, and gender on the perception of haptic pulses delivered by a pedal	Subjective perception of 16 pulse settings in two shoe conditions. The pulse was played to the participants in a stationary, but running car with a haptic pedal prototype installed	36
Warning study Pilot A	This study aimed to create a tactile warning and to determine a setting for a clearly noticeable vibration	Lab study, subjective perception of vibration feedback	3
Warning study Pilot B	This study aimed to determine the setting of the components for the MMW	Lab study, subjective perception of the MMW	9
Warning study Pilot C	This study aimed to select three similar distractor tasks in three modalities and select a similar difficulty level in all three.	Lab study, subjective perception of the difficulty and workload that were imposed by the distractor tasks on the participant	18
Warning study (question 4)	This study compared the RT to, and subjective perception of an MMW, an auditory and a tactile warning over three highly attention capturing distractor tasks (visual, auditory, and tactile).	Driving simulator study, performance (subjective and objective) of three warnings compared over three distractor tasks in a self-driving car scenario with an emergency brake event	45

This Chapter describes research questions 3 and 4 from research stream 1 (Section 1.1). Table 2 summarises the two studies and the pilot studies presented in this Chapter. The Haptic Pedal study (REGO-2014-1312) and the Warning study (REGO-2016-1741 AM01) were ethical approved by the University of Warwick's Biomedical & Scientific Research Ethics Committee (BSREC).

3.1. Understanding haptic perception of a pedal

Tactile interfaces offer an opportunity to communicate without taking the eyes off-road. Previously, tactile feedback has been evaluated as communication means for warnings, navigation, and eco-driving (Fitch et al., 2007; Adell et al., 2008; Chang et al., 2011; Birrell et al., 2013). The evaluated interfaces comprise the seatbelt, seat, steering wheel, and pedals (Chang et al., 2011; Birrell et al., 2013; Spence and Ho, 2017). A study even found tactile feedback as the preferred modality for communication (Adell et al., 2008). In order to initiate an interaction between human and machine, the machine's feedback needs to be perceptible for the human (Norman, 2002). As tactile information requires the driver to be in physical contact (Van Erp, 2002), clothes are a potential influence on the perceptibility. For a pedal this specifically concerns shoes. Naturally, shoe types vary over the year, being thin in summer, and stiff and thick in winter. Only few studies considered a potential influence of shoes previously (Abbink and Van der Helm, 2004). Other factors within a person that can influence haptic perception are age and gender. For example, females tend to be more sensitive to tactile feedback (Hale and Stanney, 2004). As with other abilities, such as vision, it appears that tactile sensitivity declines with age (Brown, 2005). In the Haptic Pedal study, the combined effects of shoe type, age, and gender on the perception of a tactile pulse delivered by a pedal are evaluated for the first time.

3.1.1. Method and procedure

Overall, 36 people took part in the Haptic Pedal study, thereof 21 males and 15 females. The participants were sampled from three age groups: "39 and younger" with 11 participants, "40 – 59" with 13 participants, and "60 and older" with 12 participants.

The Haptic Pedal study was conducted in a stationary but running Range Rover Evoque with a prototype of the haptic pedal installed as accelerator pedal. Two shoe types were provided to the participants. One type were sneakers with thin flexible soles

(approximately 8 mm) and the other type were safety boots with thicker stiff soles (approximately 14 mm). The study is limited in that the material of the shoe has not been further analysed in terms of its vibration characteristics. The shoes were available in sizes spread over 95% of the population. The shoe were provided to avoid unwanted variations.

The tactile pulse was varied in amplitude and duration, ranging from just noticeable feedback reported in previous studies by Abbink and Van der Helm (2004) and Ichinose et al. (2013) to increased amplitudes and durations (Table 3).

Table 3. Sixteen pulses were employed in the Haptic Pedal study; Top - Force amplitudes in Newton (N), Middle - Duration of the pulse as frequency in Hertz (Hz), and Bottom - Duration of the pulse in Milliseconds (ms).

7 N	9 N	14 N	18 N	7 N	9 N	14 N	18 N	7 N	9 N	14 N	18 N	7 N	9 N	14 N	18 N
1 Hz	1 Hz	1 Hz	1 Hz	15 Hz	15 Hz	15 Hz	15 Hz	30 Hz	30 Hz	30 Hz	30 Hz	50 Hz	50 Hz	50 Hz	50 Hz
1000 ms	1000 ms	1000 ms	1000 ms	67 ms	67 ms	67 ms	67 ms	33 ms	33 ms	33 ms	33 ms	20 ms	20 ms	20 ms	20 ms

The participants rated each pulse immediately after experiencing it. The pulse was rated in the same three questions, always in this order: perceived intensity, perceived urgency, and perceived comfort. In case the participant did not notice the pulse, the rating was noted in the intensity scale and no further ratings for urgency and comfort were taken.
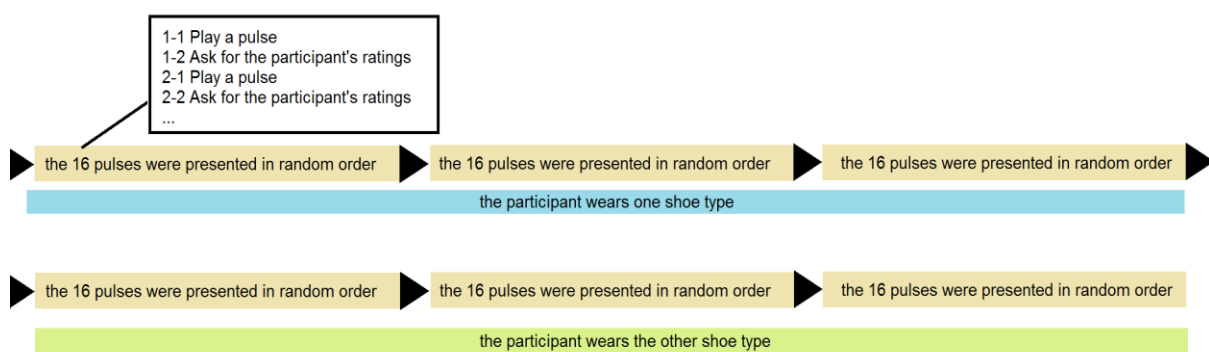
Figure 4. Procedure of the Haptic Pedal study.

The Haptic Pedal study started randomly with either one of the shoe types, whereby the order was counterbalanced. Participants sat in the car, started it and depressed the accelerator pedal. Only when the pedal was depressed and kept in a stable position at 1,500-2,000 RPM (revolutions per minute of the engine) a pulse was played. The sixteen pulses from Table 3 were presented in randomised order, three times in each shoe condition. Figure 4 visualises the procedure.

3.1.2. Results

The data analysis was conducted in R (R Core Team, 2014). An ANOVA analysis was conducted to evaluate effects of shoe, age and gender on the ratings of intensity of the haptic pulses, each with a critical value of $p<0.05$. The ANOVA included an error calculation for the within subject variables (shoe, force amplitude and duration).

The shoe type did not have a significant influence on the perception of a haptic pulse ($p>0.05$). This result differs from the findings in Abbink and Van der Helm (2004). A reason might be the different study setting. For example, the Abbink's and Van der Helm's utilised a laboratory setting and the pedal was pressed with a force amplitude against the participant's foot. Contrary, in the Haptic Pedal study there was no counterforce on the pedal. The participants themselves kept their foot in a certain position and there was a slight vibration through the running car. This slight vibration might have decreased perception of just noticeable pulse settings. Previously differences in perception between shoes were shown for just noticeable feedback (Abbink and Van der Helm, 2004). When such perceptions are diminished, differences between shoe types might no longer observed.

However, the study has a limitation concerning the shoe type. The material of the shoe sole has not been analysed for its ability to convey vibrations. The effect of different sole materials on the perception of a haptic pulse remain unknown. Future research would need to clarify the effect and its size as the pulse from the pedal appears to be perceived over joints and tendons in the leg rather than the sole of the foot.

Gender was a main significant effect on the perception of the pulses ($F(1,30)=5.05$, $p=0.03$). Females rated tactile pulses higher compared to males, similar to findings in (Hale and Stanney, 2004). Further, females tended to rate high intense tactile pulses more negative in comfort compared to males. Age did not have a main effect in this study. However, there were significant interactions between age and force amplitude ($F(6,90)=2.39$, $p=0.03$); between age and duration ($F(6,90)=2.45$, $p=0.03$); and between age, duration, and gender ($F(6,90)=2.77$, $p=0.01$). Older participants missed a higher percentage of short pulses (20 milliseconds (ms) and 33 ms) compared to the youngest age group.

In consistency with previous findings, pulses were better perceived with a higher force amplitude (Abbink and Van der Helm , 2004), except for pulses with a duration of

20 ms. In addition, an increasing frequency made a low force amplitude better perceptible, this applied to all except the 7 Newton (N) force amplitude. However, pulses rated as high in intensity and with no missed pulses over all participants tended to be rated negatively in comfort.

Besides these results, submission 2 includes recommendations for the duration and amplitude of pulse feedback. Details are presented in the submission. The findings can improve the development of haptic pedal feedback, for example for eco-driving (Birrell et al., 2013; McIlroy and Stanton, 2017), or speed warning (De Rosario et al., 2010). Haptic feedback can become annoying over longer periods of experience (Van Erp, 2002), for that reason future studies would benefit from an evaluation of the haptic experience over a longer timeframe.

3.2. Pilot studies

The following sections describe three pilot studies that were conducted to determine settings for the Warning study. The Warning study investigated the performance of a tactile, auditory (traditional) and auditory-tactile warning across three highly attention capturing tasks. Warning study Pilot A, describes the development of the haptic seat cushion that was planned to be used as interface for the tactile warning and tactile component of the MMW. The warnings were designed in pilot study B and the distractor tasks were developed in pilot study C. The pilot studies helped to determine the settings for the driving simulator study, the Warning study.

3.2.1. Pilot study A: Haptic seat cushion

The haptic seat cushion consists of two seat cushions. Six vibrating motors are placed in two rows between those cushions. It was intended to place the cushion on the driver seat and participants would have perceived the feedback as a vibration in the seat. The seat cushion was developed and a first evaluation started with the aim to determine a good noticeable vibration setting. Within the first participants it was noticed that the frequency in which the motors vibrate changed with the weight of the participant. Frequency of vibration effects the perceptibility of a tactile cue (Abbink and Van der Helm, 2004). It is not practicable to select participants with equal weight for a study. To avoid this confounding variable, it was decided to use the ButtKicker system (ButtKicker, 2017) which had been installed in the simulator in the meantime. The

ButtKicker system generates vibrations with a vibrating motor that is installed on the back of the driving simulator seat (Figure 5). When the motor vibrates, the vibrations are distributed towards the seat and to the driver. Because the driver does not sit on the motor directly, this system is robust against the influence of the driver's weight.
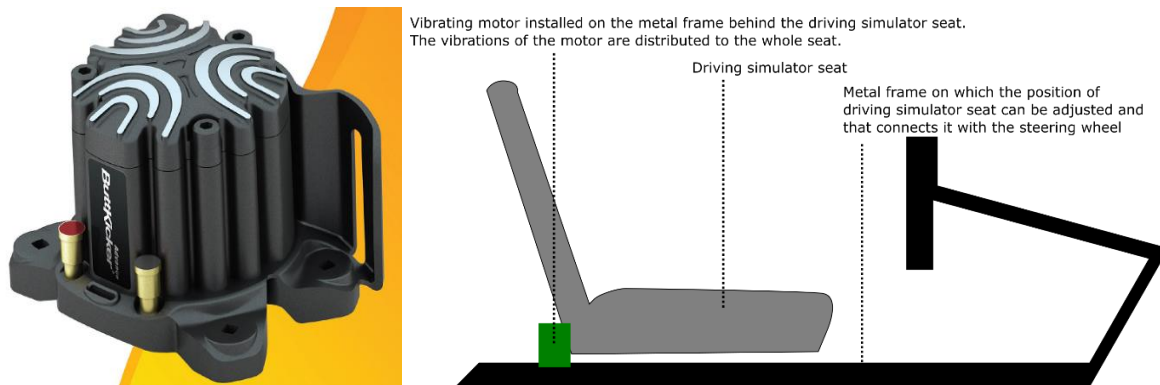
Figure 5. ButtKicker motor (left) and its implementation to the driving simulator seat (right).

Another outcome of the Warning study Pilot A is a literature review of haptic seat designs, which guided the selection of a minimal feasible design, and that can be used in the company to compare designs and use-cases of haptic seats (submission 2).

3.2.2.   Pilot study B: Warning design

In Warning study Pilot B the warnings for the study were designed and a setting was determined in which they are perceived as similar in noticeability. The auditory warning was obtained from the sponsoring company, a sound that is used as warning in a car on-road, to make the comparison realistic and relevant for automotive industry. The sound was presented for 2 s and consisted of a sequence of identical high frequency beeps. In a car, the warning is presented at approx. 70 decibel (dB), or it is should be at least 15 dB higher than the surrounding sounds (Lees and Lee, 2007). The driving simulator gave auditory output of the driving scene (the engine of the simulator car and engine sounds of other cars in the scenario). This output had a noise level between 45-60 dB. Following recommendations for the noise level of a warning, the auditory warning was planned to be presented at 70 dB.

The MMW was the combination of the auditory and tactile warning. The design followed rules for multisensory integration to ease two stimuli (an auditory and a haptic) to be perceived as belonging together:

-    All components of the auditory-tactile warning were presented non-spatially (Spence and Ho, 2017; King and Calvert, 2001)

- The association between the auditory-tactile warning components can be enhanced with a shared pattern, making auditory and haptic stimuli more similar (Keetels and Vroomen, 2012)
- All components of the auditory-tactile warning were presented synchronous (Wilson et al., 2009; King and Calvert, 2001)

A shared characteristic between auditory and tactile warning is supported by the nature of the ButtKicker system. The ButtKicker takes an audio file as input and presents its low frequencies as vibration output. Consequently, a copy of the auditory warning was created with the high frequencies filtered out. Only low frequencies, below 20 Hz, remained, that are not perceived by the ear. This altered file was then played on the ButtKicker. The result was a vibration of the same length and in the same alternating rhythm as the auditory warning.

Next, a study was conducted in the 3xD driving simulator to determine a setting for the vibration output in which it is perceived as similar in noticeability as the auditory warning. The procedure of the study was adapted from Stanley (2005), who used the procedure to design an auditory-tactile warning as well. The auditory warning was presented at 70 dB continuously, measured from the driver seat with a sound level meter. Alternating, the participants were presented with the tactile warning first either in lowest intensity (not-perceptible) or in highest intensity. Slowly the intensity was either increased or decreased. Whenever the participants perceived the tactile warning as equally in intensity to the auditory warning, they were instructed to inform the experimenter. Then, the intensity of the tactile warning was noted and the procedure started again in reverse order. Overall, the procedure was conducted six times. The six intensity scores for each participant were summarised to one average intensity score for each participant. Then the intensity scores were averaged over all participants. This average score was used as the setting for the intensity of the tactile warning in the Warning study.

Overall, nine people participated in Warning study Pilot B, six males and three females (Table 4). The lowest average setting was 3.25 and the highest average setting was 6.25 (on a scale ranging from 0 to 11). The average setting over all participants was 4.18 (approximately 100 Hz). This setting was subsequently applied in the Warning study.

Table 4. Results of Warning study Pilot B. The intensity settings for the vibration that are perceived equally intense to the audio warning presented at 70 dB from the simulator's loudspeaker.

Parti-cipant	Gender	Intensity asc. 1	Intensity desc. 1	Intensity asc. 2	Intensity desc. 2	Intensity asc. 3	Intensity desc. 3	Mean intensity / participant
1	m	3	3	3	3	3.5	4	3.25
2	m	7	7	6.5	7	4	6	6.25
3	f	3.5	6	3.5	6	3.5	6	4.75
4	m	5	5.5	3	6	3.5	6	4.83
5	m	3	3.5	4	4	6	5.5	4.33
6	m	3	4	3.5	4	4.5	4	3.83
7	f	2.5	3	2	3	2	2.5	2.5
8	m	4	3	4	3	5	4	3.83
9	f	3	2.5	6	3.5	6	3.5	4.08
							Overall mean intensity rating:	4.18

3.2.3.   Pilot study C: Distractor tasks

Warning study Pilot C was conducted to determine three high attention-capturing tasks in three modalities that are perceived similar in workload. At first, a literature review was conducted to obtain an overview of existing secondary tasks that could potentially be used for the Warning study. For the study it was aimed to select a distractor task that imposes a continuous level of high demand at a set pace, that is possible to conduct during the whole duration of the scenario, and that can be presented in a similar form in different modalities. It was aimed to present the task in the three sensory modalities that are typically utilised in interface design: vision, hearing, and touch.

Table 5. Detection tasks from research literature.

Detection task	Modality	Description	Reference
PDT	Visual	LED signals appear on the windscreen at random positions in the driver's normal field of sight. They appear within three to six seconds and have a duration of one second. When an LED signal appears the driver needs to respond, by clicking on a button on the index finger, within 200 ms to 2 s after onset, otherwise it is a miss. Drivers should detect as many signals as possible without decreasing their driving performance (driving has first priority).	(Jahn et al., 2005) (Martens and Van Winsum, 2001)
VDT	Visual	It is an extension of the PDT. An LED is presented periodically in the driver's central field of vision. It is presented every three to five seconds for a duration of at most three seconds. It disappears as soon as the driver presses a button.	(Young et al., 2012) (Santangelo and Spence, 2007)
Rapid Serial Visual Presentation (RSVP) task	Visual	A stream of random letters, and in between randomly a number, is presented to the participant. The participant needs to react only to a number by tapping on the display or pressing a button.	(Soto-Faraco and Spence, 2002)
Rapid Serial Audio Presentation (RSAP) task	Audio	This task is similar to the RSVP task. The stream of random letters is presented verbally. Randomly, numbers are spoken instead of a letter. The participant needs to react only to a number by tapping on the display or pressing a button.	(Soto-Faraco and Spence, 2002)

The distractor tasks from literature were narrowed down based on the above mentioned criteria. Table 5 shows tasks that were considered and tested for their suitability. The Peripheral Detection Task (PDT) and the Visual Detection Task (VDT) were excluded in favour of the Rapid Serial Presentation (RSP) tasks. It appeared easier to vary the RSP tasks in modality compared to the PDT or the VDT. In addition, the RSP tasks already exist in two modalities in literature: auditory and visual.

