CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG

Shaping the next generation of UX in cars

Interaction design guidelines for music streaming services in cars

Master's thesis in Interaction Design & Technologies

LOVISA JÄBERG, MIKAEL MALMQVIST

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2018

Master's thesis 2018

Shaping the next generation of UX in cars

Interaction design guidelines for music streaming services in cars

LOVISA JÄBERG, MIKAEL MALMQVIST

Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology
Gothenburg, Sweden 2018

Shaping the next generation of UX in cars
Interaction design guidelines
for music streaming services in cars
LOVISA JÄBERG, MIKAEL MALMQVIST

Supervisor: Thommy Eriksson, Department of Computer Science and Engineering
Examiner: Staffan Björk, Department of Computer Science and Engineering

Master's Thesis 2018
Department of Computer Science and Engineering
Division of Interaction Design
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2018

Shaping the next generation of UX in cars
Interaction design guidelines for music streaming services in cars
LOVISA JÄBERG, MIKAEL MALMQVIST
Department of Computer Science and Engineering
Chalmers University of Technology

Abstract

The aim for this thesis has been to deliver a set of interaction design guidelines for the user experience of music streaming applications in cars. These guidelines will provide designers with concrete guidance of how to design safer interactions with the head unit in a car to reduce driver distraction and increase the overall usability.

As the goal was to deliver a set of guidelines based on insights and knowledge from design research, the design process has followed a Research through Design approach, combined with a Goal-Directed Design approach, looking at user behaviours and needs in the context of interacting with music streaming applications while driving.

The derived guidelines are motivated through an extensive literature review, user research, and user testing where the majority of the guidelines has been derived from tested hypotheses for driver safety by using a number of attention measurements, looking at driving performance, usability, visual attention and cognitive load. The remaining guidelines are motivated through already carried out research. The result of the work stress the importance of taking driver distraction into consideration when designing for in-car experiences.

These guidelines do not stand-alone guarantee safe interactions but together with already existing guidelines for in-car HCI and best practices within interaction design they provide a solid ground for decreasing driver distraction and improving the safety of the driver while using a music service in the car.

Keywords: In-car HCI, in-vehicle systems, driver distraction, interaction design.

Acknowledgements

This thesis has been carried out by two master students within the subject of interaction design at the Department of Applied Information Technology at Chalmers University of Technology. This publication only reflects the views of the authors'.

A big thank you to Anna, one of our supervisors during this master thesis, for your dedication, enthusiasm and support during our work.

Thanks to Thommy Eriksson, our other supervisor during the project, for all your help and feedback throughout the work of this thesis.

Participants in all our tests and workshops, thank you for your time and patience with us, it was a blast!

To our friends and family, you are the best and we could not have been happier to have you in our lives.

Lovisa Jäberg, Mikael Malmqvist, Gothenburg, June 2018

Glossary

Affordance "..the perceived and actual properties of the thing, primarily those fundamental properties that determine just how the thing could possibly be used." [37]. 18, 19, 48, 77

Analysis paralysis A state where users are unable to make a quick decision, due to an overload of options.. 69, 70

Continuous scrolling Scrolling that continues a while after the user has performed their swiping gesture and gradually slows down to a stop.. 68

Cover art Image representing an album, playlist or similar.. 67, 70

Discrete scrolling Scrolling that changes content discretely by changing out all visible content as a single unit and stops after one unit of content has been changed. Sometimes referred to as pagination.. 68

Eye glance duration The amount of time of a single eye glance away from the road.. 66

HCI Human-computer interaction — branch of design concerning how humans interact with computer systems. 17, 19, 29, 41, 51, 52, 65, 84

Head unit Human machine interface for interacting with in-car systems. Typically found between the driver seat and the passenger seat in cars. 17–19, 22–25, 27, 29, 41–43, 50–52, 55, 57, 65, 71

Navigational excise "Unnecessary or difficult navigation.." [13]. 47, 70

OEM Original Equipment Manufacturer. 30

Scanability How easily the user is able to visually search or find something with their gaze.. 68

Scope creep An unwanted expansion of the project's scope that often occurs when a problem statement of a project is too vaguely defined.. 47

UI User interface. 18, 22, 27–29, 31, 51, 53, 55, 56, 66–69, 71, 77, 79

UX User experience. 17, 22, 24, 30

Contents

1

Introduction

This master thesis is set within the field of in-car HCI, exploring user interactions with modern head units found in cars today. Due to the lack of concrete guidelines in current research, and an increased focus on safety regarding use of digital devices while driving, the end goal of this thesis is to deliver a set of interaction design guidelines for the UX of modern music streaming applications in cars. These guidelines focus on safe interaction with the head unit and will ultimately help designers of music streaming applications to cater for users in a safer way. That is why this master thesis has set out to answer the following research question:

What guidelines are suitable for designing interactions with music streaming applications found in head units in cars, taking driver distractions and usability into consideration?

1.1 Scope

Due to new legislations and laws of how and to what extent drivers are allowed to interact with hand held devices in cars, makers of mobile applications today have to adapt if they want their services to stay relevant in the context of use by drivers. Designers of such applications must no longer only design for a rich user experience, but also a safe one, taking visual attention and cognitive load into consideration. Therefore, the focus of this thesis has been to explore interactions between the driver and the head unit of modern cars[1]. Such interactions are limited to applications that are run either natively in the head units of the car (as in the case of Tesla's touchscreens), or on a brought-in device connected in such a way that the majority of interactions take place in the head unit (as in the case of Android Auto[2] or Apple CarPlay[3]). There has been no focus on direct interactions with hand held devices while driving.

Moreover, as research has shown that drivers spend most of their time in the car alone [1], there has been a focus on single drivers where there would be no passenger able to perform any interaction.

[1] In this thesis, a modern car is considered a car that is equipped with a built-in touch screen device between the driver seat and the passenger seat - i.e., a head unit

[2] Android Auto, https://www.android.com/auto/

[3] Apple CarPlay, https://www.apple.com/ios/carplay/

The deliverable of the project is a set of interaction design guidelines for designing a safe user experience with music streaming applications in the car. Due to this, it is out of scope for this thesis to look into aesthetic guidelines on look-and-feel of the user interface. It is also out of scope to investigate various alternative ways of interacting with the UI (e.g., playback controls on the steering wheel) due to the lack of the possibility to influence each car manufacturer and the design of in-car hardware.

1.2 Relevant stakeholders

A number of relevant stakeholder has been identified for this master thesis, and these are all described in the list below.

- Us, as the design team - Because of an interest in carrying out the work in the best possible way.
- Chalmers University of Technology - With an interest in graduating us and seeing an appropriate academic significance in our work.
- Companies of music streaming applications - companies are stakeholders that might seek to use the work of this thesis to better adapt their products to modern laws and legislation for use of apps in cars.
- Users (drivers) - The users of such applications will be affected by the work of this thesis and are thus also stakeholders needed to be taken into consideration.
- Car manufacturers - Car manufacturers will be affected by such guidelines, as they design integrated systems in their cars.
- Artists - As the suggested guidelines can affect how songs and albums are visualized, artists and labels will also be stakeholders of the work of this thesis.

1.3 Background

The car is an important part of modern society, enabling mobility to billions of people daily. As much as 87.5 % of US residents spends an average of 47.1 minutes driving every day, shown by a study released in 2016 by AAA Foundation for Traffic Safety [56], while another study (conducted in 2010) argues that Americans spend as much as 86 minutes/day in their cars, while Europeans spend 43 minutes/day in their cars [1]. Naturally, this makes development of cars and the set of features found in them very important.

Many of the features found in cars are controlled via an embedded head unit, typically found to the right of the driver, between the driver seat and the passenger seat, offering control over various in-car systems. Before the introduction of the touch screen, analog controls dominated how drivers interacted with these units. Embedded head units with features such as analog radio and climate control was interacted with by using physical knobs, dials and buttons, which offered users intuitive ways of interacting by providing tactile feedback and clear affordance. As technology has evolved, head units have become less analog and more digitized, introducing more modern means of interacting with in-car controls – an example being the digital radio and physical, context-sensitive knobs controlling both radio tuning and audio volume. Further technological advancements within the field of touch screens and HCI have enabled car manufacturers to integrate head units with all-touch interfaces in them, rendering physical controllers redundant. These interfaces provide flexibility in showing many different types of information and new types of feedback. A drawback with such interfaces is their lack of the intuitive tactile feedback and affordance found in analog controls, thus requiring more attention of the driver, which removes focus from the primary task of driving [9].