In RSP tasks, the participants are presented with a set of random signals. Each signal appears for a predefined timeframe. After exceedance of this timeframe, the signal disappears and no signal is presented for a predefined timeframe. Thereafter, the next signal appears. The participants should react only to a sub-set of signals, targets. Whenever the participants perceived targets in the RSVP and RSAP tasks, they should tap on the screen of the tablet computer on which the task is presented (Figure 6). The whole set of signals were numbers and letters in the RSVP and RSAP task. All numbers were targets. Next, settings for the RSP task in different modalities were compared to decide about the settings for Warning study Pilot C.
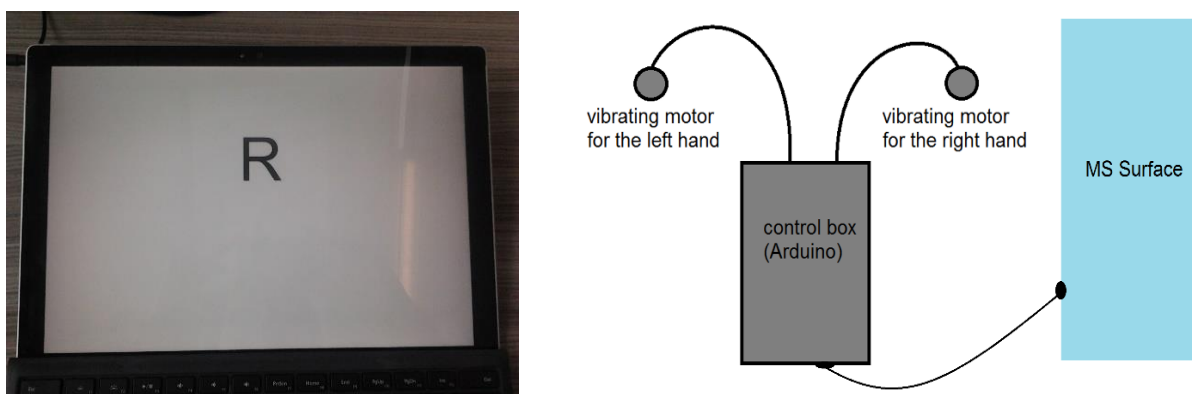
Figure 6. Visual task (left) and motors of the tactile task (right).

The settings for the visual and auditory task were adapted from literature (Soto-Faraco and Spence, 2002). This led to the following set of numbers and letters in the visual task and in the auditory task:

- Set of distractor stimuli: B, C, D, E, F, J, K, L, M, N, P, R, S, T, X, Y, Z
- Set of target stimuli: 2, 3, 4, 5, 6, 9

There appeared to be no comparable design for the tactile task in research literature. The tactile task needs to be a simple detection of change from the previous stimulus without learning. Multiple motors that vibrate alternating could generate different

stimuli, with one of them defined as target. The sponsoring company lent a haptic kit for the use in Warning study Pilot C and the Warning study. The kit consists of two vibrating motors controlled by an Arduino that was connected to a tablet computer (Figure 6). The participants put the tablet in their lap, and held one of the motors in the left and the other in the right hand for the duration of the task. The following easy distinguishable stimuli were presented in this task: left motor vibrating, right motor vibrating, both motors vibrating, and no signal presented for a predefined time frame. Two motors vibrating at the same time were defined as target and when that happened the participant needed to tap on the tablet screen.

A vibration should have a minimum duration of 20 ms to be noticeable on the foot (submission 2). In a test with friends and myself, it was difficult to remain attentive to a tactile signal of 20 ms duration on the fingers. A presentation for 40 ms was better perceptible over a longer period of time. Further, tactile receptors in the skin need time to recover from the perception of a tactile input. In a test with friends and myself, 200 ms appeared to be the minimum time between two vibrations in order for them to be perceived clearly as two vibrations. Consequently, the vibrations were presented for 40 ms with a minimum break of 250 ms between.

Three settings were selected per task to have a variety of demand, but an acceptable duration of the pilot study of 30 minutes (min) per participant (Table 6). Therewith, Warning study Pilot C consisted of nine settings. A participant experienced all nine task-settings one time for 2 min. After experiencing a task setting, a participant rated the experienced level of workload on a scale from zero (very low) to twenty (very high) (Hill et al., 1992). The workload ratings were averaged for all participants for each task setting to determine the mean workload a setting imposed on the participants.

Table 6. Settings for Warning study Pilot C.

Task	Setting	Time display stimuli	Time blank	Number of targets	Minimum signals between targets	Duration
Visual	1	40 ms	260 ms	8	17	120 s
Visual	2	40 ms	160 ms	8	25	120 s
Visual	3	40 ms	80 ms	8	42	120 s
Audio	1	120 ms	200 ms	8	16	120 s
Audio	2	120 ms	150 ms	8	19	120 s
Audio	3	120 ms	100 ms	8	23	120 s
Tactile	1	40 ms	450 ms	8	10	120 s
Tactile	2	40 ms	350 ms	8	13	120 s
Tactile	3	40 ms	250 ms	8	17	120 s

Overall, 18 people took part in the study. The majority of participants was male (14/18). The control software was still a prototype, in 21 cases (out of 162) the performance data was lost. After Warning study Pilot C was conducted, the control software implementation was improved with an automatic safe-function and a backup-log file, minimising the risk of data loss.

Participants performed considerably higher in all conditions of the visual task compared to the auditory and to the tactile task (Table 7). Participants performed lowest in the most difficult tactile task, the average detection rate was 27.7%. Performance in the auditory task did not vary much over the levels. In all task difficulty levels the sounds were played with higher speed than normal, this could have contributed to the difficulty besides the pace of the task in all conditions.

Table 7. Results of Warning study Pilot C.

Task	Duration	Signal setting	Pace setting	Workload (mean)	Workload (median)	Correct detected
Visual 1	2 min	No. of targets: 8 Min no. of signals between targets: 17	Signal appeared for: 40 ms Time blank: 260 ms	64.4	65.0	97.7%
Visual 2	2 min	No. of targets: 8 Min no. of signals between targets: 25	Signal appeared for: 40 ms Time blank: 160 ms	66.7	70.0	92.4%
Visual 3	2 min	No. of targets: 8 Min no. of signals between targets: 42	Signal appeared for: 40 ms Time blank: 80 ms	76.0	80.0	81.0%
Auditory 1	2 min	No. of targets: 8 Min no. of signals between targets: 16	Signal appeared for: 120 ms Time blank: 200 ms	72.8	75.0	46.9%
Auditory 2	2 min	No. of targets: 8 Min no. of signals between targets: 19	Signal appeared for: 120 ms Time blank: 150 ms	74.5	72.5	44.6%
Auditory 3	2 min	No. of targets: 8 Min no. of signals between targets: 23	Signal appeared for: 120 ms Time blank: 100 ms	77.3	82.5	43.4%
Tactile 1	2 min	No. of targets: 8 Min no. of signals between targets: 10	Signal appeared for: 40 ms Time blank: 450 ms	65.6	70.0	51.9%
Tactile 2	2 min	No. of targets: 8 Min no. of signals between targets: 13	Signal appeared for: 40 ms Time blank: 350 ms	65.0	70.0	36.7%
Tactile 3	2 min	No. of targets: 8 Min no. of signals between targets: 17	Signal appeared for: 40 ms Time blank: 250 ms	68.1	70.0	27.7%

The participants were not required to perform well in the distractor tasks, but the distractor tasks should keep the participants engaged. A low task performance can indicate disengagement. That is why engagement was analysed and utilised to decide about a redesign of the tasks. For calculating the engagement, the 2 min long task was

divided into sections of 30 s. When a participant clicked in a 30 s interval (one or more times) the engagement score was increased by one, which resulted in a maximum score of 4 (and minimum 0 if the participant did not make any clicks at all). In two task settings participants did not interact during the task (Figure 7, p12 in auditory task 3 and p18 in tactile task 2). Because the data was analysed after completion of the pilot study, the reason remains unclear. The difficulty participants experienced in tactile and auditory task reflects in the engagement, whereby the majority of the participants remained physically engaged in the task. For that reason it has been decided to select a setting in the current task design for the Warning study.
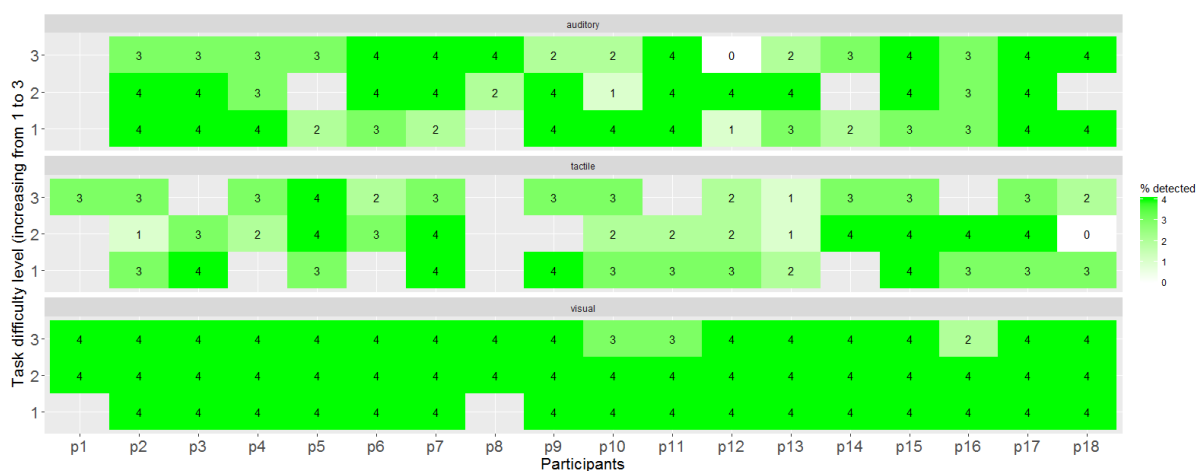
Figure 7. Engagement (a click in an intervals of 30s resulted in one score, scores ranging from 0 to 4).

The task settings were compared in workload by an average workload rating over all participants (Table 7). The highest workload rating was given for the most difficult auditory task ($M = 77.3$). The workload ratings increased with the pace of the visual task. As performance was still very good in the most difficult level (81.9 % detected) and that level had been used in previous studies, it was selected for the Warning study. Matching, comparably in workload rating the auditory task level 2 was selected for Warning study (visual workload rated median 80 and auditory median rated 72.5). The level 3 auditory task would have been a better match in terms of the workload rating (median 82.5), but auditory task 2 had a slightly better engagement score (Figure 7) and slightly better detection rate. For the same reason tactile task level 2 was selected for the driving simulator study.

3.3. Comparing tactile, auditory, and multimodal warnings

Drivers can potentially engage in manifold tasks while driving. The tasks can vary in demand and modality, ranging from listening to music, talking to another passenger, reading the weather forecast in the in-vehicle information display, adjusting the playlist in the multimedia system, to typing in an address in the navigation system. A most efficient warning would be well noticeable in as many as possible of these diverse situations. However, the perception of a warning can be decreased if a task and a warning use the same sensory channel according to the Multiple Resources Theory (Wickens, 2002).

Tactile warnings have been described as hard to ignore and direct (Van Erp, 2002) which might be an advantage as a modality for a warning. Previously also MMW were reported as being able to initiate a faster RT compared to unimodal warnings (Brown, 2005; Biondi et al., 2017; Spence and Ho, 2017). A faster RT to a warning is important for the design of a warning system, because it means more time for the driver to react in a safety critical situation. A faster brake reaction might prevent a collision. In addition, according to the Multiple Resources Theory, the noticeability of MMW could be higher across multiple distractor tasks compared to unimodal warnings. This is because a MMW consists of two modalities which increases the chance that one of the modalities in which the warning is communicated is not covered by a potential distractor task. Previously, tactile and MMW have mainly been compared in one distracting condition, for example, a phone conversation (Biondi et al., 2017).

The Warning study investigated in this gap and compared an auditory warning, a tactile warning, and a MMW over three highly attention-capturing distractor tasks in a self-driving car scenario with an emergency brake event. Settings for the warnings were utilised from Warning study Pilot A and settings for the distractor tasks were applied from results of Warning study Pilot B.

3.3.1. Method and procedure

Data of 45 participants was analysed for this study (26 female and 19 male). The majority of the participants (80%) was between 20-39 years old. The Warning study employed a three (warning) by three (task) factorial design of nine scenarios (Table

8). In each scenario the warnings were presented eight times. Previously this procedure has been utilised as RT naturally varies (e.g. in Biondi et al., 2017).

Table 8. Driving scenarios (each 5 min long).

Scenario	1	2	3	4	5	6	7	8	9
Task	Visual distraction RSVP task	Auditory distraction RSAP task	Tactile distraction RSTP task	Visual distraction RSVP task	Auditory distraction RSAP task	Tactile distraction RSTP task	Visual distraction RSVP task	Auditory distraction RSAP task	Tactile distraction RSTP task
Warning	multimodal	multimodal	multimodal	audio	audio	audio	tactile	tactile	tactile

Figure 8. Development driving simulator.

The Warning study was conducted in the development simulator (Figure 8). The study was set in context of an autonomous driving car in which the driver sits and engages in a non-driving related task. At a random point of time the autonomous driving vehicle presented a warning (auditory, tactile, or auditory-tactile) to which the driver needed to react as fast as possible by pressing the brake pedal. To embed the warning into a realistic context all warnings were presented in a convoy scenario. The driver's vehicle drove between one car following and two cars in front. At a random point in time the first car braked and in reaction to that shortly after the second car. The warning was played 2 s before the first car braked, setting a time frame in which the drivers should react. All warnings were presented in the same convoy context to minimise confounding the RTs by varying causes of the warning that can be perceived as difference in importance.

After each scenario, the participants rated the warning in the same four questions, always in this order: noticeability, motivation to respond, startlement, and annoyance.

After calculating the RT, this data was analysed for outliers and misses, similar as in (Biondi et al., 2017). RTs longer than 2.5 seconds or shorter than 0.4 seconds were excluded from the analysis. Overall, this affected 27 values (less than 1% of the data). The missing values were spread over the tasks and cue types. Those missing values were replaced by the mean RT value for the participant's existing RTs in this scenario.

The dataset was tested with a Mauchly's test and met the criteria of Sphericity. Then, a repeated measure ANOVA analysis was conducted, with the RT being the dependent variable, and cue type and task type the independent variables.

The rating data was tested for normality with the Shapiro-Wilk test. The ratings for noticeability, motivation, annoyance and startlement were not normal distributed. A paired Wilcoxon signed rank test as a non-parametric statistic was then applied for a within-subject variable comparison of the ratings across the three warning cues.

3.3.2. Results

The data analysis was conducted in R (R Core Team, 2014). The ANOVA revealed a significant main effect of task (F(2, 88), $p<$.001, generalised η^2=.81), and a main effect of warning (F(2, 88), $p<$.001, generalised η^2=.03), and an interaction effect between task and warning (F(4, 176), $p<$.001, generalised η^2=.01). The RT was then compared across the three warning types separately in each of the three distractor task conditions.

In the tactile task condition, the RT to the MMW (M = 1.28) and to the auditory warning (M =1.27) did not differ significantly, $p>$.05. Participants reacted significantly faster to the MMW compared to the tactile warning (M =1.33), $p=$.005. Participants reacted also significantly faster to the auditory warning compared to the tactile warning, $p<$.001. A slower reaction to the tactile warning in the tactile task condition is in accordance with the MRT (Wickens, 2002). In the auditory task condition, RTs did not differ significantly across the three warnings. However, participants reacted in mean faster to the MMW (M =0.75) compared to auditory warning (M =0.78) and to the tactile warning (M =0.78). However, the difference was not significant. In the visual task condition, RTs were faster for the MMW (M =1.25) compared to the tactile warning (M =1.3), $p<$.001, and to the auditory warning (M =1.23) compared to the tactile warning, $p<$.001. Previous literature indicates that visual information can decrease tactile perception (Auvray et al., 2008; Murphy and Dalton, 2016). RTs between MMW and auditory warning did not differ significantly, $p>$.05.

The MMW had a positive influence on missed warnings and false reactions. The most warnings were missed in auditory warning conditions (11 out of 19). In the MMW and in the tactile warning conditions less warnings were missed, four out of nineteen each.

False reactions occurred most in conditions with the tactile warning (43) compared to the MMW (20) and to the auditory warning (19). In terms of missed warnings and false reactions the MMW combined the best effects of both unimodal warnings.

The MMW (M = 6.3) in this study was perceived as subjectively more noticeable compared to the tactile (M = 5.6, V= 2499.5, p<.001) and the auditory warning (M = 5.9, V=1646.5, p<.001). The MMW (M = 6.2) was also rated as more motivating to respond to compared to the tactile warning (M = 5.7, V= 2437.5, p<.001) and the auditory warning (M = 5.9, V=2068.5, p<.001). However, as a trade-off the MMW (M = 4.5) was also perceived as more startling than the tactile (M = 3.95, V=3298.5, p<.001) and the auditory warning (M = 4.1, V= 2545, p<.001). However, all three warnings were rated medium startling. In accordance with previous research, all warnings were perceived as not annoying (Biondi et al., 2016).

4. Trust and distraction

Chapter 3 discussed alternative modalities to visual feedback. However, for some tasks it might be most feasible to utilise visuals to inform the driver as the information otherwise becomes hard to communicate, e.g. for a route calculated by the navigation system. For such interfaces it is important to reduce the amount of visual information so that drivers can process the information by a quick glance off-road. Besides the visual demand of the interface as such, there are other variables that influence how long drivers glance at the information, one of them is trust. Previously, it has been found that trust influences visual scanning behaviour and how much people engage with a system (Bagheri and Jamieson, 2004; Djamasbi et al., 2010; Gold et al., 2015). The aim of this research was to understand how the participants' trust in technology influences their interaction behaviour with a device that is new to the majority of them.

This Chapter describes research question 5 from research stream 1 (Section 1.1). Table 9 summarises the studies presented in this Chapter.

Table 9. Overview studies.

Study	Aim	Method	Participants
Trust study (Section 4.1)	Investigation of the relationship between trust and glance behaviour in a dual task situation while driving.	Secondary analysis of a previously conduced on-road study (Mehler et al., 2016) with variables that had not yet been investigated.	80
Trust Brake study (Section 4.2)	Investigation of the relationship between trust and experience, and trust and RT.	Secondary analysis of the data collected in the Warning study (Section 3.3) with variables that had not yet been investigated.	45
Trust 3navi (Section 4.3)	A detailed investigation of the relationship between trust and glance behaviour across three interfaces with varying visual demand.	Driving simulator study, eye tracking of the drivers entering addresses into three different navigation systems while driving on a motorway. The Trust 3navi study is based on the results of Trust study.	49

4.1. A link between trust and glance behaviour

The Trust study evaluated the relationship between pre-exposure trust and visual glance distribution while interacting with a voice-command system that was new to the majority of the participants. The Trust study was conducted as part of the international placement. The international placement report (submission 3) and the conference paper by Geitner et al. ((b) 2017) present it more detail. The research is a secondary analysis of a previous evaluation conducted by Mehler et al. (2016). It describes an analysis that has not been reported before.

4.1.1.   Method and procedure

Two in-vehicle-voice command navigation systems were used in the study, a 2013 Chevrolet Equinox equipped with the MyLink system and a 2013 Volvo XC60 equipped with the Sensus system (Mehler et al., 2016). Half of the 80 participants were assigned to the Volvo and the other half to the Chevrolet. A participant only experienced one of the voice-command systems. Participants were counterbalanced in gender and over four age groups (18-24; 25-39; 40-54; 55 and older; 20 participants in each group). After an introduction into the study, but before being exposed to the voice-command systems, the participants completed a demographic questionnaire that consisted of three questions related to trust (Table 10)

Table 10. Trust questionnaire.

1) How would you rate your overall level of trust in technology?

Very distrustful									Very trustful
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

2) How would you rate your level of trust in established car technologies?

Very distrustful									Very trustful
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

3) How would you rate your level of trust in new technologies that are being introduced into cars?

Very distrustful									Very trustful
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Each participant entered three addresses into the voice-command navigation system while driving on a motorway. Two addresses were the same for all participants. The third address was the participant's home address and varied. Previous research has shown that voice-command systems include visual components and attracted glances while interacting with them (Reimer et al., 2014). The drivers' switch behaviour between device and on-road was analysed with the glance metrics listed in Table 11. The metric of glances off-road >2 s was employed as a common metric from literature to assess driver distraction (NHTSA, 2013). The interaction between participant and navigation system was recorded on video. Later, that video was used to code the glance behaviour. First, values were calculated for each participant and then averaged. Only the two addresses that all participants had to enter were used for the analyses, to prevent differences in glance behaviour due to variations in the address. Data was analysed in R (R Core Team, 2014).

Table 11. Eye metrics for the analysis.

Eye-metric	Calculation
Percent of glance time to the device	Percentage of the total time glancing to the device (center stack region) for the tasks navigation entry 1 and 2. Calculation of the mean from both tasks.
No. of glances per minute	Number of glances (transitions across 9 coded glance regions, e.g. forward roadway, rearview mirror, device, left and right mirror/window, etc.) over the period of a task, for the tasks navigation entry 1 and 2, divided by time. Calculation of the mean from both tasks.
No. of glances to the device per minute	Separate calculation of the number of glances at the device over the period of the task for navigation entry 1 and 2. Calculation of the mean from both tasks.
No. of glances >2 s to the device per minute	Separate calculation of the number of glances at the device that were >2 s over the period of the task for the tasks navigation entry 1 and 2. Calculation of the mean from both tasks.

4.1.2.   Results

The voice-command interaction while driving was new for the majority of the participants (65%). From 80 participants, 28% had not experienced voice-command interaction before at all. For two of the three questions related to trust an association between the levels of trust and glance behaviour has been found. Trust in established car technologies is not statistically significant linked with any eye metric. High trust in car technology in general is associated with shorter total task time in the Chevrolet sample. High trust in new car technologies is associated with more frequent glances in general, over all coded glance locations (instrument cluster, mirror, road-way, centre stack, etc.). This relationship was observed in both voice-command systems.

The relationship between trust in new technology and glance behaviour included further interesting details in the Volvo system. In the interaction with the Volvo system high levels of trust were associated with a higher number of glances per minute, and with fewer longer duration glances (>2 s). This pattern was not found in the Chevrolet system.

The two in-vehicle systems utilise two different approaches for the address entry. In the Volvo system participants had to enter the address in steps for city, street and number, each with a confirmation. In the Chevrolet system the address was entered in one chunk. Both approaches have their advantages and disadvantages. The address entry was in average shorter in the Chevrolet system, but the interaction involved a higher number of user and system errors compared to the Volvo system (Mehler et al., 2016). When an error occurred in the Chevrolet system, the whole address needed to be entered again. In the Volvo system only the current step needed to be repeated in case of an error. It appears that whereas it took longer in the Volvo system to enter an address it included a better error recovery and fewer errors. Those differences in

system design could have affected glance behaviour. The glance behaviour in the Chevrolet group might have changed in the course of the interaction due to an altered level of trust by the more frequent errors.

The findings of the Trust study indicated interesting relationships between levels of pre-exposure trust and glance behaviour that can be exploited further in future research. Future research can shed more light on whether the observed glance patterns can be generalised and if they reappear in interfaces with different levels of visual demand. A better understanding of the relationship between trust and visual interaction might help human factors engineers to utilise trust better as a variable in system design and help the driver maintain a safe glance behaviour.

4.2. Link between trust and experience / reaction time

Alarms communicate critical information from the system to the user. Previous research showed that components such as the system reliability and predictability can influence a user's trust in the system and therewith how the user responds to alarms given by the system. Low reliability in the alarms led to lower alarm reaction frequency and appropriateness (Abe et al., 2002). Lees and Lee (2007) evaluated how the driver's trust and consequently behaviour is influenced by collision avoidance systems varying in utility, predictability, and reliability. Drivers who perceived unnecessary warnings in a non-critical routine driving context (such as passing an oversized vehicle or a parking vehicle that rearranged its position) became more sensitive to changes on the road situation and reacted more often and with a greater degree to non-critical driving events. In contrast, false alarms diminished trust and compliance to the alarms.

This study investigated the effect of trust and reaction to a warning that was new to the majority of the participants in a secondary analysis of data collected in the Warning study (Section 3.3). Safety critical warnings might occur seldom to the driver, but require a fast reaction. When signals that the driver is not familiar with are applied for such a warning it should be ensured that, even if the driver is unfamiliar with the signal, the RT is not increased. This study explored the link between trust and experience, and trust and RT.

This study (REGO-2017-2078) was ethical approved by the University of Warwick's BSREC.

4.2.1. Method and procedure

Data for this analysis was collected in the Warning study presented in Section 3.3 (submission 5). The study evaluated the subjective perception of, and reaction to three types of warnings in three distracting conditions (resulting in 9 scenarios). One warning was a traditional auditory warning as it is utilised in commercially available vehicles. Another warning was a vibration of the whole seat (tactile warning). The third warning was a combination of both, auditory and tactile warning presented in parallel. The study was conducted in the development simulator, consisting of a racing car seat, steering box, pedals and a set of three monitors on which the virtual environment was presented. The participants drove in a self-driving car along a virtual cross-country road and conducted a high attention capturing distractor task. While the participants conducted the distractor task, eight warnings appeared at random points of time during the scenario. The participants had to react as fast as possible to those warnings by pressing the brake pedal.

Overall, 45 people participated in the study. The majority of participants (80%) was between 20 and 39 years old, 11% of the participants were younger than 20 years, and the remaining 9% were older than 39 years. Of the participants, 26 were female and 19 were male.

Before the training scenario and start of the study, participants completed a demographic questionnaire and, thereunder, rated their level of trust in technology in five questions. The same three questions about trust from the previously presented study were utilised again (Section 4.1, Table 10). Additionally, two questions of trust related to the self-driving car scenario of this study were utilised (Table 12). It was expected that the trust would vary from overall perspective to single components. For example, the warning system of an automated or self-driving vehicle was rated higher in trust compared to the automated or self-driving vehicle such (Table 12, 3.6 and 3.7).

Table 12. Trust questionnaire.

3.6) How would you rate your level of trust in an automated or self-driving vehicle?

Very distrustful									Very trustful
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

3.7) How would you rate your level of trust in the warning system of the highly automated vehicle?

Very distrustful									Very trustful
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

Trust in each of the three warnings was rated on a seven point rating scale (Table 13). This questionnaire was completed before and after the training scenario to investigate whether trust in a warning would change after experience of the warning.

Table 13. Trust – experience questionnaire for the warnings.

1) How would you rate your trust in an auditory warning

Not very much						Very much
□	□	□	□	□	□	□

2) How would you rate your trust in a tactile (vibrating) warning?

Not very much						Very much
□	□	□	□	□	□	□

3) How would you rate your trust in an auditory-tactile warning?

Not very much						Very much
□	□	□	□	□	□	□

The data was analysed in R (R Core Team, 2014). Pre-requisite to participate in this study was normal or corrected to normal vision, normal or corrected to normal hearing, and no issues known that could influence tactile perception.

4.2.2.   Results

The majority of participants had experience with a haptic interface in general. Only one person had no experience with haptic interfaces in general. Haptic interfaces in cars are still not common. The majority of the participants had no experience with a haptic interface in a car (71.1%).
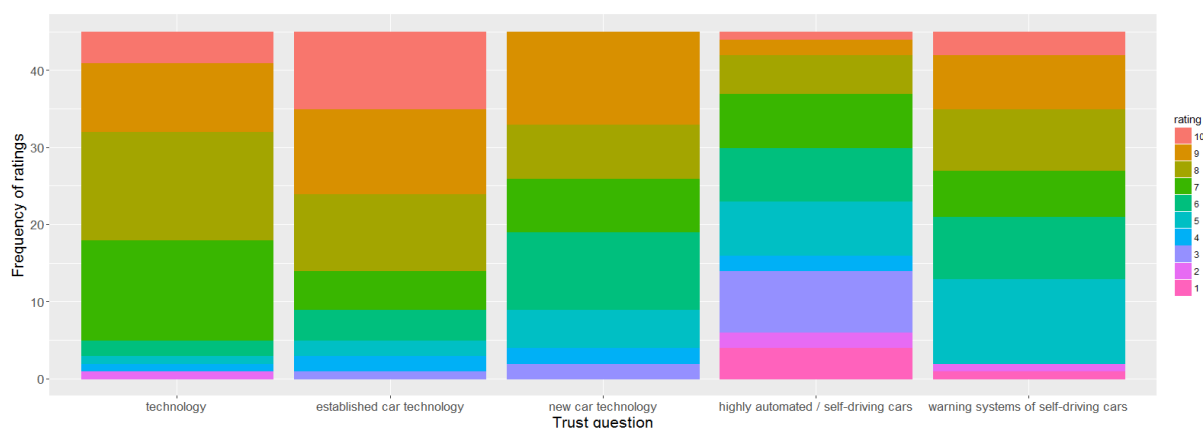

Figure 9. Distribution of ratings in trust. Colours indicate levels of trust, ranging from low (1) to high (10).

Figure 9 shows the ratings for each of the trust questions. In general, participants rated their trust in technology high ($M = 7.7$). The participants had high trust in established

car technology (M = 7.9) and new car technology (M = 6.9). Lowest trust ratings were given to trust in highly automated or self-driving vehicles (M = 5.2), but they were higher for a warning system of such a vehicle (M = 6.7).

Before and after experience trust rating

The participants were introduced to the study setting in the beginning of the study. Then, in three parts of the training scenario, the participants adjusted themselves to the simulator, practiced reacting to the warnings only, and practiced conducting the task and reacting to the warnings in parallel. The training scenario started with a driving only phase, which lasted about one minute. Thereafter, each of the three warnings appeared at a random point of time (each warning four times) and the participants had to react as fast as possible to the warning by pressing the brake pedal. The last part of the scenario was training in the distractor tasks at the end of which the participants had to react to the warning while doing the distractor task. Before and after the training scenario, the participants were asked to rate their level of trust in each of the three warnings (Table 13).

The trust ratings increased with experience, as the participants became more familiar with each warning and experienced them in all distracting conditions (Figure 10). Before the training scenario the trust was rated highest in the MMW (M = 5.8, SD=0.9), middle in the auditory warning (M = 5.7, SD=0.9), and lowest in tactile warning (M = 5.2, SD=1.2). After experiencing all warning types in the training scenario the average trust increased slightly. Trust was rated highest in the MMW (M = 6.3, SD=1.1), middle in the auditory warning (M = 5.8, SD=1.4), and lowest in the tactile warning (M = 5.4, SD=1.3). A pairwise comparison with three independent t-tests revealed that only the ratings of the MMW increased significantly between pre and post experience (t(84.9) = -2.2, p=0.03).
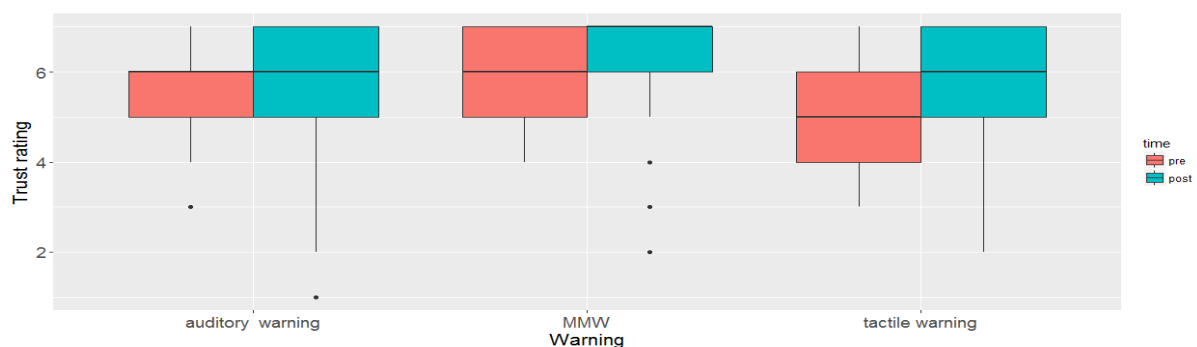

Figure 10. Trust in each of the three warnings, before (red) and after (blue) experiencing the warning.

Correlation trust and RT

The ratings in trust (Table 10, Table 12) were compared in a pairwise Pearson Product correlation analysis to the RTs to the tactile and MMW. The trust ratings and RT to the MMW were not correlated (r <0.3). The trust ratings and the number of false reactions to the MMW were not correlated (r <0.3).

4.3.   A link between trust and glances in three navigation systems

The third study is a follow-up study to the Trust study (Section 4.1). Previous research found that drivers could be divided into three groups dependent on how they distribute their glances between on-road and an in-vehicle device (Broström et al., 2015). Some drivers apply more frequent glances, but at a short duration. Other drivers apply less frequent glances, but with more glances of longer duration. Other drivers apply a balanced mix of glances, not so often and short in duration. In this research approach, it was evaluated if that glance pattern reappears over interfaces with varying visual demand and, further, if these glance patterns are linked to how much trust a driver has in a new in-vehicle device.

4.3.1.   Method and procedure

The study evaluated the relationship between pre-exposure trust and glance behaviour in three interfaces with varying levels of visual demand. Data of 49 participants was considered for the analysis. All participants experienced three navigation systems. The study was conducted in WMG's 3xD driving simulator.

Based on the findings of the Trust study (Section 4.1), a potential link between pre-exposure trust and glance behaviour, this study evaluated the relationship between trust and glance behaviour further. The same three questions related to trust were used (Table 10). In contrast to the Trust study the relationship was evaluated in three interfaces with varying levels of visual demand. One interface was voice-only, not providing any visual output. Another interface was voice-visual, a voice command system with visual feedback to guide the interaction (Figure 11). The third interface employed a visual-manual interaction and included more detailed visual information (Figure 11).
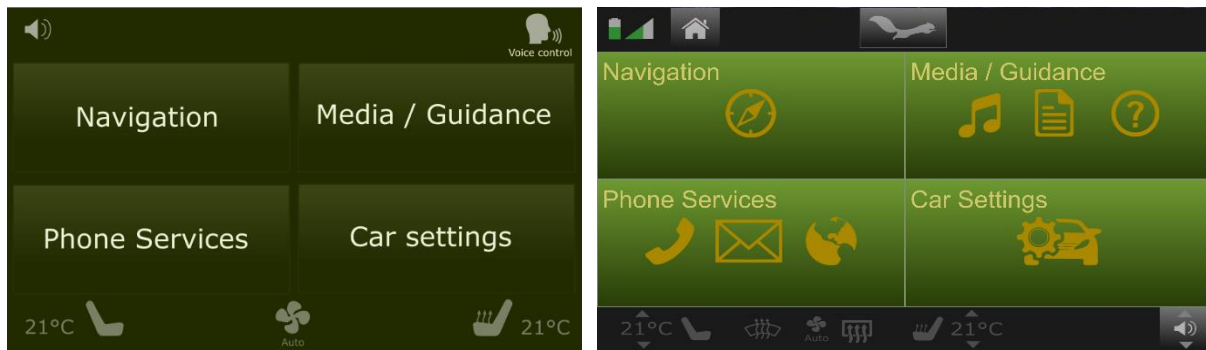
Figure 11. Voice-visual interface start-screen (left); visual-manual interface start-screen (right).

Before the start of the study, the participants were asked to complete a set of questionnaires. Those questionnaires confirmed that the participants met the criteria to participate in the study, captured demographic information, and captured the pre-experiment level of trust in technology. The participants were required to have a valid driving license in UK and to drive regularly, at least once a week. At the beginning of the study, the eye tracking system was calibrated to the participants' facial features and gaze behaviour. Therefore, the participants held a chessboard up, and looked at pre-defined locations (the middle of the speedometer and the middle of the tachometer) within the simulator car's dashboard. While the participants looked at each of the specified locations, the system learned to detect the participants' gaze. During the study, the participants drove in manual mode along a straight motorway road. While driving, the participants needed to enter addresses into the navigation system (similar to the study setting in Mehler et al. 2016). The navigation system was presented on a tablet computer that was attached to the centre console of the simulator car and gave the impression of a built-in device. One scenario consisted of three repetitions of the task with three different addresses. After each scenario the participants completed a questionnaire related to trust, usability (in the System Usability Scale (SUS; Brooke, 1996)), and user experience (User Experience Questionnaire (UEQ), adapted from Van der Laan, Heino, and De Waard (1996)). At the end of the study the participants selected their favourite interface.

The data was analysed in R (R Core Team, 2014). A subset of participants was selected for the analysis of the relationship between trust and glance behaviour to the navigation system based on the quality of the data available for each of the three navigation systems. In each section the number of participants considered is stated. The recording included missing data points, for example, when the participant blinked

no gaze data was recorded. Data with up to 15% missing data points was considered for the analysis. Durations up to 500 ms with no Area of Interest (AOI) assigned were classified as blinks and interpolated. This is a duration which was found in literature to be considered as blink (Metha and Shrivastava, 2012; Wang et al., 2011). Longer durations were not interpolated and classified as missing data.

Only the address entry with the best quality was selected, sets with more than 15% missing data points and bad data (periods longer than 500 ms of missing data) were excluded. For each interface the following glance metrics were analysed: Total task time, mean single glance duration, total time glanced to navigation display, percentage of glances to the device >2 s to navigation display, average duration of glances to the navigation system and average number of glances to the navigation display per second.

4.3.2. Results

The study showed, similar to previous literature (Reimer et al., 2014; Sawyer et al., 2017), that participants apply glances while interacting with voice-command interfaces. In this study, there were even glances applied when no visual information was shown at the interface. Jensen et al. (2010) evaluated glance behaviour interacting with three different navigation systems while driving in the real world. The three navigation systems were similar to those in this study: one audio, one audio-visual, and one visual. Similar to their finding, the number of glances to the navigation display (and number of people who glanced to the navigation display) increased from voice-only (53), voice-visual (124), to visual-manual (443) navigation display. Positively, the participants applied no glances >2 s to the navigation display in either of the voice-command interactions.