More sophisticated head units allows for activities not before found in cars, that not only assist the driver in her primary task of driving, but also introduces new services for both the driver and the passengers. Even if such services add value to the driving experience, they can potentially provide a less safe driving experience, as they add more interaction elements [57]. Examples of such would be sending text messages or streaming online music, which add value to the driving experience, but potentially increase driver distraction [46]. However, such secondary tasks must not only be ruled out as driver distractions, as they can also sharpen the attention of the driver and actually increase her alertness by decreasing driver drowsiness, fatigue and sleepiness, which can be seen in the case of listening to music while driving [8].

The possibility for drivers to perform new tasks while driving constitute new challenges for developers of in-car applications, due to the important aspect of attention [29]. Thus, interaction designers within in-car HCI need to consider both driver safety and user goals when designing applications that can be used while driving. These are novel challenges not found in traditional user interface design [10], since design of applications for embedded head units within in-car HCI is a relatively new field and differs in many ways from the design of mobile and desktop interfaces [46].

Moreover, unsurprisingly, the use of mobile devices while driving causes an increase in reaction times up to 35 % [42]. This is most likely one of the reasons why there is an increase in stricter laws and regulations against the use of hand held devices while driving [16]. Due to this, creators of mobile applications whose product fill a purpose in the context of driving, have to consider designing for embedded head units, using touch interfaces and adapt their apps accordingly, as consumers expect head units in their cars to look just like their phone [31]. Examples of such applications are Google Maps[4], Spotify[5] and WhatsApp[6]. Designers of such applications must adapt their work process and methods accordingly as methods and techniques used in typical HCI might not suit in-car HCI well at all.

This is especially important for music streaming applications, as streaming is the biggest revenue source for the music industry today and has a yearly growth rate of 41.1 % [23]. Moreover, one study showed that over 70.3 % of surveyed

[4]Google Maps - https://www.google.com/maps

[5]Spotify - https://www.spotify.com/

[6]WhatsApp - https://www.whatsapp.com/

participants preferred audio as their form of in-car entertainment [1]. This is not hard to imagine, as music listening has been an essential activity of the car driving experience since the 1920s [8] and, as of 2017, radio still dominated the in-car listening experience [36]; however, such traditional audio services found in cars are gradually being replaced by online streaming services, similar to the ones found in smartphones [14]. Moreover, it is of high interest to investigate this area, as people expect to find their music in their car [8] and the number of sold cars running Android Auto and Apple CarPlay is expected to be around 40 million vehicles respectively in 2020 [14].

2
Theory

This chapter will introduce theoretical frameworks and relevant concepts used as a base for the work of this thesis. It will thus introduce the result of the literature review carried out during the project. The chapter will also offer an insight into the current situation of in-car HCI and what research has been conducted already on the subject.

As briefly touched upon previously, drivers today tend to engage in other in-car activities than just driving. An activity, implying obvious safety risks regarding driver distraction, is using your smartphone. This is highly discouraged and in some countries directly illegal [16] [41] and the National Highway Traffic Safety Administration (NHTSA) have expressed concerns regarding use of such portable devices, while driving [33]. Due to this, it is important to provide the ability for drivers to instead of interacting with their phone, perform secondary tasks as safely as possible though built-in devices in the car.

Providing thoroughly tested guidelines for designing interaction with in-car interfaces is therefore a crucial step in moving forward within this area to make sure new systems is implemented in a safe way. Because of this, it is of interest for software companies, with products of relevance in cars, to follow such guidelines when developing devices and software to be legally used in cars [60].

2.1 In-car user interfaces

It is not hard to believe that listening to some form of audio is the most common secondary task people engage in while driving. According to Brodsky, people listen to music more often in their car than in any other daily situation [8]. As of today, increasingly more music content is consumed through the use of music streaming services [23] and by comparing such services found on smartphones to traditional radio, the selection of content to choose from is huge, interactions are more advanced and users have a greater ability to curate listening sessions according to their own preference. On the contrary, radio has the clear advantage of a simplified interaction model that is hard to beat, as controls (e.g physical knobs and buttons) are often very easy to understand and familiar to most people. This can be an advantage for an in-car use. Moreover, radio traditionally has few features, limited to switching between stations and changing of volume. For audio streaming applications on smartphone, such as Spotify, Apple Music or podcast apps, the user are presented with much more alternatives and choices, which sometimes can be overwhelming. Such feature overflow often results in users reaching for the shuffle or instant play button, to quickly put on some music [7].

Today many car manufacturers include head units with touch display in their new cars, offering more information and interaction than ever before. There are also a plethora of aftermarket head units that can be installed in cars without an already embedded head unit [19]. The appearances and functions of UIs of head units are moving towards that of smartphones. As can be seen in the case of Android Auto and CarPlay, head units are used to extend the phones functionality to the car, with all that comes with it. With screens getting bigger, (e.g the 17 inch display in the Tesla Model X[1]) this also provides for the possibility to engage the user in more activities requiring new and different content and interactions. By increasing the ability to engage in other activities than driving in the car, questions regarding safety needs to be taken into consideration. How can one ensure that drivers are not behaving in dangerous ways while interacting with applications on the head unit and how does one define what behaviour and interactions are safe?

Moreover, as will be touched upon later, as existing guidelines from some parties focus more on look-and-feel of in-car applications, defining a unified visual and navigational language, it is interesting to see how this affects the user's mental model of applications they are used to from other platforms. Most people are familiar with applications from the use of smartphones and know that the UX and appearance is shifting between different apps, but mostly stays the same across platforms (e.g switching from smartphone to desktop, or tablet). An interesting question here is "Is this also the way to go for the in-car use as well?". It is commonly known within interaction design that designers should design for the user's anticipations and what they expect. By following existing interface standards, users will quickly learn the interface and improve their productivity, as they can predict behaviours of the application based on previous experience [13]. If the user expects an application to function in a certain way on a new platform, because they are used to it from another platform, maybe it should. Instead of focusing on look-and-feel, designers of in-car application should be provided with clear interaction design guidelines of what is the safest way of interacting with a head unit while driving; however, keeping a consistent look-and-feel throughout the whole car experience might be equally as important to avoid confusion and ensure driver safety.

2.1.1 Brought-in solutions

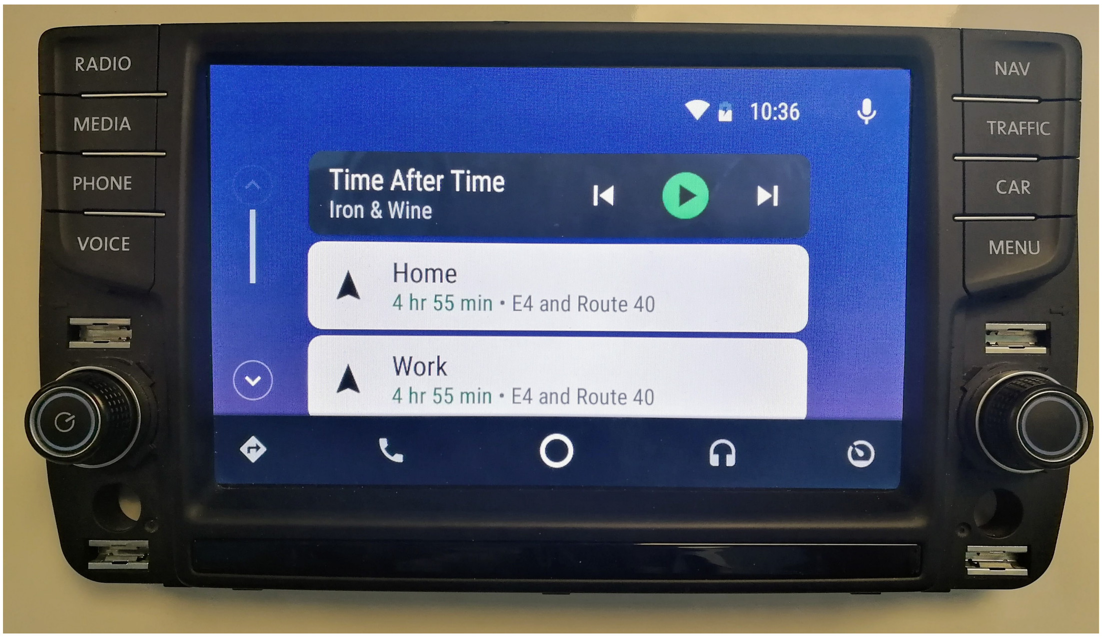
A brought-in solution refers to when the driver brings an external device into the car, which is connected to the car in such a way that interactions are done with the embedded head unit found between the driver and the passenger. The start of such an interaction can be by connecting a smartphone to the car using Bluetooth- or USB-technology. In those cases, the use of a pairing app (also referred to as a companion app) can sometimes be demanded to provide the user with an interface displayed in the head unit of the car instead of on the phone. The application is still being run on the phone, and thus using the phone's mobile data, but the majority of interactions are moved from the phone to the head