A number of participants glanced to the navigation display for all three of the navigation systems. However, only very few glances to the navigation system and of a very short duration were observed in the voice-only and in the voice-visual condition. Those glances were considered as insufficient to be analysed in detail for the glance patterns described in Broström et al. (2015).

In the interaction with the visual-manual navigation display the group described as optimisers by Broström et al. (2015) was found: participants who applied short duration glances with less than 15% of glances off-road >2 s. There was a general

tendency to apply more frequent but shorter duration glances, or glances of higher duration but less frequent.

Over all three interfaces there was a positive relationship between trust and total task time, participants with higher trust in the interface took longer to complete the task. For the voice-only navigation system this was the only observed pattern.

While interacting with the voice-visual navigation system participants with a higher rating of trust tended to apply shorter glances, but glanced more frequently at the navigation display. Participants with lower ratings of trust tended to apply longer glances, but less frequent at the navigation display. This observed relationship is similar to the previous correlation reported (submission 3). However, the difference in glance patterns between high and low trust was not significant in this study. Sawyer et al. (2017) also reported longer duration glances off-road for participants with a low level of trust. They also reported that participants with a low level of trust glanced more frequently off-road, which did not appear to be the case in this study, however differences for this glance metric in this study were not significantly different.

While interacting with the visual-manual navigation system, participants with high trust tended to spend more time glancing at the navigation display and tended to spend a higher number of glances >2 s at the navigation display, compared to participants with lower trust. However, the difference was not significant.

5. HF Toolkit project

New in-vehicle interfaces are developed in a prescribed design process. Evaluation of the HMI with users is part of the design process. Measures associated with those studies are called HF related measures as they provide an understanding of the human machine interaction. HF related measures provide information about various aspects of the interaction: ergonomic design of the interface, usability of the interface, or the subjective experience with the interface. Each of the aspects is associated with another set of measures. Some measures, such as for driver distraction, are suggested by guidelines (NHTSA, 2013). There exist several measures for driver distraction which differ in required effort of application, skills and measurement equipment. For a company it is important to employ an efficient way to test their products most suitable for the project time-schedule and available equipment. The herein developed HF toolkit aims to present HF related measures in a way that eases retrieval of a specific measure from a great number of measures and that eases the comparing them to each other. By this, the HF toolkit could support HMI engineers in the understanding, comparison, and selection of the most suitable measure for their user trial. The following example illustrates the usefulness of the HF toolkit:

"Kevin and Linda are planning a user trial to evaluate the distraction of a new voice-based in-vehicle navigation system. Both are familiar with user trials and experts for workload measurement, but they do not have experience with driver distraction measures yet. They start gathering information from different sources and arrange them in a table. But with many measures in the growing table, Kevin and Linda are not sure which of the measures is the best to use."

Kevin and Linda could have saved time searching for papers and consolidating information from various research papers, if there had been a collection of measures about driver distraction. The example from Kevin and Linda points to two requirements for the HF toolkit. A simple collection of measures in a database already helps. However, a good structure that characterises the measures in the database would decrease the time to find a measure for a specific study. Keywords to search for measure characteristics help to identify the most suitable measure for a purpose fitting to the project circumstances, e.g. limited equipment availability, or a time critical project. The second aspect in the example points to the next step of Kevin and Linda's task, that is to consolidate the information and compare the measures among each

other to identify the most suitable for their study. Aiding this step the HF Toolkit should present the information in a format that eases comparison. Tables are a typical way to compare information. Consisting of text, instead of visuals, tables can become cumbersome to compare a large amount of information. Concluding from Kevin and Linda's example, it was aimed for an interface design that presents a large amount of information easy to retrieve and that aids the comparison of that information in a visual format.

It appears there exists no such electronic aid for selection and comparison of HF related measures in the published literature of automobile manufacturers. There are websites with collections of measures for usability or user experience (Table 14). However, those collections are for a specific set of HF measures, often including methods for design as well, and they concern web-design rather than in-vehicle interface design. The websites are designed in common web layouts, at times emphasising visuals and easing information retrieval with icons that represent the sections of the website.

Table 14. Existing collections of HF related measures – web-based HF toolkits.

Source	Topic	Presentation
('Service Design Tools', 2016)	Service design tools – covering usability, user experience, and creativity in the design process	Web-design layout with an icon representing a tool Tools are ordered in four categories (e.g. design activities and recipients) and their sub-categories A click on a sub-category shows the tools therein, each represented by an icon and its name A click on a tool then opens a description
('All About UX', 2016)	Measures related to user experience	The measures are organised in categories which are presented with sub-categories in a list, each list item is a hyperlink A click on a sub-category hyperlink opens a list of measures each presented with a sentence on description A click on one of the measures then opens a separate window with the detailed information about the measure: description, strength and weakness
('Usability Toolkit', 2016)	Usability measures, design methods	The website is organised in five areas which are presented as "tabs" as horizontal menu on the top of the website A click on one of the menu point opens it and shows, e.g. for methods to reach a usable design – each described by a sentence. A click on a method then opens its description, including: preparation, execution, and analysis
(Usability.gov, 2016)	User experience, usability, design methods, project management	The website aims to be visual, its content areas are presented by icons A click on the dedicated icon opens the area The measure description includes implementation, analysis and writing of a report – including templates for, e.g., usability test plans and reports

Aim of the HF toolkit project was to develop a novel interface that would support HMI engineers in the task to understand, compare, select and utilise HF measures for the evaluation of in-vehicle devices. The HF toolkit interface was planned to be developed in a process of co-design with future users (HMI engineers), to ensure their needs are

met. At the projects end the final interface concept was transferred into an interaction flow-chart and handed over to the sponsoring company for the development of a computer-based application. An interaction flow-chart describes how users can interact with an interface. In this project, it was based on the interactions from the paper prototype studies.

Summarising, this Chapter describes the research questions from research stream 2 (Section 1.1). The next sections describe the design process of the HF toolkit, the underlying methods of the conceptual design, the evaluation of the conceptual design and the final conceptual interface.

These studies (REGO-2015-1719 and REGO-2016-1795) were ethical approved by the University of Warwick's BSREC.

5.1. Method and procedure

The toolkit was developed with the Human-Centred Design (HCD) process to ensure that the interface reflects the future users' needs, the needs of the sponsoring company's HMI team members. The HCD process describes how future users should be involved in each design stage in the product development process (Maguire and Bevan, 2002). This involvement guides the designer in the decision-making process to select a design option that serves the users' needs best. Otherwise, the designer would create an interface with best intention, but unknown benefit for the users. The interface might not support certain tasks or users find it difficult to interact with it. Consequently, each design stage involved feedback from future users. Figure 12 shows the HCD stages: Planning, Context of Use and Requirements, Conceptual Design, and Design Evaluation.
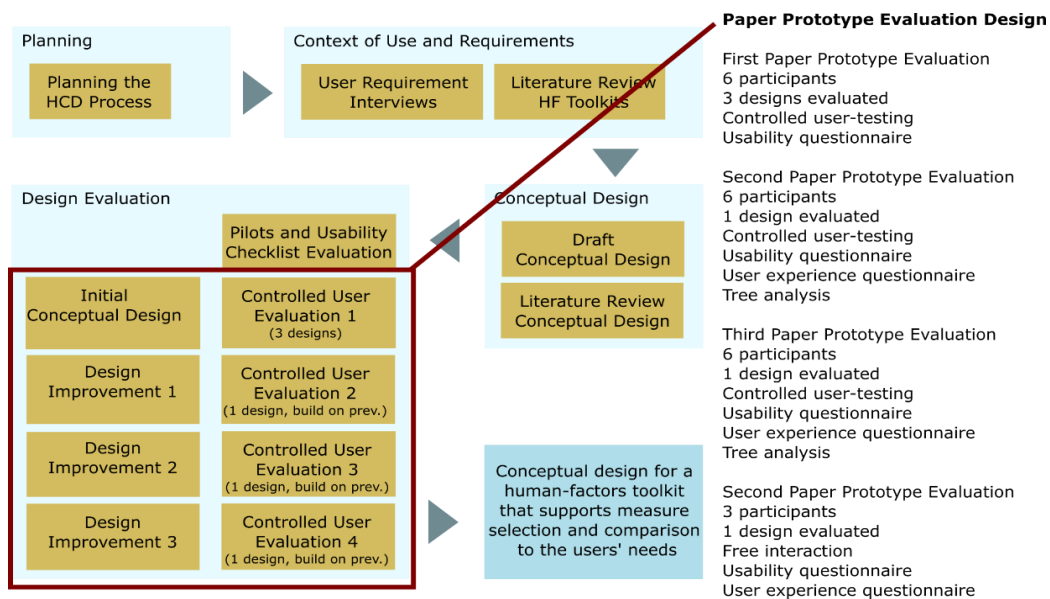
Figure 12. HCD process design stages and selected methods for each of the design stages.

5.2. Planning

The HCD process started with planning (Figure 12, top left). That meant to decide about the methods in each of the design stages. The future users are members of the sponsoring company's HMI team. To participate in the toolkit design process, they needed to take time from their industry project. Time is a critical factor in industry. To facilitate participation and to gain the most from the time each participant contributes, cost-efficient methods were selected to reach a usable design. Nielsen (2009) proposed usability methods that can be utilised with a minimal investment of resources and participants: expert evaluations, user trials, and paper prototypes. Those methods were applied in this project.

Expert evaluations before a user trial help to identify major usability flaws. Experts are easier available in a company for a participation than users. Through preliminary elimination of major usability issues, user trials can become more efficient with comparably little investment. However, because experts have a different mental model of the interface and a different interaction with devices, perhaps more technical than a user, they do not replace users. Consequently, it was planned to let the interface be first evaluated by three experts with a usability checklist, and then to perform an iterative evaluation with users. The user evaluation was planned with four paper prototype iterations with evaluation and redesign (Figure 12, bottom left).

The second stage of the HCD process included additional methods to obtain an understanding of the context of use and to obtain user requirements as basic principles for the design (Figure 12, top conter). This involves an understanding of the situations in which the interface will be used, the users' expectations towards the interface and the most important functions for the interface that support the users' tasks. There are two potential user groups for the toolkit: HMI engineers and managers. An HMI engineer uses the toolkit to compare and select HF measures for a user trial. A manager might be interested in information to communicate the best practices and establish a certain way to measure, e.g., driver distraction. For the gathering of the user requirements users from both groups were interviewed. The interviews were semi-structured with slightly different keywords dependent on the role of the interviewee.

Interviews with HMI engineers comprised the following topics:

- Understanding of the process of comparison and selection of HF related measures to evaluate in-vehicle devices in user trials
- Understanding of implicit knowledge involved in the measure selection process
- Understanding what tools HMI engineers currently use and expectations towards a newly developed tool

Interviews with HMI team managers comprised the following topics:

- Expectations, benefits, and concerns towards an electronic aid for planning of user trials
- Expected concept of use

For the initially conceptual design (Figure 12, centre) a literature review was conducted to identify existing interfaces and strategies to visualise information. The initial conceptual design was planned to improve iteratively in four paper prototyping rounds (Figure 12, bottom left). In each paper prototyping round, the HMI engineers conducted a set of typical tasks, selected from the previously introduced use-cases. The paper prototyping evaluations consisted of objective and subjective measures. Objectively, the interaction between HMI engineers and interface was analysed, comparing the HMI engineers' inputs while completing a task to a path of shortest interaction. Subjectively, the HMI engineers rated the usability (in the System Usability Scale (SUS) (Brooke, 1996)) and user experience (in the User Experience Questionnaire (UEQ) adapted from Van der Laan et al. (1996)) of the interface after

completion of the tasks. After each paper prototyping round the collected data was analysed to improve the usability and user experience of the design. Each improvement was implemented under consideration of several sources of user input: interaction, comments, and subjective rating of the interface.

For the first paper prototyping round it was planned to evaluate three interface designs to compare two interfaces with visual metaphors to a tabular interface. In the other three evaluations, only the favourite of those three interfaces was iteratively improved. Overall, the four paper prototype evaluations focused on:

- Understandability of the visualisation concept
- Efficient navigation through the tool, obtaining information about different measures
- Provision of useful information about a measure in overview and as details
- Support of typical tasks derived from user requirements (Section 5.3)

5.3. Context of use and user requirements

Overall, three HMI engineers and two managers of the HMI team of the sponsoring company were interviewed to gather user requirements. HMI engineers and managers both found the idea of a toolkit useful. HMI engineers saw value in a quick way to compare measures and to be able to find descriptions of the best practices. The managers specifically found that a toolkit could help to consistently apply a measure and thereby improve user trials. The interviews were analysed with coding. Outcome of the interview analyses was a set of use-cases that was subsequently implemented into the conceptual design:

Use-Case 1:   The user explores measures in the toolkit content.

Use-Case 2:   The user quickly obtains general information about a measure.

Use-Case 3:   The user quickly compares measures.

Use-Case 4:   The user can practically apply a measure (implementation and analysis) based on information provided in the toolkit.

Use-Case 5:   The user adds a new measure in the toolkit.

Use-Case 6:   A new employee finds support in the measurement selection process for a user trial. – This means the integration of existing guidelines for new employees in the toolkit.

5.4.   Conceptual design

The development of the conceptual design focused on the creation of an interface easing the comparison of multifaceted data. A visualisation was a design option early on, as visuals are faster retrievable compared to text and the visual sense transfers the most information. A design concept with a visual interface and fast retrieval of large amounts of data is Visual Information Seeking (VIS) (Shneiderman, 1996). Previously it had successfully been applied to organise large amounts of information in, for example, patient records (Plaisant et al., 1996) and a movie database (Ahlberg and Shneiderman, 1994). VIS has seen limited application since its development over 20 years ago. However, in the current era of Big Data where Information Visualisation and Infographics are essential methods for conveying large amounts of information to users, the principles of VIS could be considered to be more important than ever. It is for this reason, especially given the proven effectiveness to avoid information overload and structured principles to visualise information, that the method was adopted for this research.

A central concept of VIS is the mantra: "overview first, zoom, and filter, then details on demand". The VIS concept is visualised for HF measures in Figure 13. Users can see all "items" of a database on the start screen of a VIS interface. For a database of HF related measures that means all HF measures are presented on the start screen. The measures are only shown by their characteristics. By sorting the measures along their basic characteristics, the users learn about the content of the toolkit and can decide to reduce the presented set of HF measures to a set most interesting for them. In this reduced set, they can obtain more information about a measure on demand. This process eases information retrieval and avoids information overload, compared to a table that presents the measures in all their detail.
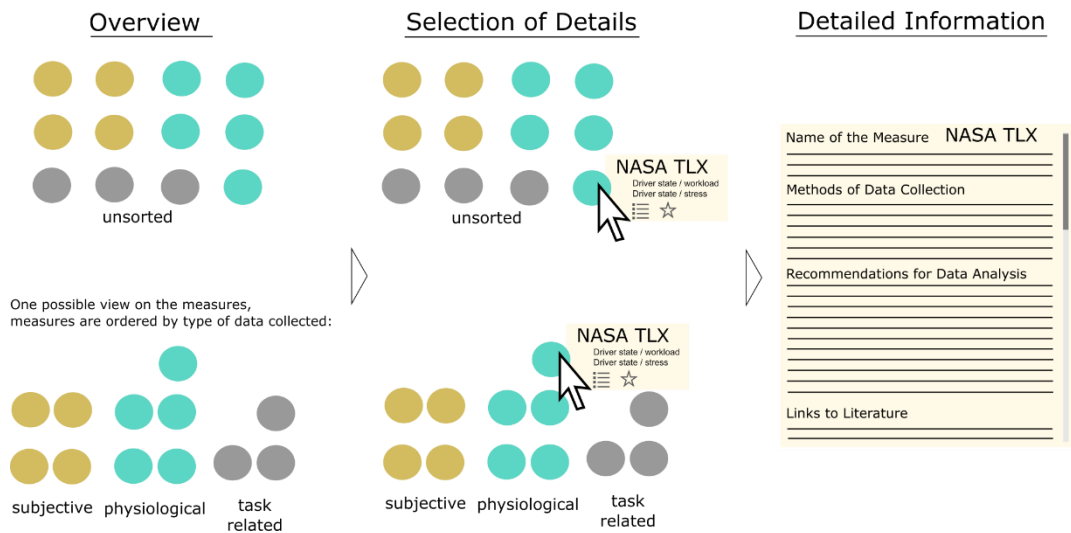
Figure 13. The VIS concept applied to the comparison and selection of HF measures, each dot is a HF measure.

The VIS literature does not provide advice for the creation of the characteristics by which the measures are presented in the overview. The characteristics are organised and presented in categories. The process to create those categories is comparable to the organisation of information in websites, typically faceted classifications are used (Broughton, 2005). A faceted classification describes data from multiple points of view and can represent the manifold aspects of a measure, such as the collected type of data, equipment needed, reliability, and practicability. Table 15 describes the initial category names for characteristics of measures in the HF toolkit. The same categories were used in all three initial conceptual designs.

Basis for the initial conceptual design were the use-cases derived from interviews with future users of the HF toolkit (Section 5.3). Those use-cases were implemented in interfaces. After an expert review three interfaces were selected for the paper prototyping iterations with HMI engineers.

Table 15. Initial category names for the HF toolkit in Paper Prototype 1.

Category name	Explanation of the category
Experimental Target	Psychological constructs for which a measure can be interpreted, e.g. usability, user experience, driver distraction or workload.
Type of Data Collected	The type of data that is collected with the measure, e.g. subjective, task related or physiological data.
Product Design Phase	The design phase in which a measure can be applied. Some measures are more suitable to be used in the early stage of a product, others require a certain prototype, e.g. measures for glance duration or driver distraction.
Period Measured	Describes at which point in the study a measure can be applied and if it measures a point in time or collects data over a period of time.
Location of Use	The experimental locations in which a measure can be applied, e.g. driving simulator, test track, or on-road.
Practicability and Quality	Whether a measure easy to use and easy to administer or it defines the quality of the measure.
Disadvantages	Disadvantages a measure might have, e.g. if it could interfere with the task that the participant is asked to conduct.

One initial interface adhered to principles of VIS and employed a diagram as well-known metaphor for visualisation (Figure 14). Each tab next to the diagram represents a category (Table 15). For an overview of the measures and an initial comparison, users can sort the measures along categories. When a tab is selected, each tab presents the user with a diagram. The measures in the diagram are sorted in subcategories along the diagram's axes, exemplary shown for practicability and quality, whereby a measure is represented by a dot. The user can select a filter from the filter menu on the left side to reduce the shown set of measures.
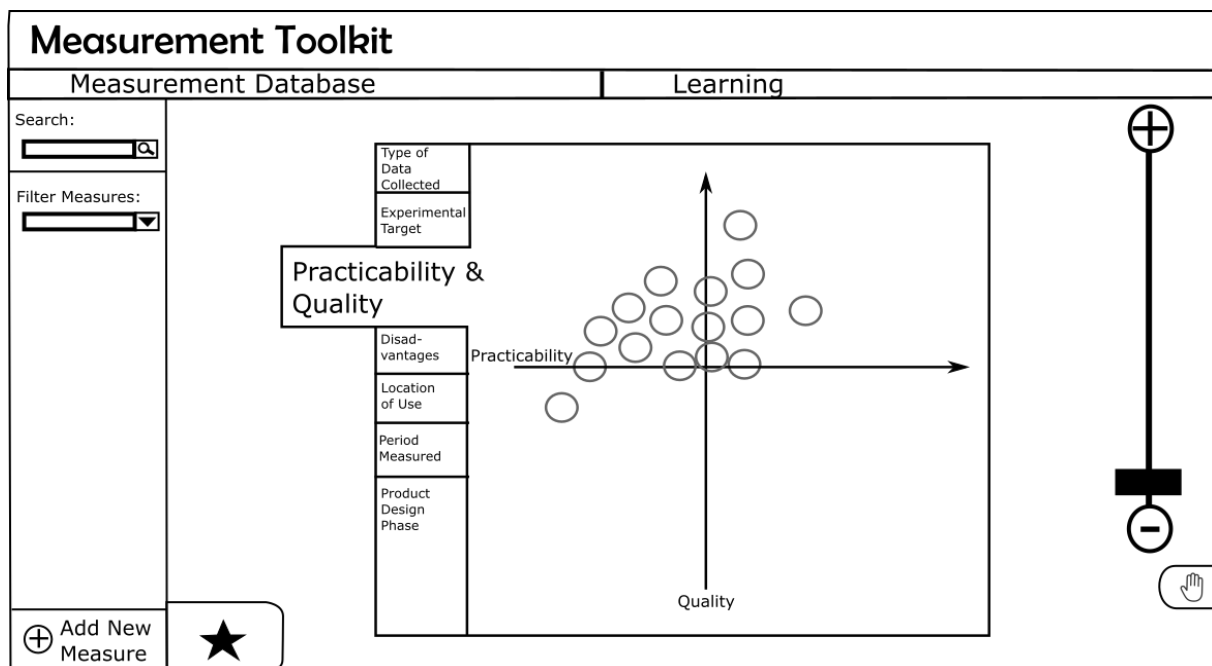
Figure 14. Initial Diagram interface.

Figure 15 shows the two other initial conceptual designs, the Bubble (top) and the Spreadsheet (bottom) concept. The Bubble interface employs VIS principles, but uses a more abstract visualisation with a circular menu. The categories are presented in the outer circle. A click on one of the categories shows its subcategories in the inner circles. When a sub-category is selected connecting lines between this category and the measures show to which measures it applies. Measures can be filtered on the left side, and on the bottom is an area where measures can be dragged and dropped for comparison. The Spreadsheet interfaces lists the measures alphabetically at a list (left, top), measures can be filtered (left, bottom), and measures can be selected for a comparison in tabular format (right).
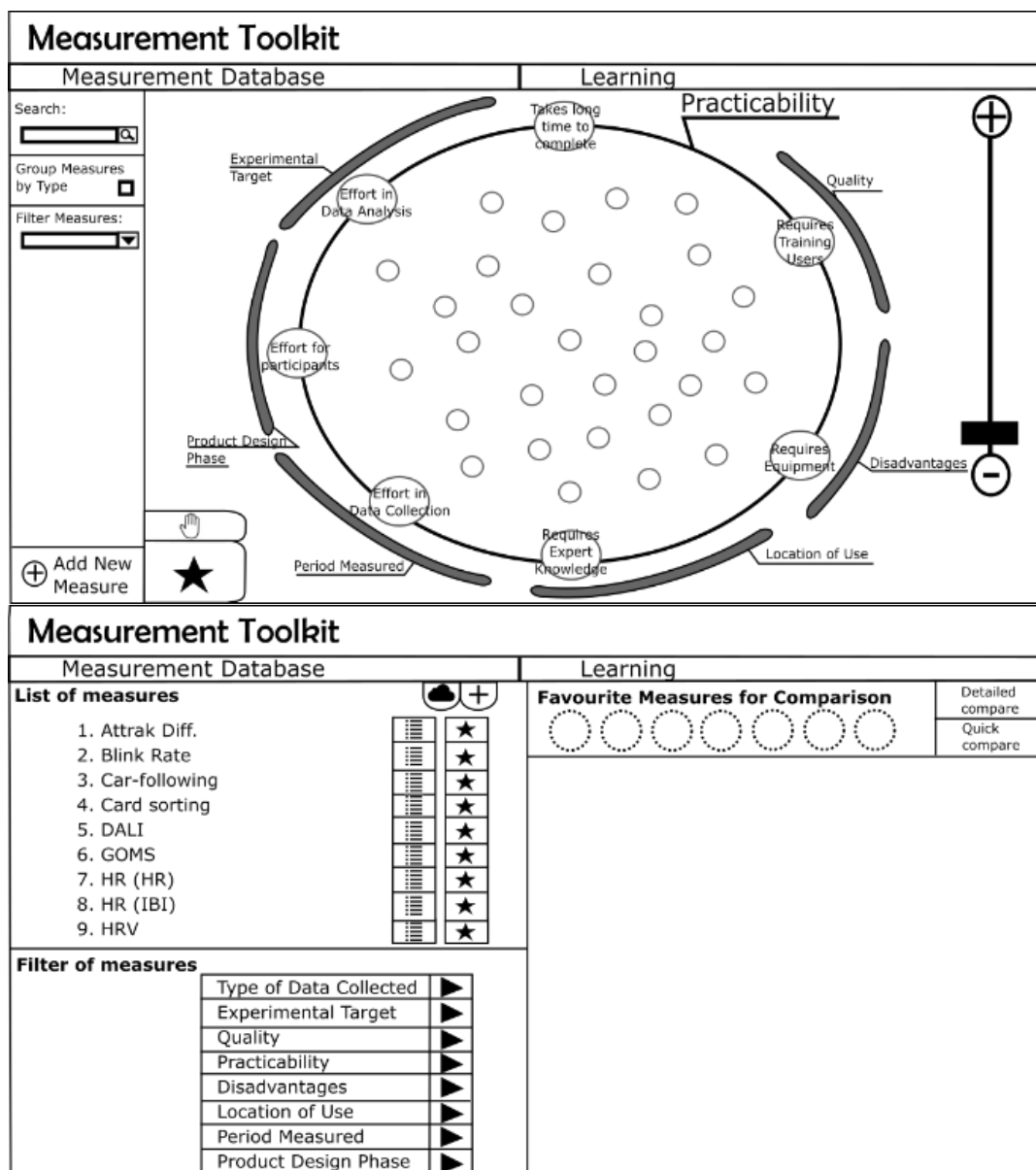
Figure 15. Bubble interface (top) and Spreadsheet interface (bottom).

5.5. Design evaluations

Each of the four paper prototype evaluations had another focus (Table 16). The first and the last paper prototype evaluation covered all use-case. Paper prototype evaluations two and three focused on the improvement of the information structure.

The first paper prototype evaluation compared three interfaces (Figure 14 and Figure 15). The results indicated that the participants preferred a known element for the visualisation and a good structured interface. The Bubble interface was popular with half of the participants, but over half of participants also claimed that it contains too much information. The difficulty reflected in the interaction. A higher number of

participants needed help to complete the tasks in the Bubble interface compared to the Diagram interface. The Bubble interface received the lowest usability rating (SUS score: 53). The Spreadsheet interface received a SUS score of 65. The Diagram interface was rated most positive of all three interfaces (SUS score: 76). Compared to the other interfaces, the least number of participants needed help interacting with the Diagram concept. Perhaps the intuitiveness of the interface increased by the measures being ordered in the known metaphor of a diagram, one participant mentioned that explicitly. The participants switched easily between overview and detailed information. A difficulty was the association between the categories for overview and underlying information. For example, participants did not associate the category name "study target" with measures grouped into those for usability, user experience, driver performance (person), ... This issue was addressed in the next paper prototype evaluations.

Table 16. Paper prototype evaluations.

Paper prototype evaluation	Aim	Participants
Paper prototype evaluation 1	Evaluation of three different interface designs. The first paper prototype study covered all use-cases.	6 participants (HMI engineers)
Paper prototype evaluation 2	Improvement of the preferred interface from iteration 1. The second paper prototype evaluation focused on use-cases one, three, and four.	6 participants (HMI engineers)
Paper prototype evaluation 3	Improvement of the interface from iteration 2. Aim was to evaluate filtering and sorting of information in different categories. The third paper prototype evaluation focused on use-cases two, three, and four.	6 participants (HMI engineers)
Paper prototype evaluation 4	Improvement of the interface from iteration 3. The fourth paper prototype study considered all use-cases.	3 participants (HMI engineers)

Paper prototype evaluations two and three focused on the improvement of the categories for overview. The categories were evaluated to be clearly distinguishable from each other and to have meaningful names that represent the underlying information. For example, the category "Disadvantages" was deleted as it was not clearly distinguishable from the other categories and the category "Product Design Phase" was renamed to "Design Phase" to represent its content better. In all paper prototype evaluations the participants switched easily between overview and detailed information, which led to the conclusion that the VIS principles can be used to present information about HF measures adequately.

The usability and user experience ratings varied over the paper prototype evaluations whereby the Diagram concept received the highest usability rating in the first evaluation and was then successively improved (Table 17). The final SUS score of 68

does indicate that the interface is not immediately intuitive but its usability is between ok and good (Bangor et al., 2009). The final conceptual design achieved the goal of a good user experience with a clearly positive rating UEQ of 0.9 (on a scale of -2 to 2). The usability of the interface can be further improved in the implemented version.

Table 17. SUS and UEQ ratings for the paper prototype evaluations.

Diagram SUS and UEQ scores		Paper prototype one	Paper prototype two	Paper prototype three	Paper prototype four
Number of participants		6	5	5	3
SUS Score	Average	76	55	70	68
	Range	(60-88)	(33-83)	(48-88)	(38-85)
Number of participants		6	5	6	3
UEQ Score	Average	-	0	1	0.9
	Range	-	(-0.7-0.9)	(0.5-1.9)	(0-1.4)

5.6. Final conceptual interface

The final conceptual design consisted of four main functional parts (Figure 16): overview, filters, detailed information, and measure comparison. In the overview section HMI engineers obtain an understanding of the database content and learn about the measures. Based on that understanding the HMI engineers can then apply filters to reduce the shown set of measures. In the reduced set of measures, HMI engineers can then obtain more information about a selected measure. In a dedicated section of the interface, HMI engineers can compare measures. There is a quick comparison and detailed comparison available. The following paragraphs will explain each functional part in more detail.
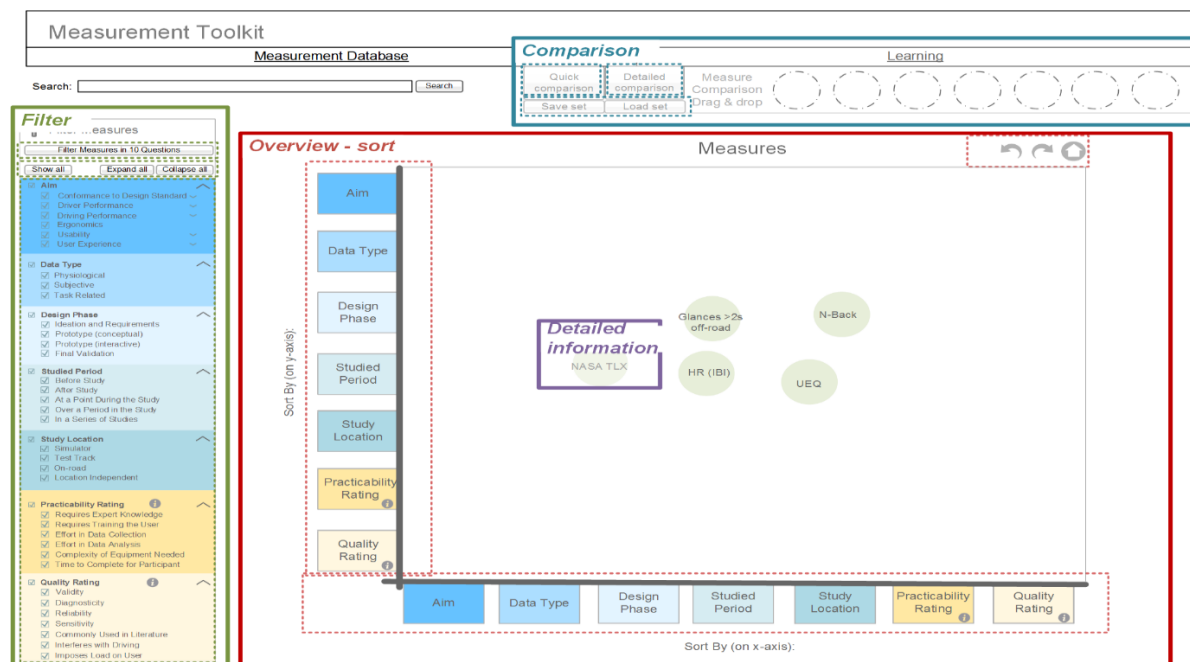

Figure 16. Final interface design of the HF toolkit.
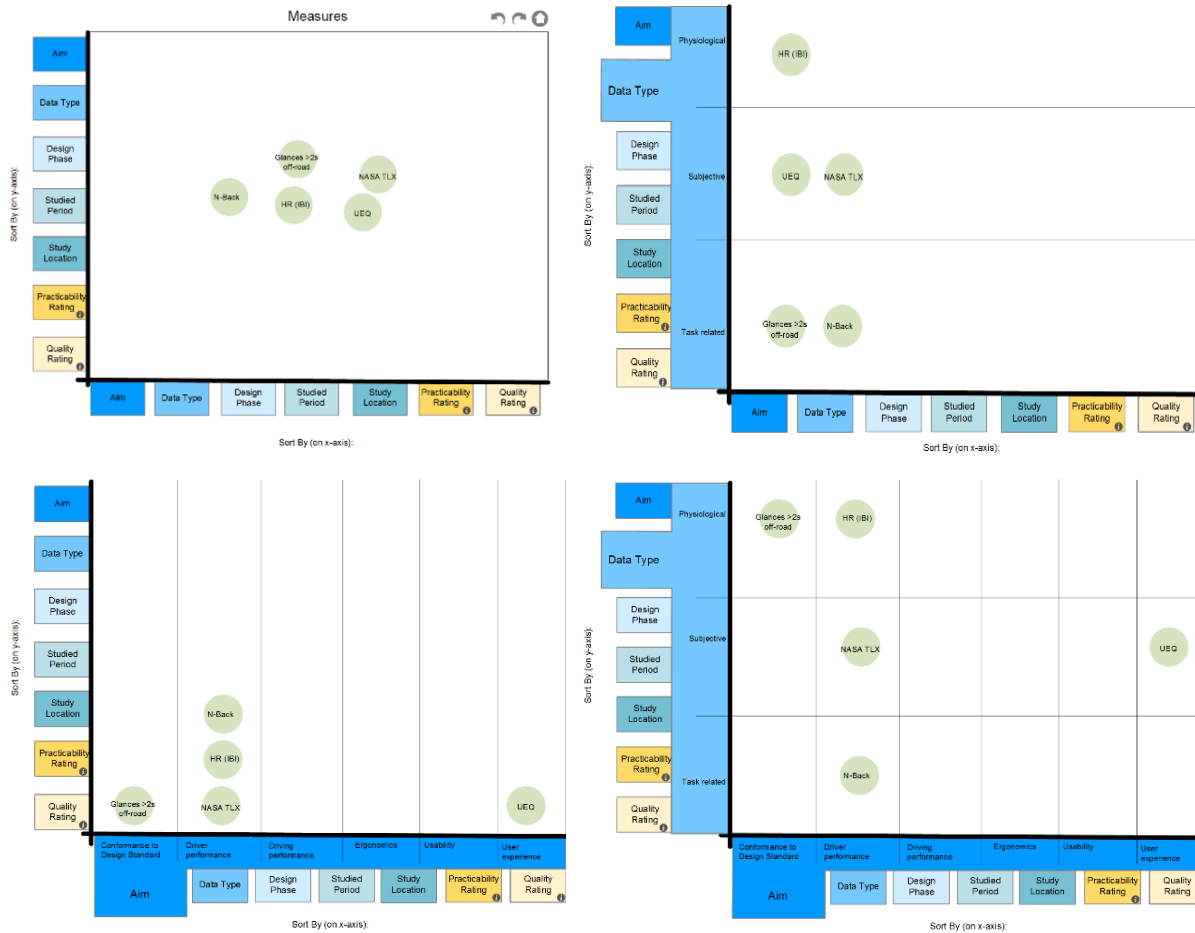
Figure 17. Overview information in the scatterplot.

- *Overview sort:* The measures are presented as dots in a scatterplot (Figure 17). With tabs next to the axes it is possible to sort the shown measures. A measure's spatial position in the scatterplot derives from its characteristics. Each tab presents a different category and reveals another aspect about the measure. Thus, the HMI engineers can retrieve information dependent on the circumstances of their study. For example if the study is time critical, HMI engineers might be interested in measures that are fast to apply and analyse. Therefore the HMI engineer selects the category "Practicability", and select only the sub-category about effort a measure requires. The diagram will then show the measures sorted by the required effort. In another case, when there is the requirement to conduct the study in the simulator, they could select the category "Study location" which sorts measures dependent on the location in which they can be used.

- *Filters:* On the left side of the HF toolkit interface are filers to reduce the shown amount of information. The filters consisted of the same categories, matched in colour, as the overview to facilitate the learnability of the interface.

- *Comparison:* The interface presents options for a detailed or a quick comparison of measures. To conduct a comparison the measures are dragged and dropped into the comparison field. A detailed comparison opens a new window and shows the details of the selected measures in a table. A quick comparison opens a new window and shows (only) the selected measures in the diagram.

- *Detailed information:* Detailed information about a measure is available on demand by selecting the dedicated option in the measure's context menu, which appears on click on the measure or when the mouse pointer hovers over the measure. The detailed information consists of information from all overview categories and a practical information of the measure, for example how to implement the measure, how to adjust equipment and how to analyse data obtained from the measure.

The final conceptual design supports the use-cases in the following ways:

Use-Case 1: The user explores measures in the toolkit content: The visual presentation of information supports exploration. HMI engineers see, initially, all measures from the toolkit, which could raise an interest in exploring a yet unknown measure.

Use-Case 2: The user quickly obtains general information about a measure: This use-case is implemented as view options by which HMI engineers can acquaint themselves with characteristics of a measure (red area "Overview – sort" in Figure 16). The set of measures can be reduced rapidly to the most interesting by filters presented next to the diagram (green area "Filter" in Figure 16).

Use-Case 3: The user quickly compares measures: This use-case is implemented as overview information, by the views and by the option to compare a small set of measures in the comparison box (blue area "Comparison" in Figure 16).