[1]Tesla Model X, https://www.tesla.com/modelx

unit [4]. Both Google and Apple have developed such software, Android Auto and CarPlay, where the in-vehicle system of the head unit is powered by Android or iOS respectively, which makes it possible to interact with your phone apps via the head unit in the car or on an aftermarket head unit [4], instead of the phone directly.

The apps supported by these softwares are often strictly limited in various ways to ensure safety and to decrease driver distraction. Writing a text message is for example completely limited to using voice interaction, both for Android Auto [2] and Apple CarPlay [5]. Drivers cannot interact with such apps via the touch interface at all; however, the feature of making a phone call can be accessed using the touch screen at all times. In both systems, a number of music, podcast and radio applications are supported and can be interacted with via the possible alternatives provided by the car and the display. Using Android Auto, Apple CarPlay or similar system gives the user a simplified user interface with large touch areas and limited interaction, as can be seen in Figure 2.1.

Figure 2.1: Home screen in Android Auto, where users are presented with personalized content in the form of recommendations of what to engage with - in this case, continuing playing music and start a route to drive home or to work.



2.1.2 Embedded solutions

Embedded solutions refer to head units with an integrated operative system that natively runs applications hence eliminating the need to connect a phone to the car. This can already be seen in cars such as the Tesla Model X, which is equipped with a 17 inch touch screen [52], where apps can be run exactly like they would on a smartphone. Embedded solutions makes it easier for users to access applications directly in the car, as they remove the needs previously mentioned for brought-in devices, simplifying interactions by removing the steps of connecting the phone and the need to have a companion app for pairing the phone with the system in the car. Depending on the maker of the embedded solution, for software companies developing applications, this means more freedom in how to design the in-car UX, but also more responsibility of making their applications safe for use while driving.

Furthermore, with more cars being connected to the internet [50] and a higher demand of car displays looking like their phones [31], embedded solutions, where the ability to run apps are built into the cars head units, will most likely increase. If the driver does not have to use data from their own data plan, this might eliminate some obstructions towards shifting from radio to other services for streaming music or other audio while driving.

2.2 Driver distraction

People spend a significant amount of their time today inside their car, which have lead to an increased demand for in-car entertainment [1]. Viewing the activity of driving as the driver's primary activity, engaging in any other activity that is not crucial for the task of safe driving can be considered as a secondary activity and thus be regarded as a driver distraction [17]. More formally, the term "distraction" will be used to describe "a specific type of inattention that occurs when drivers divert their attention away from the driving task to focus on another activity" [33].

Just as with any secondary task, interacting with an ever increasing amount of digital content requires more attention of the driver. At the same time, while providing users with what they want, developers of such systems must comply with stricter safety and reliability standards [14]. Although this raises issues of crucial factors such as road safety connected to driver distraction [1], as driver inattention to the road is one of the leading causes of car crashes [17], some research actually point towards clear benefits in engaging in some secondary tasks while driving. A benefit that has been observed is a decreased risk of crashing when carrying out a cellular conversation while driving [60]. It has also been shown that engaging in secondary activities can sharpen the attention of the driver by increasing her alertness and decreasing driver drowsiness, fatigue and sleepiness [8]. This is important to take note of, as driver drowsiness is one of the biggest causes of car crashes [14].

A study that looked into how to solve the issue of making in-car HCI safer, concluded that personalized vehicle user interfaces can make driving safer, as well as make the overall user experience more pleasant. In their testing, they found that drivers using personalized user interfaces completed test tasks quickly and efficiently [38]. However, a drawback with their study was the lack of comparison to in-car interfaces without personalization. An example of a personalized user experience outside of in-car HCI is Daily Mixes introduced by Spotify in their application, personalizing the music listening experience by presenting content based on the music taste of the user. As concluded by Normark [38], personalized user experiences is relevant for in-car HCI, not only as it is a sought-after feature in car UIs, but also as it is crucial for drivers to be able to complete their secondary tasks as quickly and efficiently as possible and it is a common behaviour among users to turn to recommended content when they do not want expend effort [7]. By providing users with a personalized experience it is possible to present the right information at the right time. An example is how Android Auto displays personalized content directly on the home screen for the driver (See Figure 2.1). This can, as mentioned, also be found in Spotify's Daily Mixes, where users are presented with relevant playlists made only for them. This creates a shift from looking for content to being presented with content, which can be essential in the context of in-car HCI. Although, designers have to be careful to not lose the group of users who actually wants to discover content by themselves and are not interested in being presented with it.

2.2.1 Measuring driver distraction

Driver attention and safety are factors that both seem difficult to measure directly. Instead, some research suggest that factors such as speed maintenance and lane keeping performance should be measured as indicators of driver distraction, when carrying out a secondary tasks while driving [29]. Other research have found that factors of visual distraction, auditory distraction, biomechanical distraction (engaging in other physical activity, not related to driving) and cognitive distraction (thinking of something not related to driving) is more important when looking for potential problems with in-car interfaces [53]. Others have found that visual factors, cognitive factors, activation factors (actively engaging in something that takes attention from driving) and anticipation factors (task related to explicit knowledge of other things, such as recalling a name of an album, causing drivers to focus on this) should be considered as most important [8].

In the extensive literature review carried out by Bach *et al.* [29], five attention measures were identified as important when looking at how secondary tasks (e.g., interacting with a head unit) affects driver distractions. These were concluded to be *primary task performance*, *secondary task performance*, *eye glance behaviour*, *physiological measures* and *subjective assessments*. Primary task performance refers to how well the user maneuvers the car measuring *lateral control*, concerning steering and how well the user stayed inside the lines, and *longitudinal control*, concerning speed maintenance. Secondary task performance concerns how well the user completes a given task requiring manipulation of an in-vehicle system, where the metrics of *task efficiency* and *task effectiveness* are studied. Eye glance behavior concerns the visual attention of the user, measuring *eye glance frequency*, *eye glance duration*, and *eye scanning pattern*. Physiological Measures refers to measures of stress and attention capacity, but also bodily measures such as heart rate and body temperature. Subjective assessments allow for test users to self reflect on their performance in completing certain given tasks. This enables discovery of the own perception of workload and attention of the participants.

2.2.2 Driver distraction guidelines

There exists some guidelines for measuring if interactions with electronic devices while driving is safe - an example being the driver distraction guidelines developed by NHTSA. These concern "effects of distraction due to drivers' use of electronic devices" [33]. NHTSA developed their guidelines in two phases, where the first phase (released in 2013) concerns driver distractions caused by devices that have been "built into a vehicle when it is manufactured" [33] and second phase (released in 2016) concerns driver distraction of aftermarket devices that are "intended to be permanently installed in the vehicle" [34], among other things. It was, however, found that the guidelines that were developed for the built-in devices during the first phase, also applies for aftermarket devices investigated during the second phase [34].