Use-Case 4: The user can practically apply a measure (implementation and analysis) based on information provided in the toolkit: The detailed

information about a measure was selected to guide HMI engineers in the implementation and analysis of the measure (purple area "Detailed information" in Figure 16). Specifically, it is aimed to gather best practices of application in the description.

Use-Case 5:   The user adds a new measure to the toolkit: New measures can be added in the section about detailed information of any measure in the toolkit.

Use-Case 6:   A new employee finds support in the measurement selection process for a user trial: Existing guidelines can be implemented in the toolkit.

The interface as presented in this section was described in an interaction flow-chart in MS Visio and handed over to the sponsoring company. The implementation, as it does not involve conceptual novelty, was not part of the project. Later, an initial software prototype of the interface was developed by a summer intern at WMG. The prototype is implemented with Python and the Django web framework. It runs platform independently in a web-browser. The implementation comprises the layout, and basic functions of sorting and filtering the measures. After further fine-tuning of the functions, the interface can be implemented in the sponsoring company's HMI lab. Further user feedback could then be obtained when the toolkit is actually used.

6. Combined research impact

This Chapter summarises the practical impact of the research conducted in the frame of this EngD project. Section 6.1 describes the research impact for research stream one and Section 6.2 that for research stream two.

6.1. Research impact of research stream one

Research stream one comprised research that explored the tactile modality as an alternative less distracting way of in-vehicle communication and that investigated factors that can influence how drivers interact with a new device. The following sections describe how the research was relevant to the sponsoring company (Section 6.1.1), and to the wider automotive industry (Section 6.1.2). The section concludes with potential future research arising from the presented studies (Section 6.1.3). All sections include subsections for the research question in research stream 1.

6.1.1. Applicability for the sponsoring company

1) What is driver distraction, and when and how does driver distraction occur?
2) What strategies could an automotive company employ to mitigate driver distraction?

The first two research questions aimed to set the frame for the EngD and to ensure that mitigation strategies are selected that are interesting for automobile manufacturers. The research focused on the tactile modality as it does not require visual attention. Further, the tactile modality is comparable new to drivers in an in-vehicle setting and can serve the request for new technology that customers (particularly young customers) have.

3) Haptic feedback has been shown to be less visually distracting for the driver, however, what variables influence the perception of haptics?

The Haptic Pedal study was a user trial evaluating the influence of shoe type, gender, and age on the perception of haptic pulse feedback (Section 3.1). In preparation, a literature review was conducted to identify knowledge gaps in the design of haptic pedals. This literature review is described in submission 2 and includes use-cases of haptic pedals and settings for other types of haptic feedback such as vibrations. It can be used as an overview of existing use-cases and gaps for potential new use-cases. The

majority of use-cases was safety related. At the time of writing submission 2, only two studies related to tactile feedback on a pedal and eco-driving.

The results of the Haptic Pedal study indicate that shoe type does not influence the haptic perception, but gender and age do. This knowledge has implications for future studies involving a haptic pedal in the company. The development process of a new interface involves understanding the HMI aspect. This includes the evaluation of variables that can influence the interaction to customise the experience of the interface towards the drivers. Those variables are evaluated either in separate studies or as conditions in the same study. Each study costs time and resources to conduct. Each condition in a study adds to the duration of the study. Since the results of the Haptic Pedal study suggest that the shoe type does not influence the perception of a pulse on a haptic pedal it may not need to be considered in future studies.

Following another result from the Haptic Pedal study, user trials for haptic pedals should include young and old people and people of both genders to ensure a good noticeable feedback. For an industry study it is suggested to recruit people from a young and an old age group, for example "30 years and younger" and "60 years and older". Knowing there is a difference in haptic perception, the company might decide to focus on the age group of their target audience in user trials.

Some feedback settings received a high intensity rating and no pulses were missed over all participants. However, such a highly noticeable feedback received negative comfort ratings in the study. Such a setting could be implemented as warning. A warning does not require comfort, but it requires accurate detection.

4) Can a tactile warning as such or a tactile warning enhanced by another modality initiate a faster reaction time compared to a traditional auditory warning?

In the Warning study, the tactile modality was further analysed for its usefulness as warning (Section 3.3). Newly introduced warnings would only be economic if they were more efficient than an existing warning in a commercially available car. A tactile warning was compared to an auditory warning and an auditory-tactile warning. Literature suggests that warnings presented in two modalities have a higher perceptual level than a presentation of a warning in a single modality. Criteria for an ergonomic MMW were collected in a literature review and are summarised in submission 5.

The results of the Warning study suggest that an MMW is as effective as an auditory warning and is even perceived as being more noticeable. In contrast, the tactile warning was less effective and was rated comparably lower in noticeability. MMWs received high ratings of noticeability, but were also rated high in startlement. Because of that, MMWs might be unsuitable for an informative use-case when comfort becomes more important as opposed to a warning.

This knowledge about MMWs can be applied to different use-cases of warnings and within cars of different levels of automation. Examples for potential use-cases are lane-departure warning, frontal-collision warning, or in combination with driver state monitoring a warning for microsleep. It might even be used as an indication of a take-over scenario.

The warnings were compared over three distractor task conditions in three modalities. The tasks were artificial, so that distraction could be compared over different modalities, at a continuous level of demand. The distractor tasks were designed for a highly automated driving scenario and are not suitable to be conducted during manual driving. However, the setting can be adjusted. The design opens an opportunity to utilise the tasks for further research in distraction in an increasingly automated vehicle, in accordance with the sponsoring company's research goals.

5) How does a driver's trust in technology effect the visual interaction with a new in-vehicle device?

The Trust Brake study showed that the subjective ratings of trust increased between pre- and post-experience The Trust study and the Trust 3navi study showed the importance and positive effect that voice interaction can have (Section 4.1 and 4.2). In this study no glances off-road >2 s were applied in the voice interaction scenarios. A reason for that can be the design of the navigation system. In this study a Wizard-of-Oz navigation system was utilised to control for potential errors. Errors can increase the task completion time significantly and might contribute to larger glance durations (submission 3). Results of the Trust 3navi study further suggest that glance patterns might differ between types of interaction. In the voice-visual interaction similar glance patterns appeared as observed in the Trust study, however, the glance patterns were different in the visual-manual interaction. More research is required to understand those differences.

6.1.2. Applicability for the wider industry

1) What is driver distraction, and when and how does driver distraction occur?

2) What strategies could an automotive company employ to mitigate driver distraction?

The mitigation strategies for driver distraction investigated in this research are applicable to the wider automobile industry. The increasing implementation of electronics into the car is not only driven by the research on autonomous driving vehicles, it emerges directly from customer demands (McKinsey and Company, 2013). The decision whether to buy a car is not only driven by its perceived safety but also which technology it offers (Woodward et al., 2017).

The tactile modality does not require taking the eyes off-road, but yet it is still new to most drivers. Therefore, it can mitigate distraction and serve the demand for new technology. Similarly, the interaction between the drivers' trust in technology and interaction with new in-vehicle technology is relevant to all automobile manufacturers.

3) Haptic feedback has been shown to be less visually distracting for the driver, however, what variables influence the perception of haptics?

Results from the Haptic Pedal study (Section 3.1) have been published at a conference (Geitner et al., 2015) and are further planned to be published as journal paper. The journal paper is under review at the time of writing this report. The content describes how shoe type, gender and age affected the perception of a haptic pulse in the Haptic Pedal study, the pulse settings, and recommendations for settings in terms of noticeability and comfort. Similar to application of knowledge in the sponsoring company, other companies or research teams can use the information from the paper to define variables for their user trial. They can learn from the paper that they might not have to control for shoe types, and therefore focus on a balance of genders and include old participants. Further, the pulse settings could be taken forward to be applied in an actual use-case – such as a speed limit warning.

4) Can a tactile warning as such or a tactile warning enhanced by another modality initiate a faster reaction time compared to a traditional auditory warning?

Research on driver distraction remains relevant in the development of higher automated cars. The Society of Automotive Engineers (SAE) defined six levels of automation, ranging from zero with a completely manually operated car to five with an completely self-driving car that does not require any input from a human driver / operator (SAE, 2016). According to the SAE levels of automation there will be some sort of feedback from a human driver required up to level four, only level five does not require any input. This means that the design of the interaction between user and automated vehicle needs to incorporate the risk of a distracted user with potentially delayed reaction and accuracy of response. The Warning study (Section 3.3) pushes towards a different set of tasks for the evaluation of distraction in automated cars. The distractor task can become a primary task in a self-driving car scenario, in contrast to distraction in a manually driven car. The distractor tasks selected and developed for the Warning study contribute to the evaluation of driver distraction in a self-driving car scenario with highly-demanding tasks. The distractor tasks occupy the three typically used modalities for HMIs in a comparable workload. Therewith it is possible to evaluate the effectiveness of warnings comparing situations where different sensory channels are occupied. For example, naturalistic tasks that drivers engage in occupy different sensory modalities. Reading or watching a video occupies the visual sense. Listening to music, having a phone conversation or listening to an audio book occupies the auditory sense. Typing or playing a game on the mobile phone can occupy the haptic modality.

Further, the results of the Warning study are relevant for the design of warnings. Participants reacted significantly slower to the tactile warning compared to the MMW and to the auditory warning. A warning that requires an urgent reaction might be more effective in the auditory modality or as MMW. The results of the Warning study are planned to be shared with the wider research community and industry in a paper which is under review at the time of writing this report.

5) How does a driver's trust in technology effect the visual interaction with a new in-vehicle device?

Results of the Trust study have been published and shared with the research community in a conference paper. It is planned to publish the results of the Trust 3navi study was well.

6.1.3. Future Research

3) Haptic feedback has been shown to be less visually distracting for the driver, however, what variables influence the perception of haptics?

A limitation of the Haptic Pedal study (Section 3.1) is that a sound was perceptible in some pulse settings. This sound could have influenced the participants' perception of the pulse. Future studies should cover any mechanical noise delivered by a pedal with tactile feedback. A noise-cancelling headphone was in consideration for the study design, but had been abandoned because some of the noise was perceptible nevertheless. Playing white noise through the cabin loudspeakers would have been a better way to cancel potential mechanical noise from the pedal.

The Haptic Pedal study focused on an evaluation of perception which is the first step of the human-machine interaction process (Norman, 2002). Norman describes the interaction between human and machine as a series of steps beginning with perception of the state of the situation to the execution of an action sequence to complete a task. Future studies would need to specify a use-case for a pedal pulse feedback and test later stages of the interaction process. Specifically, future studies should evaluate the other steps of the human-machine interaction: if the pulse feedback communicates the information adequately and helps the user to select a suitable response.

4) Can a tactile warning as such or a tactile warning enhanced by another modality initiate a faster reaction time compared to a traditional auditory warning?

The Warning study (Section 3.3) focused on the perception of a tactile and a MMW in a highly distracting driving scenario in a self-driving car. The artificial distractor tasks were specifically selected and designed to create demand in three different modalities. Such a task design can be used to evaluate the robustness of the perception of signals further. Following a result from the Warning study, tactile feedback could be compared to traditionally utilised feedback forms in an informative use-case. For example, such a use-case could be navigation information.

A limitation of the Warning study was the auditory task condition. In the Warning study Pilot C (Section 3.2.3) the task settings were selected so that the tasks were perceived as similar in demand. However, considering that the RT was slower in the auditory task condition compared to that in the other tasks, this task needs to be improved. The participants mentioned that some of the sounds were difficult to

perceive. A remedy could be to rerecord the audio of the letters and digits, but this time already with a short duration in mind. In general, the task required the participants to adjust to fast spoken sounds. The original audio files were in normal conversation pace, some even a bit slower. If the audio would be recorded with a short duration in mind it could increase perceptibility of the letters and numbers. This strategy has been implemented in a new version of the task, but still needs to be evaluated (Github, 2017).

More research is needed to understand which "distractor" tasks drivers would be willing to perform in a highly automated or self-driving car. This is important in order to estimate response times and determine response procedures for take-over manoeuvres or emergency interventions that a user of an autonomous car is expected to do, besides evaluations with artificial tasks such as in the Warning study.

5) How does a driver's trust in technology effect the visual interaction with a new in-vehicle device?

In future research to the Trust and Trust 3navi study it would be of highest interest for a safe interaction with voice command systems to evaluate the effect of errors and different types of errors (e.g. a user related error such as a confusion of the address and system related error such as failed understanding of the user's command). A next step would then look into the feedback a system can provide to counteract potential effects of errors in the interaction dialogue. This study showed that it is possible to design a voice command interaction that does not cause glances off-road >2 s, despite providing visual feedback in all steps of the interaction dialogue to the participant.

6.2. Research impact of research stream two

Research stream two comprised the development of the HF toolkit, a database interface to support HMI engineers in their task of comparing and selecting measures for user trials. The following sections describe the research impact to the sponsoring company (Section 6.2.1), and to the wider automotive industry (Section 6.2.2). The section concludes with potential future research arising from the presented studies (Section 6.2.3). All sections include subsections for the research question in research stream 2.

6.2.1. Applicability for the sponsoring company

Research stream two, comprised the development of the HF toolkit – the innovative visual interface concept to support HMI engineers in the comparison and selection of measures for the evaluation of in-vehicle devices (Chapter 5). The HF toolkit is an incremental innovation. The concept of VIS has previously been successfully applied to present large amounts of information in a fast retrievable way in, for example, health-care (Plaisant et al., 1996), a movie-database (Ahlberg and Shneiderman, 1994), and a library system (Heilig et al., 2008). In this research project the concept of VIS was adapted to the measure comparison and selection process of HF measures to support HMI engineers in the automotive industry in their planning and conduction of user trials. The development process was a co-design process with the sponsoring company's HMI engineers as they were the intended users.

1) How do designers select measures for user studies?

At the beginning of this research stream, user requirement interviews were conducted to identify how designers in the JLR HMI research team select measures for user studies. The selection process is described based on their experience. Because, the interviews revealed that no electronic aid was available for this process and the process would benefit from a knowledge database, the project proceeded with the development of an electronic aid for measure selection.

2) Can measure selection benefit from electronic support, and, if so, how can designers be supported in their task in a usable way?

Based on the user requirement interviews it was determined that the measure selection process would benefit from an electronic aid (toolkit). To design a usable product, the user centred design process was employed and the toolkit was developed in four paper prototype evaluations together with designers from JLR as the future users of the toolkit. In the end, the usability and user experience of the toolkit was rated good.
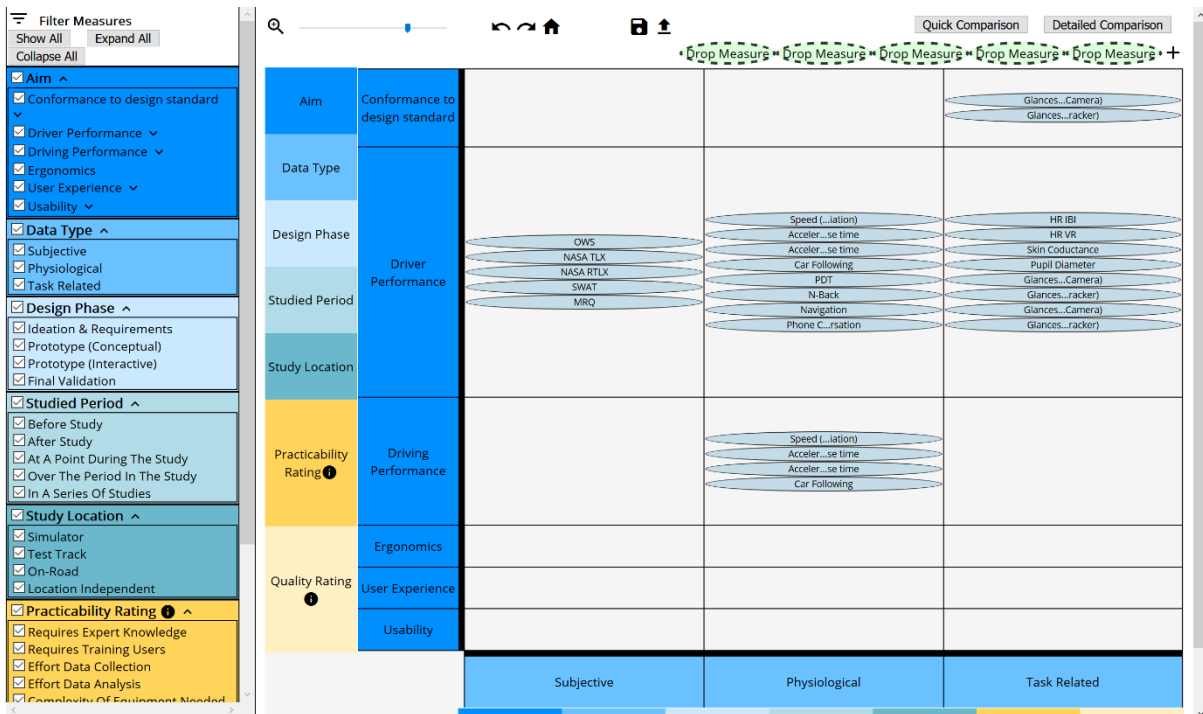
Figure 18. HF Toolkit initial software implementation.

An undergraduate student implemented the conceptual interface as presented in Chapter 5 as a software prototype during a summer internship of eight weeks (Figure 18). This prototype is a first feasibility test to investigate the best way to implement the concept. The interface consists largely of typical interface design elements. However, the presentation of tabs to sort measures is an interface element specific to the HF toolkit. Decisions about the implementation were made in consideration of work procedures in the sponsoring company. An HMI engineer works in several distributed locations, e.g. the test ground, the HMI lab, or at the dedicated desk. For example, it is typical that a user trial is planned in office, is conducted in the HMI lab, and is analysed in office. The toolkit can offer support in all those steps (measure comparison / selection when the user trial is planned, templates for data collection, a best practice description for the data analysis). To be most usable, the access to the toolkit would need to be location independent. Another constraint to consider is that HMI engineers use company owned computers with restricted rights to install software and therefore the HF toolkit should not require Administrator rights.

In consideration of these requirements, it was decided to implement the toolkit as a web application that runs in a browser. Web applications are independent from the operating system, and HMI engineers can access the application from any computer with network access without an installation. The prototype is implemented in the

Django framework. The framework offers a modular approach to build websites in the programming language Python. The Django framework only needs to be installed on the computer that will be used as server. The framework creates a local server process on the computer that handles communication with the client web browsers. The backend of the website is a database in which the measures will be stored. There is a browser-based interface to the database, so that information about measures can be amended and new measures can be entered without detailed technical knowledge. The first prototype implementation was tested with the two commonly used browsers Google Chrome and Mozilla Firefox.