NHTSA have created a number of testing protocols that are intended to be used to ensure that a product adhere to the NHTSA guidelines. These protocols look at somewhat different factors contributing to driver distraction, have different acceptance criteria and are suitable in different testing venues [33]. The protocol within the NHTSA guidelines labeled as EGDS (Eye Glance Testing Using a Driving Simulator) is of particular interest in the work of this thesis, as research suggest a focus on eye glance behavior when measuring driver distraction. This is because insufficient perception, due to a withdrawal of attention to the primary task of driving, affects vehicle control and object and event detection negatively [29]. As the name suggest, the EGDS protocol is suitable when measuring performance by looking at eye glances in a driving simulator. In addition to the requirements of a specific driving scenario and exactly 24 test participants, this protocol contains a number of acceptance criteria that concerns the duration of eye glances of-the-road, as well as the total sum of all individual of-the-road eye glances. According to these acceptance criteria, for any test participants, 85% of individual of-the-road eye glance durations should be less than 2 seconds, while the sum of all individual of-the-road eye glance durations should not exceed 12 seconds in total. Additionally, for at least 21 of the 24 test participants, the mean of all individual of-the-road eye glances durations should be less than 2 seconds [33]. This is also supported by other research, that have found a significant increased crash risk of tasks causing eye glance durations above 2 seconds [27]. According to the guidelines, a task that fails to meet the acceptance criteria for more than 50 % of the test trials (with a sample of 24 test participants) is not suitable to be performed by the driver while driving [33].

2.3 Cognitive load

While the concept of cognitive distraction (also referred to as driver distraction) generally refers to decreased attention from the primary task of driving, cognitive load refers to the amount of cognitive resources needed to complete a task. This means that cognitive distraction can occur both in situations of high and low cognitive load [17]. As an example, cognitive distraction can occur whilst engaging in repetitive and monotonous activities causing low cognitive load, as the mind wanders, and it is not hard to imagine a situation where distraction occurs due to high cognitive load.

Due to the nature of this thesis, it is only relevant to look into situations where cognitive distraction is caused by high cognitive load, due to interaction with in-car head units. Moreover, cognitive load is an important concept within in-car HCI, as it has a negative effect on tasks directly related to primary task of driving [17].

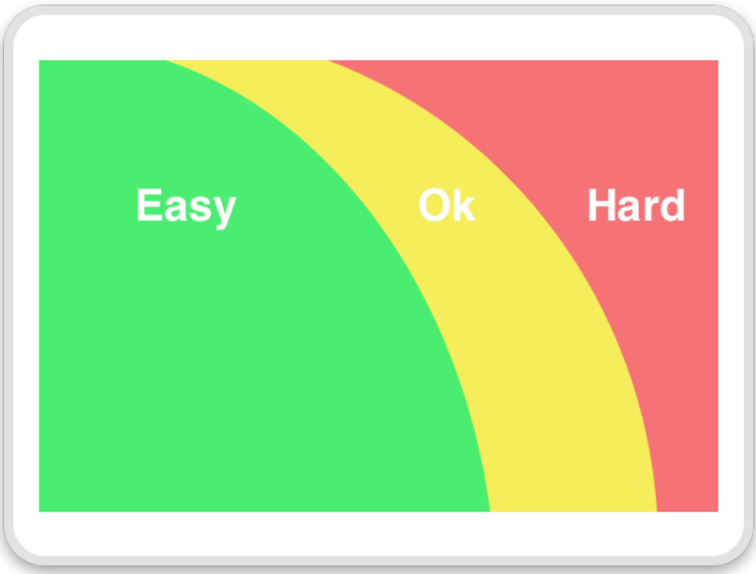
To design for such applications, transient posture is an important concept to consider. A transient application is a single-function application [13] and very applicable in the context of the car. Tasks should be carried out easily without much attention from the driver, to eliminate excise and keep the cognitive load as low as possible for the driver to be able to focus on driving. An example of what can cause high cognitive load in the context of listening to music is having too many alternatives to choose from [7].

2.4 In-car interaction models

It is crucial to avoid excessive complexity in in-car UIs as this is regarded as "..major contributors to unmanageable physical and cognitive demands for drivers." [60]. It has also been found that designers of in-car UIs must take touch zones and reachability into great consideration, as these are drastically different in head units, compared to those of handheld devices and ignoring this may add to the manual task load of drivers [60]. Therefore, thorough placement of interactive components in the UI is crucial both for reachability and viewability of the driver.

Research suggest that the touch interface of (landscape) head units can be divided into three touch zones depending on how difficult they are to reach - easy, medium and hard [60]. For a right-hand drive vehicle, it would be easy to reach to the far left of the touch screen, it would be okay to reach in the middle of the touch screen and it would be hard to reach to the far right of the touch screen (See Figure 2.2).

Figure 2.2: Touch zones of a (right-hand drive vehicle) head unit.



Other research has also shown that reachability and precision can be improved by placing interactable UI-components at edges and corners of the display [20]. However, even if designers spend more time in improving reachability of interactable UI-components, driver distraction is inevitable when looking at anything else than the road. Thus, there need to be more effort spent on designing non-visual interaction models to ensure driver safety [14]. Due to this, it is highly relevant to look into research on alternative ways of interacting with in-car interfaces. Such an area, researched by Tashev *et al.*, is voice interaction. In their research they saw that, when interacting using voice, it is possible to reduce cognitive load and driver distraction by limiting the number of keywords needed to control the interface. It has also been found that voice controlled UIs increase usability and thus decreased driver distraction [51]. This is also supported by a number of studies that have shown that the overall risk of crashing is lower when interacting using voice, compared to using visual-manual interfaces [60]. This is not hard to believe as research suggest that having a mobile conversation actually reduces the risk of a crash [60].

There has also been interesting findings that show benefits of combining voice interaction with personalization, as it increases recognition and usability for in-car user interfaces [51]. Contradictory to this other research report potential negative effects of interacting with in-car systems using voice. One study from 2015 concluded that voice control in cars caused surprisingly high levels of mental workload. Giving test users tasks such as sending and receiving texts and email, there was findings showing that voice control produced high cognitive distractions among drivers. This is especially interesting, as this study evaluated voice control using Wizard of Oz technique, providing the test users with perfect reliability. The study also conducted similar user tests using Apple's voice assistant Siri, which produced even higher cognitive distractions among drivers [49].

Another paper thoroughly compared and evaluated three other interaction models – tactile interactions with tangible controls, touch interactions and gesture interactions. In their study, they primarily look at how such interaction models affect driver distraction and driver performance. They conclude that eye glances on interfaces, where secondary tasks are carried out can be reduced if gesture interaction is used. Comparing this interaction model with touch interactions, they found that touch interaction made tasks faster to carry out, but produced more eye glances. Finally, they concluded that tactile interaction (using knobs and dials) were less intuitive and less efficient, as it took the longest to complete tasks and it did not require fewer eye glances than touch or gesture interaction. However, a benefit of using the tactile interaction model was the ability to physically being able to "scan" for the appropriate button [29].

Related to this, a study on unidirectional swipe gestures showed how the majority of test users prefer to interact using gestures. This study concluded that, for music control, there was high agreement on swipe up/down to make music louder/quieter and swipe left/right to play next/previous song. The study concluded that this most likely followed interactions transferred into the context of driving [9]. This is also suggested by other research on the topic, that concluded that it is appropriate to design gestures that users are familiar with. The same study also suggested to use audio feedback, rather than visual feedback when interacting with in-car head units [20].

2.5 Evaluating In-car user interfaces

Research on the topic of how to conduct proper user testing of in-car UIs while driving have also been reviewed. As in any field of user-centered research, evaluating entertainment systems within in-car HCI requires some form of user testing, if qualitative data shall be gathered. Conducting such evaluations while actually driving a real car has proven to be challenging, due to safety issues [1] and there are clear advantages of using a simulated driving environment instead [29]. This might not come as a surprise, but besides the clear benefit of allowing for user testing of risk-filled scenarios without endangering the test participant, a simulated environment adds a level of control of scenarios that would otherwise not be possible, and it furthermore allows for repetition of specific scenarios of interest [29].

During his study on personalized user interfaces in cars, Normark used a car simulator and reported safety and the ability to clearly see how a prototype works as being two great advantages with this approach [38]. Tashev *et al.* also worked with a simulated driving environment during their study to avoid the dangers of testing a prototype in real vehicles. In doing so they experienced benefits such as being able to record accidents, driver behaviors, and correctness of driving. They concluded deviation of distance between the car and the road lane to be the most useful parameter when measuring driver performance [51]. Another study showed that while performing user tests in a simulated driving environment, changes in driver performance were more apparent when driving along a rural road than along the motorway [17].

2.6 Existing design guidelines and best practises

There currently exist a number of design guidelines and best practises within HCI and to some extent within in-car HCI as well. As can be found in the extensive literature review of 100 papers on the subject of in-car HCI, carried out by Bach *et al* [29], safety guidelines for designers focus more on benchmark metrics regarding task completion and eye-glance behavior, than how designers should design for a safe driving experience.

Two companies that have developed guidelines for how they believe that designers should design for their respective driver-friendly in-car platforms are Google[2] and Apple[3]. It should, however, be said that while Google's guidelines for Android Auto do, Apple's guidelines for CarPlay do not mention driver distractions in any form. Even though these guidelines give clear instruction on

[2] Android Auto Design Guidelines, https://designguidelines.withgoogle.com/android-auto/

[3] Apple CarPlay Human Interface Guidelines, https://developer.apple.com/carplay/human-interface-guidelines/overview/introduction/

how to design for the respective platform, a potential problem with platform-specific guidelines are that many of the third-party applications found on the platforms look and function differently than they do on other platforms, making users unfamiliar with the experience of applications they are familiar with from other devices (such as their smartphone).

At the moment, the look-and-feel and overall UX varies a great deal between different embedded implementations by different car brands. This might not be a problem per se but it becomes an obstruction in the way that it complicates ensuring safe interaction in the context of driving. As stated by Schmidt *et al.*, the responsibility of designing in-car interfaces has been on the car manufacturers and OEMs, but is now shifting towards a more shared responsibility [46] where other software companies have to have their say on the design and implementation caused by the introduction of apps in the car's native system. Here, the use of common interaction guidelines for implementation becomes an important part.

Human Interface Guidelines - Apple CarPlay

The following guidelines has been derived from Apple, and is a subset of the Human Interface Guidelines for design of audio application on their platform CarPlay [6].

- Limit your content hierarchy to three levels or fewer.
- Use multiple tabs to organize content and ease navigation
- Show the most relevant content first
- Include single-touch playback actions at the top level of your hierarchy
- Intelligently filter content when the vehicle is moving

Android Auto Guidelines

Here follows an overview of the the Android Auto UI Guidelines, these are more thoroughly explained at the Android Auto website [3]. They do enhance the difference between designing applications for in-car usage compared to in other contexts.

- The interaction pace on Android Auto must be controlled by the driver.
- Touch targets must be larger on car interfaces so they are easier to glance at and tap.
- Appropriate color contrast helps drivers quickly interpret information and make decisions.
- Apps must adapt for nighttime driving where bright screens can be distracting.
- Roboto is used throughout the system for consistency and to help with legibility.
- Text length - Strings should not exceed 120 characters in any language.
- Touch-initiated pagination should be used to supplement swipe-controlled scrolling.
- Imagery should only be used sparingly and selectively.
- UI animation may be used sparingly to indicate transitions between 2 states.

10 Usability Heuristics for User Interface Design - Jakob Nielsen

In an article released by the Nielsen Norman Group in 1995 [35], Jakob Nielsen offers general interaction design principles for design of UIs. These principles was created as heuristics and broad rule of thumb for interface designers, and not as usability guidelines and has thus been regarded as such during the work of the thesis. A subset of these principles has been listen in the bullet list below. The subset contains the principles that are most relevant to the work of this thesis.

- Visibility of system status
- Match between system and the real world
- User control and freedom
- Consistency and standards
- Error prevention
- Recognition rather than recall
- Flexibility and efficiency of use
- Aesthetic and minimalist design

Shneiderman's Eight Golden Rules - Ben Shneiderman

In his book *Designing the User Interface: Strategies for Effective Human-Computer Interaction* [47] Ben Shneiderman's offers a number of practical guidelines applicable for designing UIs. A subset of the most relevant guidelines has been used during this master thesis and can be found below.

- Strive for consistency
- Enable frequent users to use shortcuts
- Offer informative feedback
- Design dialogues to yield closure
- Reduce short-term memory load

The guidelines that have been introduced in this section have been used throughout various parts of this thesis. They have been helpful both as rule of thumb when designing the various prototypes for testing purposes, but also served as guidance and inspiration to many of the design guidelines created during the later parts of the project. Jakob Nielsen's 10 design heuristics was especially useful, as they, as mentioned, were also used as heuristics when evaluating various design concepts for the final prototype, before testing with real test users in the driving simulator.

3
Methodology

The following chapter will offer an overview of methodology and scientific approaches relevant for the work of this master thesis. The methods will be briefly introduced and how they were actually applied along with their result will be presented in the Process chapter.

A relevant approach to design practise for the work of this thesis is Research through design, which is a scientific approach to design studies, where, through the use of methodological and theoretical approaches, researchers aim to derive a result more than just a concrete design or artifact. Thus, the goal of projects engaging in research of design is not only to deliver concrete design solutions, rather it is also to gain knowledge and understanding through design practice. As the approach centers around learning and contributing to knowledge, it is suitable in a context where researchers and designers are still practicing their skills within their respective practise [18].

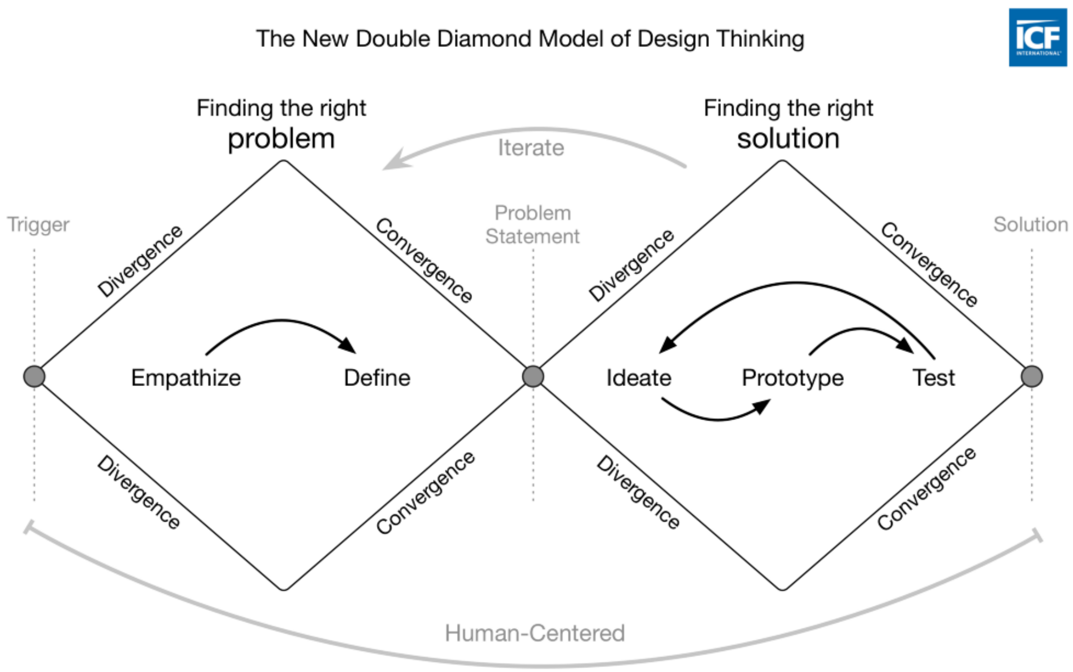
3.1 Goal-directed design

As the deliverable of this thesis is a set of interaction design guidelines for music streaming applications in cars, it made sense to engage in a design process that centered around the user, her goals, behaviors and expectations in a given context. Therefore, an iterative goal-directed design process [13] was utilized during the project. Goal-directed design is a design process created by Alan Cooper that focuses around the goal of the user and understanding their purposes and values [59]. The reason why a goal-directed approach was more suitable for this project was due to the project's focus around understanding *why* certain activities are carried out and to gain insights into the underlying values and purposes those activity have for the user. This was believed to be important in order to deliver interaction design guidelines catering for specific users goals and behaviours in the given context.