The prototype foresees a link to the company's existing information landscape. The HMI engineers can export a Comma-Separated Values (CSV) file of a measure list. This CSV file will be imported in an interface that comprises options to collect physiological data in the HMI lab. Dependent on the selected measures in the CSV file different options in the interface appear. For example if eye tracking measures are selected in the CSV then those options will be shown in the interface. Besides this, existing procedures and templates for measures can be linked to the toolkit. Examples are electronic versions of questionnaires, such as the NASA TLX, electronic versions of distractor tasks, such as the n-back task, and procedures on best practice for measures. Additionally, the toolkit offers the option to link papers about a measure. This option offers the chance for HMI engineers to share their local collection of papers with other HMI engineers. HMI engineers will find information about the measure from use-cases (in literature), how to use it (implementation), up to analysis of data obtained with the measure.

One of the sponsoring company's business goals is to innovate by learning more about their users. The toolkit can support this by providing a best practice application of measures. Because all measures are presented in the initial interface, HMI engineers are exposed to measures they may not have used yet which can provide additional information about the interaction process between driver and device. The quality of user trials can be increased by the description of the best practice for the application of a measure and its analysis. For example, HMI engineers can inform themselves about how to adjust eye-tracking equipment and which eye metrics best to use to detect visual distraction.

A limitation is that whereas the toolkit interface was developed in a user-centred design process and the interface applying the VIS approach was preferred subjectively and objectively compared to the others, the toolkit yet needs to prove how much it can ease an HMI engineer's work in practice. The implemented prototype consists of the visual design of the conceptual interface (Figure 18). The structure of the database foresees that all information that the toolkit should contain about a measure can be stored in it – overview and detailed information - as it is described in the concept (Section 5.6). Filter and sorting functions are implemented in the prototype. It is also possible to export a list of measures into a CSV file. However, due to the time-constraints of the six week summer internship the measure comparison function is not yet implemented.

Another potential challenge of the HF toolkit, as with any software based knowledge management system, is to keep the database up-to-date. At best, the interaction with the toolkit would be integrated in routine processes such as team meetings where information available in the toolkit can be discussed. A first step is a link to existing infrastructure, the implementation of the toolkit in the HMI lab and a link from existing software to the toolkit.

6.2.2. Applicability for the wider industry

The automotive industry is changing with increasing automation and fewer profit margins in the satisfied markets in Europe, Japan, and U.S. (McKinsey and Company, 2013). It is expected that profit arises increasingly from customisable products and extended services, such as for updates of in-vehicle software, in those markets. The development has impact on the internal processes of a company. The automobile as product needs to be manufactured compartmentalised in modules that customers can choose from dependent on their demand. This results in a larger amount of data in the development process. There is an ongoing trend in the development of new tools to suit that process and that are better capable to deal with larger amounts of data. The HF toolkit is such a tool for HF measure selection and comparison.

1) How do designers select measures for user studies?

The user requirement interviews were conducted with designers from JLR only. Whereas it is assumed that the process would be similar in other teams that run user studies, it cannot be concluded from this research.

2) Can measure selection benefit from electronic support, and, if so, how can designers be supported in their task in a usable way?

The toolkit was designed together with HMI engineers from the sponsoring company to support them in their task of planning and conducting user trials. However, the toolkit can be applied to other areas that utilise human factors measures, for example, the design of medical devices and the design of control rooms. Some categories remain the same, others may require changes, such as the database content that contains measures prescribed in industry specific standards and guidelines. To ensure a usable interface of the HF toolkit it is suggested to adapt the toolkit's interface in the process shown in Figure 19.
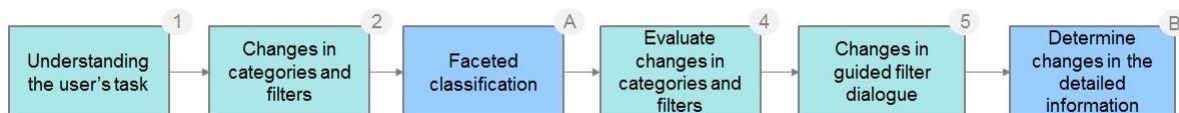

Figure 19. Process to adapt the HF toolkit interface to another company or research team (the letters mark two optional steps).

The adaptations mainly concern changes of labels or detailed information. Those changes can be conducted by changes in the database backend without detailed technical knowledge. However, it might be necessary to increase the number of filter categories or to increase the categories to sort measures. Such changes would require changes in the code of the HF toolkit.

First, it is suggested to conduct a series of interviews and find out how users select measures and how they would like to be supported (Figure 19, (1)). Based on the selected information user requirements can be generated. Those requirements can then be applied to determine changes in the categories for filter and overview (Figure 19, (2)). Small changes can be incorporated immediately; before larger changes are implemented it is recommended to conduct a faceted classification (Figure 19, (A)). Filter and overview categories are a crucial part of the interface and it needs to be ensured that the users know what they mean to quickly retrieve information from the interface. The revised interface should be evaluated in quick prototype sessions with users (Figure 19, (4)), three to six users are sufficient to discover major usability issues

according to Nielsen (2009). Changes to filters and overview categories require changes in the guided filter dialogue (Figure 19, (5)). In the last, optional, step the user requirements need to be applied to determine changes in the detailed information about the measures (Figure 19, (B)).

This process to adapt the HF toolkit interface to other industries or research teams is planned to be shared with the wider research community in a paper which is under review at the time of writing this report. The concept of the HF toolkit as such and the result of the first paper prototype study have been presented at a conference (Geitner et al. (a), 2017).

6.2.3. Future research

2) Can measure selection benefit from electronic support, and, if so, how can designers be supported in their task in a usable way?

Whereas there exists a prototype of the toolkit, it is not yet fully functional. A step further in this project is develop the prototype into a fully functional tool. Mainly this concerns the measure comparison function. The prototype toolkit includes a set of dummy measures with overview information. Detailed information and existing templates or procedures need to be added to those measures.

The HF toolkit is planned to be implemented in the sponsoring company's HMI lab when the functionality is developed further. In first trials, HMI engineers can use the toolkit in its actual context of use. Its information and interface can then be further refined, if necessary. For this evaluation it is suggested to use the logging function of the toolkit to log the interaction and further let HMI engineers complete a usability and a user experience questionnaire. Subjective feedback from the HMI engineers has been gathered with the SUS (Brooke, 1996) and the UEQ (Van der Laan et al., 1996) during the paper prototype evaluations. Those two questionnaires are recommended to be used as they are commonly used and fast to complete and analyse. Additionally, HMI engineers should be able to write free-text comments on positive and negative feedback. Such feedback might include difficulties HMI engineers experienced or things they specifically enjoyed which is not covered by the pre-defined answers given in the questionnaires. Such comments combined with an analysis of the interaction can make it easier to identify elements in the interface dialogue that should be

improved. The questionnaires give a general informative feedback on how supportive the HMI engineers experienced the toolkit in their task. The HMI engineers' comments should be compared to observations of the interaction process and feedback in usability or user experience. Comments can give detailed information why the HMI engineers gave a negative feedback or encountered difficulties in the interaction. The more detailed information can make it easier to decide on improvements of the interface.

In a wider context, it would be interesting to see the toolkit implemented as an online research tool, for example to collect information about measures for driver distraction. It could be used as a platform for information exchange about measures and their best practice application. Researchers might share software templates they used in their studies, for example implementations of distractor tasks and templates for data analysis. This would help to make studies more comparable and increase study quality by discussing best practices.

6.3. Summary

Main innovations from research stream one are:

- New knowledge and recommendations for design of tactile communication with the driver over pedals (Haptic Pedal study)
- Benchmark study for the performance (subjective an objective) of a traditional auditory warning compared to a MMW and tactile warning in a self-driving car scenario with highly attention capturing tasks - whereas the auditory warning performed equally as good as the MMW, the MMW had advantages in a lower false alarm rate and lower rate of missed alarms (Warning study)
- Combination and refinement of the RSVP and RSAP task from literature with a newly created tactile equivalent for the evaluation of distraction in a self-driving car scenario (Warning study)
- The Trust study and the Trust 3navi study showed there is a link between trust and glance behaviour. The Trust 3navi study showed that interaction with the voice-only and voice-visual system both involved glances at the navigation display, however, none was >2 s. The visual-manual interaction, in comparison, involved glances off-road >2 s. The observed relationship between trust and

glances was not consistent in all three interfaces, it might be that it differs
dependent on the nature of the task (voice command vs. visual-manual)

Main innovations from research stream two: The HF toolkit is directly applicable for
the sponsoring company. The sponsoring company intends to use the HF toolkit in its
driving simulator. Summarising, the final conceptual design supports HMI engineers
in the following ways in their process of planning and conducting user trials (Figure
20):

- Sharing knowledge about measures in the company
- Integration of measures that are required by standards, such as NHTSA
- Easier measure comparison and information retrieval
- Determining and collecting practical implicit knowledge about a measure
- Aid to implement a best practice for measure's implementation and analysis
- Integration of templates for a measure, e.g. electronic version of a questionnaire

In this project a new tool has been developed that did not previously exist in the
sponsoring company and in the published literature of automobile manufacturers. The
final interface design is applicable to automobile manufacturers to organise HF related
measures for the evaluation of in-vehicle devices, but could be adapted to other areas
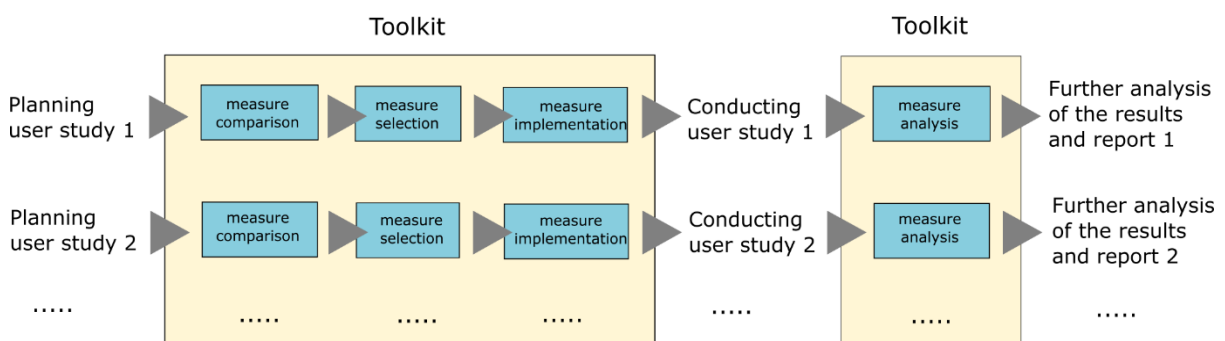in industry and research that utilise HF measures.

Figure 20. Toolkit support in relation to the process of conducting a user trial.

7. Conclusions

This EngD project started with highlighting the importance of driver distraction for the automotive industry. When drivers divert their attention away from the driving task to engage in another task their accident risk can increase due to reduced situational awareness and increased reaction time (RT) to sudden changes in on-road events. An objectively measurable indicator for driver distraction, for example proposed in National Highway Safety Administration's guideline for less distracting in-vehicle design (NHTSA, 2013), is glance behaviour. Research stream one focused on the reduction of glances off-road , by investigating tactile feedback as non-visual way to communicate with the driver. Considering that the utilisation of the haptic sense as an interface modality in a car is relatively new to drivers, it was then investigated how drivers interact with comparably new interfaces.

Contributions from research stream one can be described by the studies conducted within the frame of this research stream, listed by objectives defined in Section 1.1:

- The Haptic Pedal study extends the knowledge about tactile in-vehicle communication by evaluating pulse feedback delivered by an accelerator pedal (Section 3.1):
  - o Shoe type, plimsolls vs. safety boots, did not influence the perception of a pulse feedback, and in consequence might not be controlled in future studies involving a tactile pedal
  - o Gender and age can influence the perception of a pulse feedback delivered by a pedal. Females rated tactile feedback higher in intensity and high intense tactile feedback more negatively compared to males. Older participants (60 years and older) missed a higher percentage of short duration pulses (20 ms and 33 ms) compared to the younger participants (39 years and younger).
  - o Based on the ratings, it is recommended to utilise durations longer than 33 ms and amplitudes greater than 9 N for a good noticeable tactile feedback
  - o Tactile feedback that was rated high in intensity, where participants did not miss pulses, tends to be rated negatively in comfort

- The Warning study compared the performance of a tactile warning to a traditional auditory warning and a multimodal (auditory-tactile) warning in a self-driving car scenario with a distractor task (Section 3.3):
    o A tactile warning led to a slower RT compared to the auditory-tactile (multimodal) and to the auditory warning – it might not be most effective to be used as warning
    o Enhanced with an auditory component tactile feedback can lead to faster RTs, even slightly better (but non-significantly) than a traditional auditory warning in an auditory distractor task condition. The auditory-tactile warning led to fewer missed alarms than the auditory warning and fewer false alarms compared the tactile warning.
    o A procedure from literature was applied to adjust the setting of multiple warnings quickly, so they are perceived as equally intense. This procedure is described in the submission and recommended to be used whenever warnings are evaluated, otherwise potential performance differences of warnings might be confounded by a difference in intensity.
    o Lessons learned for the design of a multimodal warning
    o Combination and refinement of the Rapid Serial Visual Presentation task and Rapid Serial Auditory Presentation task from literature with a newly created tactile equivalent for the evaluation of distraction in self-driving car scenario (Warning study), implementation of the tasks in an easily adjustable interface so the tasks can be presented in varying levels of demand – a platform independent web interface that can be used on a computer, tablet, or smartphone dependent on the study setting
- The Trust study and the Trust 3navi study evaluated effects of a driver's trust in new technology (Section 4.1):
    o Trust and glance behaviour were linked
    o The Wizard-of-Oz voice interaction with no error did not involve any glances to the navigation system >2 s

The interaction between the driver and another device can be observed with a number of measures. In the trend towards a personalised automobile, in which customers are able to select their in-vehicle design and services modularised, it becomes more important to learn about the customer. This learning is part of the design process and reflects in an utilisation of the wide range of measures available for the evaluation of the interaction between user and interface. However, to obtain a useful result from a user trial the utilised measures need to be applied correctly to keep the study comparable and avoid confounding variables. Research stream two addressed this problem by developing an interface concept for a Human Factors (HF) measure database, which supports Human-Machine Interface (HMI) engineers in the process of understanding, comparing, selecting and utilising HF measures for a user trial (Chapter 5). Further, the proposed database functions as knowledge management tool in which information about equipment required for a measure and how data obtained from a measure can be analysed in best practice can be stored at one place. The interface is called HF toolkit. The overall contribution can be divided into sub-contributions, listed by the objectives for research stream two defined in Section 1.1:

- First, it was important to gain an understanding of the measure selection process and the potential for support (Section 5.3):
  - HMI engineers and managers both found the idea of a HF toolkit useful. HMI engineers saw value in a quick way to compare measures and to be able to find a description of how the measure is utilised in best practice. The managers specifically found that a toolkit could help to consistently apply a measure and thereby improve user trials.
- Then a concept for the interface was developed by applying the Visual Information Seeking (VIS) principle to the area of measure selection and comparison (Section 5.4):
  - Whereas there are collections of usability measures online, it is the first visual based measure collection dedicated to the evaluation of in-vehicle devices and the first visual tool to support HF measure comparison and selection from published literature.
  - For an easy information retrieval, the VIS principle was applied to measure comparison and selection, which has not been done before.

- Then the concept was iteratively improved in a series of four paper prototyping evaluations together with the future users – the HMI engineers of the sponsoring company (Section 5.5):

    o In the first paper prototype iteration three interfaces were compared, two interface designs based on the VIS concept and one interface designed as spreadsheet. An interface based on the VIS concept with a known metaphor of a diagram as visualisation was the preferred interface (Diagram concept), based on objective and subjective measures.

    o The Diagram concept was then improved iteratively in three paper prototyping evaluations. The final paper prototype interface design was developed into an interaction flow-chart for a software implementation. Thereafter, the interface concept was developed into an initial software prototype in the course of a summer internship and handed over to the sponsoring company.

This research extended knowledge about a less distracting communication between drivers and in-vehicle devices and about the interaction between drivers and newly introduced in-vehicle devices, further, this research improved the evaluation process of in-vehicle devices by describing a procedure for the comparison, selection and utilisation of HF measures and the development of an electronic aid for that process.

References

Abe, G., Itoh, M., and Tanaka (2002). Dynamics of drivers' trust in warning systems. In *15th Triennial World Congress of the International Federation of Automatic Control,* p. 1612. New York, NY: Elsevier. doi:10.3182/20020721-6-ES-1901.01614

Abbink, D. and Van der Helm, F. (2004). Force perception measurements at the foot. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, Vol. 3, 2525–2529. doi: 10.1109/ICSMC.2004.1400709

Adell, E., Várhelyi, A., and Hjälmdahl, M. (2008). Auditory and haptic systems for in-car speed management – A comparative real life study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 11(6), 445–458.

Ahlberg, C., and Shneiderman, B. (1994). Visual information seeking using the FilmFinder. *Proceeding CHI '94 Conference Companion on Human Factors in Computing Systems, 433-434.*

'All About UX' (2016). Retrieved from http://www.allaboutux.org/

Aloisio, L. and Mrasek, N. (2017). Autotrader study finds 48 percent of car buyers prioritize in-vehicle technology over brand or body style. Autotrader. Retrieved from http://press.autotrader.com/2017-01-10-Autotrader-Study-Finds-48-Percent-of-Car-Buyers-Prioritize-In-Vehicle-Technology-Over-Brand-or-Body-Style

Bagheri, N., and Jamieson, G. (2004). *Considering subjective trust and monitoring behavior in assessing automation-induced "complacency".* In D. A. Vicenzi, M. Mouloua, and P. A. Hancock (Eds.), Human performance, situation awareness, and automation, HPSAA II, pp. 54–59. Mahwah, NJ: Erlbaum.

Bangor, A., Kortum, P., and Miller, J., 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4 (3), 114-123.

Biondi, F., Leo, M., Gastaldi, M., Rossi, R., and Mulatti, C. (2016). How to drive drivers nuts: Effects of auditory, vibrotactile, and multimodal warnings on perceived urgency, annoyance and acceptability. *Transportation Research Board 96th Annual Meeting*, February 2017.

Biondi, F., Strayer, D., Rossi, R., Gastaldi, M., and Mulatti, C., 2017. Advanced driver assistance systems: Using multimodal redundant warnings to enhance road safety. *Applied Ergonomics*, 58(2017), 238-244.

Birrell, S., Young, M., and Weldon, A. (2013). Vibrotactile pedals: Provision of haptic feedback to support economical driving. *Ergonomics*, 56(2), 282–292. doi:10.1080/00140139.2012.760750

BMW (2015). Voice Control System. BMW company website. Retrieved from http://www.bmw.com/com/en/insights/technology/technology_guide/articles/voice_control_system.html (accessed December, 2017)

Brooke, J., (1996). *SUS – A quick and dirty usability scale.* In P. W. Jordan, B. Thomas, B.A. Weerdmeester and I.L. McClelland, Eds.; Usability Evaluation in Industry, pp. 189-194. London: Taylor & Francis.