3.2 Design Thinking

Having a well structured design process and an appropriate design approach will give the project a great chance of succeeding. According to Jones [26], design can be broken down into three three stages, namely Divergence, Transformation, and Convergence, where it is important for the designer to understand and be able to go through each stage during the design process. One way of structuring the design process is by using design thinking. Design thinking is a way of creating a more non-linear and flexible process where you go through the stages of Empathize, Define, Ideate, Prototype, and Test in an iterative fashion [43]. Similar to this, The New Double Diamond Model of Design Thinking [25] (see Figure 3.1) is an approach of applying Design Thinking to a project, with five distinguishable steps, each having a specific purpose of either diverging, transforming or converging insights, as described by Jones [26].

Figure 3.1: The New Double Diamond Model of Design Thinking, as proposed by Jasper Lui of ICT International, splitting the activity of design thinking into 5 distinct iterative steps – Empathize, Define, Ideate, Prototype and Test. Image reference: Jasper Liu, Medium [25]



In both approaches, the process starts with an empathize phase, where the design team tries to understand the problem they are to solve and gain insights into their users and their goals and needs. After this step, the design team engage in a define phase where they analyze findings and define a problem statement, based on the research from the empathize phase. Using the results from this phase, an ideate phase is entered, where the design team ideate upon possible solutions to the defined problem statement. When a number of ideas have been generated and prioritized, the design team enters a prototype phase, where they preferably create inexpensive and quick-to-make prototypes based on their ideas. These prototypes are tested, either inside of the design team, or outside with a small sample of test users. Creating and testing such prototypes, the design team can get better insights into constraints of the solution, what works and what does not work. Based on insights from the prototyping phase, solutions are refined and eventually evaluated during a test phase. The purpose of this phase is to rule out issues with the design solution and ensure its usability.

The three last stages of the process (the ideate, prototype and test phase found in the second diamond of Figure 3.1) will throughout this report be referred to as the design phase of the project. By nature this design process is a non-linear one, meaning that stages can be carried out concurrently by different people of the design team and that insights from one stage may be used iteratively in another stage [43]. This can also be seen in Figure 3.1.

3.2.1 Empathize

The Empathize phase has the purpose of understanding and gaining insights about the problems to solve. This is where the design team must, through user research, identify user goals and needs, and collect data which should be further analyzed during the define phase. This phase is a crucial part of the design process where the design team must engage themselves in activities to deeply understand the issues of a problem [15]. According to the Goal-directed design approach, the empathize phase would consider a literature review as well as user research using qualitative methods, for example observations and contextual interviews [13].

3.2.1.1 Literature review

As an early part of a goal-directed design process, a literature review is suggested to be carried out. This is done to gain knowledge in the domain the research is carried out within and can be used to check against gathered user data [13].

3.2.2 Define

The define phase is where the information gathered during the empathize phase is to be analyzed and the problem statement should be clearly defined [15]. Here, the design team engage in methods to structure the data from the qualitative research and literature study. When user insights and data have methods engaging in understanding them come to use. This is the important part of bridging the Research-Design gap mentioned by Cooper [13] and the reason for a define phase is to be able to evaluate findings from the previous empathize phase in order to focus and set the aim for the upcoming ideation phase.

3.2.2.1 Affinity diagramming

Affinity diagramming is a design method used to group observations, concerns, insights or requirements from research in meaningful ways. This is usually done using individual sticky notes, allowing the team to evaluate the design implications of each idea independently. From this, the sticky notes are grouped based on affinity, creating larger overarching themes. Affinity diagramming can be done both during a usability test and for contextual inquiry. The overall idea with the method is to explore how to group contextual inquiries and issues, instead of placing notes into pre-defined groups [30].

3.2.2.2 Impact mapping and behavioural mapping

Impact mapping is a method used to concretize findings from user research. By engaging in this method, a team can group user needs and goals by mapping them to specific behaviours. Divided into categories of *Why*, *How*, and *What*. Answering the question of Why creates a shared vision and aligns the future work – "Why are we doing this work?". Answering the question of How groups all user needs and goals in order to form broader user behaviors — "How do users want to engage with the product?". Answering the question of What, presents concrete design solutions for what can enable users to reach their goals — "What design solutions would help users to reach their goals?". The needs are mapped out and clustered together to find areas of interest to reach the overall vision of the product or service [39].

3.2.2.3 Task analysis

Task analysis [30] is used to break down a user's work flow into actions and subtasks. These can then be put in a hierarchical diagram to get a good overlook of what subtasks has to be performed to complete a certain task, helping in understanding where things can go wrong and what steps might be problematic for the user.

3.2.2.4 Cognitive mapping

Cognitive mapping [30] helps to visualize the users' patterns of reasoning. It is done by mapping out interactions with a system and connecting these in a way that shows how interactions affect how the user feels and behaves. This ought to be helpful to understand what underlying problems that exists and how these impact the overall user experience.

3.2.2.5 Scenarios

Being a narrative, describing how the user engages with the product or service, scenarios [30] help to communicate how the preferred interaction between user and product or service would look like. Scenarios are created as written stories, making design concepts and interactions concrete. This helps the design team to empathize with the user and look beyond technical requirements when designing for culturally meaningful situations for day-to-day human activity.

3.2.3 Ideate

The following section will introduce design methods with the purpose of generating ideas of how to solve the problems found previously. During the ideation phase, the knowledge from previous phases should be taken into consideration and help in structuring the work. The problem statement from the Define phase should be used to focus the work of ideation [15]. Ideation is a stage of diverging where a greater amount of ideas is requested.

3.2.3.1 User journey map

User journey map [30] is an evaluative method for defining how users interact with a product. It is often ordered as a timeline that maps up a series of events that occur when the user interacts with the product. Each event is documented regarding potential opportunities, problems and feelings of the user. This allows for each moment of the interaction with a product to be evaluated and improved upon and is of great help when ideating on solutions.

3.2.3.2 Brainstorming

Brainstorming is a method for rapid idea generation that is often a semi-structured activity, performed in a group [22]. The members of the group gather to understand the problem statement of the brainstorming session and is then given time to generate a large number of ideas of how to satisfy the problem statement.

For an efficient brainstorming session, Tom Kelly has summarized some wisdom in the paper *The Perfect Brainstorm* - Approximately an hour is sufficient, there should be a clear problem statement, playful rules help, ideas should be numbered to easily reference them and to motivate the team, build upon previous ideas, make all ideas visible for the whole team, do some mental warm-up before the brainstorming session and be physical [55].

3.2.4 Prototype

Prototypes are used to try out the solutions and concertizing design ideas from previous phases, finding problems with ideas and see what works and what do not [15]. A prototype is an early version of an interface that is created for the purposes of testing critical features of a design concept. Prototypes can be differentiated based on their fidelity and a low-fidelity prototype is one that most commonly appears as paper sketches or a storyboard at an early stage of a project [30], which is useful for usability testing. Using this method, users are presented with realistic tasks to be carried out in a paper version of the concept. As the name suggests, the prototype is created using low-fidelity materials, such as pen and paper [48].

Towards the end of a project, more high- fidelity prototypes are often created. In contrast to low-fidelity prototypes, high-fidelity ones are more alike the final product feature-wise and/or in look-and-feel and thus more suitable in later phases of a project [30].

3.2.5 Test

Testing concepts is of great importance in the design process to receive feedback from users outside the design team. This helps to improve solutions and is often performed in an iterative manner where insights from testing is used to redefine the problem statement to understand the users even more and explore issues further [15]. To put a design in front of the actual user gives the designer the opportunity to identify major problems in the design [13] and should therefore not be saved until the end of a project, but be done along the way to get the best result. According to Nielsen [24] five test participants is enough for covering 80 % of usability problems of a design concept.