Broström, R., Bengtsson, P., and Ljung Aust, M., (2015). Individual glance strategies and their effect on the NHTSA visual manual distraction test. *Transportation Research Part F: Traffic Psychology and Behaviour*, 36(2016), 83-91.

Broughton, V. (2005). The need for a faceted classification as the basis of all methods of information retrieval. *Aslib Proceedings*, 58(1), 49–72.

Brown, S. (2005). *Effects of haptic and auditory warnings on driver intersection behavior and perception.* Virginia Polytechnic Institute and State University. Retrieved from: http://vtechworks.lib.vt.edu/handle/10919/31844

ButtKicker (2017). Webpresence of the company Guithammer. Manual of the ButtKicker system. Retrieved from http://thebuttkicker.com/advance

Chapanis, A. (1965). *Man-machine engineering.* Wadsworth Pub. Co. (1965)

Chang, W., Hwang, W., and Ji, Y. (2011). Haptic seat interfaces for driver information and warning systems. *International Journal of Human-Computer Interaction*, 27(12), 1119-1132. doi: 10.1080/10447318.2011.555321

De Rosario, H., Loureda, M., Diáz, I., and Jornet, J. (2010). Efficacy and feeling of a vibrotactile Frontal Collision Warning implemented in a haptic pedal. *Transportation Research Part F: Traffic Psychology and Behaviour*, 13(2010), 80-91.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., et al. (2006). The 100-Car naturalistic driving study, phase II – Results of the 100-car field experiment. (NHTSA Report DOT HS 810 593). Retrieved from the National Highway Traffic Safety Administration website: http://www-nrd.nhtsa.dot.gov/departments/human%20factors/driver-distraction/PDF/100CarMain.pdf

Djamasbi, S., Siegel, M., and Tullis, T. (2010). Efficiency, trust, and visual appeal: Usability testing through eye tracking. In *Proceedings of the Forty-Third Annual Hawaii International Conference on System Sciences (HICCS), Computer Society Press*, 2010, 1-10. Retrieved from: http://digitalcommons.wpi.edu/uxdmrl-pubs/21

Engström, J., Johansson, E., and Oestlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F-Traffic Psychology and Behaviour*, 8(2), 97-120. doi: 10.1016/j.trf.2005.04.012

Fitch, G., Kiefer, R., Hankey, J., and Kleiner, B. (2007).Toward developing an approach for alerting drivers to the direction of a crash threat. *Human Factors The Journal of the Human Factors and Ergonomics Society*, 49(4), 710-20. doi: 10.1518/001872007X215782

Geitner, C., Birrell, S., Krehl, C., and Jennings, P. (2015). Good vibrations: driving with a haptic pedal. In Proceedings of AutomotiveUI '15 Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Nottingham, United Kingdom.

Geitner, C., Birrell, S., Skrypchuk, L., Biondi, F., Jennings, P. (a) (2017) Research knows best, but how to communicate distraction measures practically in an industrial context. In Proceedings of the 5th International Conference on Driver Distraction and Inattention, Paris, France, 20-22 Mar 2017.

Geitner, C. Sawyer, B. D., Birrell, S., Jennings, P., Skrypchuk, L., Mehler, B., and Reimer, B. (b) (2017). A link between trust in technology and glance allocation in on-road driving. Proceedings of *the 9th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Manchester Village, Vermont.

Github (2018). Online repository for the detection tasks. https://github.com/cognitiveresources/Detection-Tasks

Gold, C., Körber, M., Hohenberger, C., Lechner, D., and Bengler, K. (2015). Trust in automation – Before and after the experience of take-over scenarios in a highly automated vehicle. In *Proceedings of the 6th International Conference on Applied Human Factors and Ergonomics, AHFE 2014*, Las Vegas, US.

GoogleVR (2018). Google Cardboard – Experience virtual reality in a simple, fun and affordable way. Retrieved from: https://vr.google.com/cardboard/

Gov (2017). Using mobile phones when driving. Retrieved from: https://www.gov.uk/using-mobile-phones-when-driving-the-law

Greenberg, J., Tijerina, L., Curry, R., Artz, B., Cathey, L., Kochhar, D., Kozak, K., Blommer, M., Grant, P. (2003). Driver distraction: evaluation with event detection paradigm. *Transportation Research Record*, 1843(2003), 1-9.

Hale, K. and Stanney, K. (2004). Deriving haptic design guidelines from human physiological psychophysical, and neurological foundations. *Computer Graphics and Applications, IEEE*, 24(2), 33-39.

Hancock, P., Lesch, M., & Simmons, L. (2002). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis and Prevention,* 35(2003), 501-514.

Hawkins, A. (2017). This self-driving van concept from Volkswagen looks like a pissed-off toaster. The Verge. Retrieved from http://www.theverge.com/2017/3/6/14832354/volkswagen-sedric-self-driving-van-concept-mobility (accessed December, 2017)

Heilig, M., Demarmels, M., Koenig, W., Gerken, J., Rexhausen, S., Jetter, H.-C., and Reiterer, H. (2008). MedioVis – Visual information seeking in digital libraries. . In *Proceedings of the working conference on Advanced visual interfaces, AVI '08*, pages 490–491, New York, NY, USA, 2008. ACM

Highway Code (2017). The Highway Code. Department of Transport. Retrieved from https://www.gov.uk/guidance/the-highway-code/general-rules-techniques-and-advice-for-all-drivers-and-riders-103-to-158 (accessed December, 2017)

Hill, S. G., Lavecchia, H. P., Bittner, A. C., Zakled, A. L., and Christ, R. E. (1992). Comparison of Four Subjective Workload Rating Scales. *Human Factors - Special Issue: Measurement in Human Factors Archive*, 34(4), 429-439.

Hope, K. (2016). UK 'has never been more addicted to smartphones'. BBC news. Retrieved from http://www.bbc.co.uk/news/business-37468560 (accessed November, 2017)

Horrey, W. J., Lesch, M. F., and Garabet, A. (2008). Assessing the awareness of performance decrements in distracted drivers. *Accident; Analysis and Prevention*, 40(2), 675-682. doi: 10.1016/j.aap.2007.09.004

Ichinose, A., Gomikawa, Y., and Suzuki, S. (2013). Driving assistance through pedal reaction force control with consideration of JND. *In RO-MAN*, 2013 IEEE, 484–489. doi: 10.1109/ROMAN.2013.6628551

ISO/IEC (1998). 9241-14 Ergonomic requirements for office work with visual display terminals (VDT)s – Part 14 Menu dialogues, ISO/IEC 9241-14: 1998 (E), 1998.

Jaguar Land Rover (a) (2014). New discover sport achieves five star Euro NCAP rating. Jaguar Land Rover company website. Retrieved from http://www.landrover.com/experiences/news/euro-ncap.html (accessed December, 2017)

Jaguar Land Rover (b) (2014). Land Rover introduces world first laser head-up display for Range Rover Evoque. Retrieved from http://newsroom.jaguarlandrover.com/en-gb/land-rover/news/2014/09/rr_evoque_2015_hud_190914/ (accessed January, 2015)

Jaguar Land Rover (c) (2014). Jaguar Land Rover Company Information. Information Leaflet, Subchapter "Our Facilities" (page 6). Jaguar Land Rover company website. Retrieved from: http://www.jaguarlandrover.com/media/23076/jlr_company_information.pdf (accessed May, 2016)

Jaguar Land Rover (a) (2017). Introducing the Range Rover Velar. Retrieved from: media.landrover.com/en-gb/news/2017/03/introducing-range-rover-velar (accessed December, 2017)

Jaguar Land Rover (b) (2017). Jaguar Land Rover leads in MOVE_UK and UK-CITE projects. Jaguar Land Rover company website. Retrieved from: https://www.sbdautomotive.com/en/jaguar-land-rover-move_uk (accessed December, 2017)

Jaguar Land Rover (c) (2017). Jaguar Land Rover PLC – Annual report 2016/17. Jaguar Land Rover company website. Retrieved from: http://annualreport2017.jaguarlandrover.com/assets/files/jlr_ar16_17.pdf (accessed December, 2017)

Jaguar Land Rover (d) (2017). What's new for Land Rover: 2017/2018/2019 model year highlights. Retrieved from: http://media.landrover.com/en-us/news/2017/11/whats-new-land-rover-2017-2018-2019-model-year-highlights (accessed December, 2017)

Jahn, G., Oehme, A., Krems, J., and Gelau, C. (2005). Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F*, 8(2005), 255-275.

Japanese Automotive Manufacturers Association (JAMA) (2004). Guideline for in-vehicle display systems, Version 3.0. Retrieved from: http://www.jama.or.jp/safe/guideline/pdf/jama_guidelines_v30_en.pdf

Jensen, B., Skov, M., and Thiruravichandran, N. (2010). Studying driver attention and behaviour for three configurations of GPS navigation in real traffic driving. In *Proceedings of the CHI 2010: Driving, Interrupted*, Atlanta, Georgia, USA.

Keetels, M. and Vroomen, J. (2012). *Perception of synchrony between the senses*. In Murray M. M., and Wallace M. T. (Eds.): Frontiers in the neural basis of multisensory processes (pp. 147–177). London, UK: Taylor & Francis Group.

King, A., and Calvert G. (2001). Multisensory integration: perceptual grouping by eye and ear. *Current Biology*, 11, 322–325.

Klauer, S., Dingus, T., Neale, V., Sudweeks, J., and Ramsey, D. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-Car Naturalistic Driving Study data*. (Report No. DOT HS 810 594). Washington, DC: National Highway Traffic Safety Administration.

Lees, M., and Lee, J. D. (2007). The influence of distraction and driving context on driver response to imperfect collision warning systems. *Ergonomics*, 50(8), 1264-1286.

Liang, Y. (2009). *Detecting driver distraction* (Doctoral dissertation). Retrieved from University of Iowa http://ir.uiowa.edu/etd/248

Liu, Y. (2003). Effects of using head-up display in automobile context on attention demand and driving performance. *Displays*, 24(4-5), 157-165.

Maguire, M. and Bevan, N. (2002). User requirements analysis - A review of supporting methods. In *Proceedings of IFIP 17th World Computer Congress*, 25-30 August 2002, 133-148.

Martens, M. and Van Winsum, W. (2001). Effects of speech versus tactile driver support messages on driving behaviour and workload. TNO Human Factors. Retrieved from http://www-nrd.nhtsa.dot.gov/pdf/ESV/esv17/Proceed/00080.pdf

McIlroy, R., and Stanton, N. (2017). Good vibrations: Using a haptic accelerator pedal to encourage eco-driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 46, (Part A), 34-46.

McKinsey and Company (2013). The road to 2020 and beyond: What's driving the global automotive industry? Retrieved from: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-road-to-2020-and-beyond-whats-driving-the-global-automotive-industry

McMahon, C., Lowe, A., and Culley, S. (2014). Knowledge management in engineering design: personalization and codification. *Journal of Engineering Design*, 15 (4), 307-325.

Mehler, B., Kidd, D., Reimer, B., Reagan, I., Dobres, J., and McCartt, A. (2016). Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics*, 59(3), 344-67.

Metha, R. and Shrivastava, M. (2012). An automatic approach for eye tracking and blink detection in real time. *International Journal of Engineering and Innovative Technology (IJEIT)*, 1(5), 140-143.

Mohebbi, R., Gray, R., and Tan, H. (2009). Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Human Factors*, 51, 102–110.

National Highway Traffic Safety Administration (NHTSA) (2013). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices* (National Highway Traffic Safety Administration Publication No. NHTSA-2010-0053). Retrieved from http://www.nhtsa.gov/About+NHTSA/Press+Releases/U.S.+DOT+Releases+Guidelines+to+Minimize+In-Vehicle+Distractions

Nielsen, J. (2009). Discount Usability: 20 Years. Article on Nielsen Norman Group. Retrieved from: https://www.nngroup.com/articles/discount-usability-20-years/

Norman, D. A. (1990). The problem of automation: Inappropriate feedback and interaction, not over-automation. *Philosophical Transactions of the Royal Society of London*, B, 1990. Retrieved from http://www.jnd.org/dn.mss/Norman-overautomation.pdf

Norman, D. (2002). *The Design of Everyday Things*. Basic Books, Inc. New York, NY, USA, 2002.

Pettitt, M., Burnett, G., and Stevens, A. (2005). Defining driver distraction. Paper to be presented at *World Congress on Intelligent Transport Systems*, San Francisco, November 2005.

Plaisant, C., Milash, B., Rose, A., Widoff, S., and Shneiderman, B. (1996). Lifelines: Visualizing personal histories. In *Proceedings of ACM CHI '96*, 221-227. doi: 10.1145/238386.238493

Press, G. (2014). Internet of things by the numbers: Market estimates and forecasts. Forbes. Retrieved from https://www.forbes.com/sites/gilpress/2014/08/22/internet-of-things-by-the-numbers-market-estimates-and-forecasts/#5ea27144b919 (accessed October, 2017)

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria. Retrieved from http://www.R-project.org/

Recarte, M. and Nunes, M. (2003). Mental workload while driving: Effects on visual search, discrimination and decision making. *Journal of Experimental Psychology: Applied*, 9(2), 119-137. doi: http://dx.doi.org/10.1037/1076-898X.9.2.119

Reimer, B., Mehler, B., Son, J., Pohlmeyer, A., Orszulak, J, Long, J. and Coughlin, J. (2009). A cross-cultural comparison of younger and older adults simulated highway driving performance under single and dual task conditions. In *Proceedings of the 5th International Driving Symposium on Human Factors in Driving Assessment, Training, and Vehicle Design*. Retrieved from Research Gate Account of Brian Reimer.

Reimer, B., Mehler, B., Dobres, J., McAnulty, H., Mehler, A., Munger, D., and Rumpold, A. (2014). Effects of an 'Expert Mode' voice command system on task performance, glance behavior & driver physiology. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM.

SAE (2016). *Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Standard published by SAE International, SAE J3016_201609.

Santangelo, V. and Spence, C. (2007). Multisensory cues capture spatial attention regardless of perceptual load. *Journal of Engineering Psychology*, 33(6), 1311-1321.

Sawyer, B., Seppelt, B., Mehler, B., and Reimer, B. (2017). Trust impacts drive glance strategy in multitasking. In *Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting*, 1441-1442.

Scuro, G. (2017). Automotive industry: Innovation driven by electronics. Embedded computing design. Retrieved from http://embedded-computing.com/articles/automotive-industry-innovation-driven-electronics/ (accessed December, 2017)

'Service Design Tools' (2016). Retrieved from: http://www.servicedesigntools.org/taxonomy/term/4

Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualisation. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, 336-343. DOI: 10.1109/VL.1996.545307

Soto-Faraco, S., and Spence, C. (2002). Modality-specific auditory and visual temporal processing deficits. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 55(A), 23–40.

Spence, C. and Ho, C. (2017). *The multisensory driver: Implications for ergonomic car interface design*. CRC Press.

Stanley, L. (2006). *Haptic and auditory interfaces as a collision avoidance technique during roadway departures and driver perception of these modalities* (Doctoral dissertation). Retrieved from: https://scholarworks.montana.edu/xmlui/bitstream/handle/1/2338/stanleyl0506.pdf?sequence=1

Statista (2018). Number of mobile phone users worldwide from 2013 to 2019 (in billions). Retrieved from: https://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/ (accessed January, 2018)

Stutts, J., Reinfurt, D., Staplin, L., and Rodgman, E. (2001). The role of driver distraction in traffic crashes. *Prepared for AAA Foundation for Traffic Safety*. Retrieved from http://www.forces-nl.org/download/distraction.pdf

The, E., Jamson, S., Carsten, O., and Jamson, H. (2014). Temporal fluctuations in driving demand: the effect of traffic complexity on subjective measures of workload and driving performance. *Transportation and Research Part F: Traffic Psychology and Behaviour*, 22, 207-217.

Toyota (2014). Toyota unveils latest advances in automated driving and safety technology. Toyota company website. Retrieved from http://blog.toyota.co.uk/toyota-unveils-latest-advances-in-automated-driving-and-safety-technology (accessed December, 2017)

Turkus, B. (2014). Apple announces CarPlay in-car iPhone interface. Autoblog. Retrieved from http://www.autoblog.com/2014/03/03/apple-carplay-in-car-iphone-interface-official/ (accessed January, 2015)

Usability.gov (2016). Usability.gov – Improving the user experience. Retrieved from: https://www.usability.gov/

Usability Toolkit (2016). usability-toolkit.de: Usability für Webprojekte. Retrieved from: http://usability-toolkit.de/

Van der Laan, J., Heino, A., and De Waard, D. (1996). A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research. Part C: Emerging Technologies, 5(1), 1-10.*

Van Erp, J. (2002). Guidelines for the use of vibro-tactile display in human computer interaction. *In Proceedings of Eurohaptics 2002*. Retrieved from: http://www.eurohaptics.vision.ee.ethz.ch/2002/vanerp.pdf

Volvo (2016). V40 Cross Country, 2016, Owner's manual online: Pedestrian airbag. Retrieved from: http://support.volvocars.com/uk/cars/Pages/owners-manual.aspx?mc=y556&my=2016&sw=15w17&article=7fceb4e7544b4fbbc0a801e800b0ef6b (accessed December, 2017)

Volvo (2017). Volvo car group global newsroom – Press releases. Volvo company website. Retrieved from: https://www.media.volvocars.com/global/en-gb/media/pressreleases/list (accessed December, 2017)

VW (2017). VW Magazine. VW company website. Retrieved from: http://magazine.volkswagen.co.uk/driving-assistance-systems.html (accessed December, 2017)

Ward, N., and Parkes, A. (1994). Head-up displays and their automotive application: An overview of human factors issues affecting safety. *Accident Analysis and Prevention*, 26(6), 703-717.

Wang, Y., Toor, S., Gautam, R., and Henson, D. (2011). Blink frequency and duration during perimetry and their relationship to test–retest threshold variability. *Investigative Ophthalmology & Visual Science*, 52(2011), 4546-4550. doi:10.1167/iovs.10-6553

Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159-177. doi: 10.1080/14639220210123806

Wilson, E., Reed, C., and Braide, L. (2009). Integration of auditory and vibrotactile stimuli: Effects of phase and stimulus-onset asynchrony. *Journal of the Acoustical Society of America*. 126, 1960–1974

Wilson, F. and Stimpson, J (2010). Trends in fatalities from distracted driving in the United States, 1999 to 2008. *American Journal of Public Health,* 100(11), 2213-2219. doi: doi: 10.2105/AJPH.2009.187179

Woodward, M., Dixon, S., Hamilton, J., and Goodall, W. (2017). *Deloitte – Driving connectivity. Global automotive consumer study: Future of automotive technologies*. United Kingdom Insights, March 2017. Retrieved from: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/manufacturing/deloitte-uk-driving-connectivity.pdf

Young, K., Salmon, P., and Cornelissen, M. (2012). Distraction-induced driving error: An on-road examination of the errors made by distracted and undistracted drivers. *Accident Analysis and Prevention*, 58(2013), 218-225.