3.2.5.1 Formative vs summative evaluation

During a project, evaluating created prototypes can be done using formative or summative methods, depending on where you are in the process and what insights you want to gain from the evaluation.

As defined by Saettler [45], formative evaluation is used during an ongoing project to refine goals and to make sure you work towards achieving these goals. During a projects earlier stages, formative evaluation can be of great help to evaluate early design concepts or in other ways gaining information about how to continue working on the project in the best way. Here, the most important part is feedback used to be able to improve the process [12].

For this thesis, to evaluate the initial design concepts, formative evaluation methods like cognitive walkthroughs were to be used. This would help the design team in being able to derive hypotheses about intuitive interactions and how different concepts were received by the test users.

Towards the end of a project, summative evaluation can be used to get a good understanding of the projects outcome and how to continue the work, if it should be modified for further improvement or not, and if the project met the set up goals [12] [45].

For summative evaluation, the project were to use mainly user testing to evaluate the final hypotheses, where both quantitative data, measuring decided metrics of performance and eye glance behaviour, and more qualitative data, from comments of the test users and their own experience of performing the different tasks, were to be collected.

3.2.5.2 Heuristic evaluation

Heuristic evaluation is an evaluative design method used to detect usability issues of a design concept before test users are brought in to test the concept further. Using this method, a design is assessed by looking at a number of agreed-upon best practises and guidelines - i.e., a number of heuristics. This way, the design team will not take decisions based meaningful principles and not their own preferences. The evaluation does not require external test users as it is carried out by members of the design team internally, which makes later usability tests more effective, as baseline usability issues can be found and ruled out early on [30].

3.2.5.3 User testing

User testing is an activity focused on how real users perform real tasks [11]. The participant is observed while carrying out certain tasks and is often asked to verbally describe what they are doing and what they are thinking. It is an effective way of finding problems with your design and gives the designer great insights about the actual users and how they approach and carry out the tasks given to them [11].

3.2.5.4 Cognitive Walkthrough

Cognitive walkthrough is an inspection method used when evaluating the ease-of-use of a design concept and to understand if it functions in the way that the user expects it to, i.e., how well it matches the user's mental model. The method focuses around answering a set of questions for each step in an interaction sequence of interest [30].

- Will users want to produce whatever effect an action has?
- Will users see the control (button, menu, label, etc.) for the action?
- Once users find the control, will they recognize that it will produce the effect that they want?
- After the action is taken, will users understand the feedback they get, so they can confidently continue on to the next action?

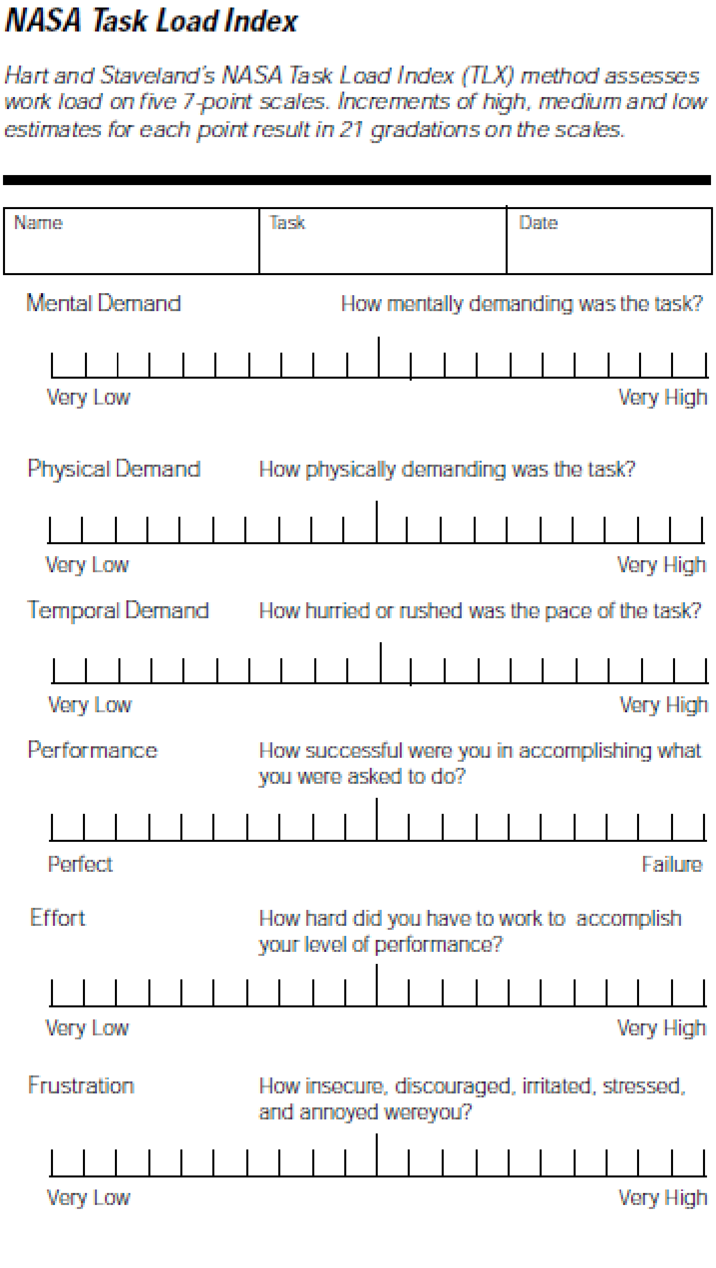
3.2.5.5 NASA-TLX

NASA task load index (NASA-TLX) is a self-assessing protocol for subjective assessment of workload or stress linked to a specific tasks, based on six individual factors - **mental demand**, **physical demand**, **temporal demand**, **performance**, **effort** and **frustration** (see Figure 3.2). It is used to gain understanding of the user's own perception of how they performed the task being evaluated. The six mentioned factors are evaluated subjectively by the user on a scale, which may or may not be weighted based on the task [21]. By summarizing results of a number of test users, it is possible to obtain an overall average workload of the different tasks, which can be valuable to compare to observed findings by the design team. NASA-TLX is a commonly used protocol for subjective assessments of in-car systems [29].

3.2.5.6 Think-aloud

Think-aloud is a method used during testing with users. When using this method, the participant verbally explain what they are thinking while carrying out a given task [30]. This method can be used during user and usability testing to understand the thought process of the participants, as well as motivations behind their interactions. The method is also good to gain insights into the participant's perception of carrying out a task instead of solemnly relying on observations by the design team.

Figure 3.2: The NASA-TLX protocol, where test participants grade the mental demand, physical demand, temporal demand, performance, effort and frustration of carrying out a specific task. Image reference: NASA



NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

4 Process

This chapter will describe and justify the execution of the methodological approaches and research methods used during each step of the project. Each phase of the project that produced a significant result will also be elaborated upon in this chapter.

As the focus of this master thesis was to deliver a set of interaction design guidelines grounded in knowledge and insights into safety and usability, and not a concrete design solution, it was found suitable to apply a research-through-design approach to the project. In doing this, the focus of the methodological approaches lay on gaining insights and knowledge around the user, her goals and behaviours in the context of interacting with head units in cars.

The overall design process for the project followed that of design thinking, as explained in the Methodology chapter. The only difference of the design process utilized during the project was how it was tweaked in order to allow for three iterations of the design phase (ideate, prototype and test), referred to as *Finding the right solution* stage in Figure 3.1, each focusing on a specific user behaviour, as elaborated upon later in this chapter. The design team (referring to the authors Jäberg, L and Malmqvist, M) also iterated upon the first two stages of the design process (The empathize and define phase) in order to identify these behaviours.

4.1 The empathize phase

The whole project started off with a pre-study, carried out during the Empathize phase, centered around user research and a literature review on the subject of in-car HCI. The goal of the pre-study was to gain an understanding of the research problem, what research had been done, and to concretize a number of user goals and needs connected to certain user behaviors found to be important when interacting with music streaming applications in cars. During this phase, user testing on existing design solutions of music streaming applications in cars was also carried out.

The user research was carried out to gain insights into what goals and needs users have while interacting with music streaming applications in cars, while the literature review focused on providing an understanding of what interaction models currently exist for head units, how they work and what benefits and risks follow such models. The literature review also provided insights into current trends in in-car HCI that helped to focus the work of the user research, as well as an understanding of current design guidelines and best practises within HCI in general. Using an empirical approach [28], the focus of the user research was on summarizing existing qualitative research on the needs and goals of people using music streaming applications. As *Do It Yourself Social Research* stress the importance for researchers to actually talk to users and that "It won't be good enough to just analyze statistical data about how many use existing services.." [58], qualitative research was also carried out in the form of user testing of one music streaming application on existing in-car platforms.

4.1.1 Exploring usability problems in existing in-car platforms

As mentioned previously, as part of the pre-study, the design team looked into existing in-car platforms by studying Android Auto and Apple CarPlay, as well as other embedded solutions, such as the head unit of Tesla Model X. To disclose what usability issues exist in current design solutions and understand how users interact with these designs, we engaged evaluation of music streaming applications on Android Auto and Apple CarPlay. Observations from these tests gave insights into what user goals could not be reached and which tasks caused unsafe user behaviors.

As user testing was done during the second iteration of the empathize and define phase, a number of tasks were set up that covered the essentials of each user behaviour found during the first iteration of the phases (more on this later). The test was carried out in a simulated environment and during the tests, participants were given tasks to complete, while driving the simulated car on a test track. While carrying out these tasks, users were asked to engage in think-aloud technique where they verbally motivated their interactions and explained their thought process throughout the whole test. At the same time, the design team observed the test participant's performance regarding the attention measures presented in the Theory chapter (*primary task performance*, *secondary task performance*, *eye glance behaviour*, *physiological measures* and *subjective measures*).

During the test, one person in the design team acted as an observer, taking notes of how well the test user performed while the other person acted as the narrator, leading the test and explaining the tasks to the test user. To complement the observations, individual interviews [58] were conducted with the test users directly after the test. During the test, to evaluate cognitive load, NASA-TLX (see Figure 3.2) [21] was used for the *subjective assessments*, as explained in the Methodology chapter. The results of the initial user testing can be found in Appendix A and B.

4.1.1.1 Test participants

In order to obtain representative insights of users of music streaming applications, test participants were selected according to the demographic of music streaming applications in general. To do this, requirements for participating included to be an active user of a music streaming application and to match the age demographic of users of music streaming applications. Also, participants were also required to have a driver's licence. In total, five participated were selected for the user testing, as this number has been found to be sufficient for covering 80 % of usability issues [24]. To get an even split between the genders, three women and two men were selected.

4.1.1.2 Simulated driving environment

For the sake of user safety, the user testing was not conducted while driving a real car, instead tests were set to a simulated driving environment. For this, the simulated environment consisted of a gaming steering wheel, gas- and break pedals, a computer screen, and a portable car head unit (see figure 4.1). The head unit both run Android Auto and Apple CarPlay, thus allowing for testing of the music music streaming applications on both of these platforms.

Figure 4.1: The simulated driving environment used for the initial user tests of existing music streaming platforms.

4.1.1.3 Tasks

The tasks that were given to the participants during the user test were a mix of exploratory and specific tasks, based on the user needs and goals of the user behaviours identified during the previous iteration of the Define stage. As mentioned, the tasks given to the participants were created with the intention of evaluating the overall usability of music streaming applications on existing in-car platforms. Intentionally, the participants were not only given tasks that were possible to complete, but also tasks that were not. This was done as it also was of interest to understand how clear current designs communicated the inability of carrying out certain actions. All given tasks can be found in the bullet list below.

- Put on some music
- Just put on some new music
- Skip 3 songs ahead
- Skip 20 seconds of this song
- What artist makes this song?
- Put on specific artist
- Go to the album of this song
- Add song to playlist
- What song is next in queue?
- Put on any other content than music
- Put on a user made playlist by you
- Put on a song you listened to recently during this test
- Switch to podcast app
- Put on a podcast
- Add this podcast to your favourite podcasts
- Put on one of your favourite podcasts
- See if you can find a podcast about cats
- Go back to the music streaming application
- Put on a specific playlist
- Save song
- Stop the music

4.1.1.4 Unsafe driving

Tasks were considered problematic if they resulted in unsafe behaviour by the driver, based on the five attention measures for driver distractions introduced in the Theory chapter - looking at *primary task performance*, *secondary task performance*, *eye glance behavior*, *physiological measures* and *subjective assessments*.

As explained, when looking at primary task performance, steering and keeping the vehicle within the lines were studied as well as speed maintenance since it has been proven that drivers reduce their speed while engaging in visually demanding secondary tasks [17]. Regarding the secondary task performance, task efficiency (how fast the users solved the task) and task effectiveness (how little errors the users had in solving the task) was observed. For eye glance behaviour, eye glance frequency and eye glance duration were studied. Studying physiological measures, indications of stress (such as swearing and rushed interactions) was determined by the design teams' observations during the user tests. As explained, for the subjective assessments, NASA-TLX was used.

4.1.2 Summarizing findings

The findings from the user tests of the existing in-car platforms was summarized and analyzed by looking at all tasks and their respective score in regards to how the test user performed according to the five attention measures. For all measures, except the subjective assessments (using the NASA-TLX), a scoring system of 1-3 was used, with 3 being the worst and 1 being best. By doing this it was possible to determine how critical the interactions connected to each task were. This was used to emphasize on what tasks were more problematic and to prioritize these problems to focus the upcoming design phase.

4.2 The define phase

The define phase of the project centered around understanding the data gathered during the empathize phase. This was done in order to define what hinders users from reaching their user goal, either by finding concrete causes of usability problems or gaps in current solutions hindering the user from being able to reach a certain goal.

4.2.1 Defining user behaviors

As goal-directed design focuses on users' goals, expectations and attitudes [13], it made sense to engage in understanding existing behaviours of users of music streaming applications. Just as mentioned in *Do It Yourself Social Research*, "To plan a really good research design, start at the beginning: with your research questions and your inquiry group and critical reference group.." [58]. Therefore, existing data on the use of music streaming platforms was analyzed by creating an impact map, identifying user needs and behaviours.

The process of impact mapping started off by a brainstorming session where the design team, along with a designer with insights into in-car HCI ideated upon the most important dimensions of the in-car user experience of listening to music through a streaming platform while driving. All ideated dimensions were grouped into a number of overarching groups and a total of four dimensions were decided upon by voting. By concretizing these dimensions into a sentence, the design team defined the core values of the project, which were used to guide the rest of the work. Focusing on these values, the design team investigated the existing user needs and goals. These needs and goals were printed on small pieces of paper which were grouped into larger themes. Having a number of larger themes, the design team were able to identify a set of user behaviors and, based on their relevance for in-car experience, four were decided to proceed with.

These user behaviors helped to guide the design work, functioning much like personas [30], but with the clear advantage of allowing the design team to focus on what was most important - how users behave in the context of driving and listening to music. Things defined in an ordinary persona (e.g., hobbies and age) was not considered as relevant as specific behaviors (e.g., behavior during stress) in the context of driving. Another advantage observed from using user behaviours instead of personas was how a person could identify with several behaviours at the same time, which would be harder when using personas, due to the extensive profile of a persona.

In total, eight different user behaviours were identified and as mentioned, four of these eight were prioritized to look further into because of their higher relevance for the context of driving. These four are explained more in detail below. In the end, only three out of these four were used in the design iterations - **The Curator**, **The Safety Conscious** and **The Casual Listener**.

Behaviour 1: The Curator

Users identifying with the behaviour of The Curator are keen on being able to curate and control their music. They want to decide for themselves what to listen to and usually have very specific requests. They are engaged in the music and are fine investing time in organizing their music and creating playlists for different occasions. This behaviour can be compared to that of the listener profile **Engaged** in *Engaging with Mobile Music Retrieval*, where users of that profile have high initial engagement in interacting their music, e,g selecting specific albums to listen to [7].

Behaviour 2: The Safety Conscious

Users of this behaviour are very conscious about safety and do not want to interact with other things while driving if not necessary. Safety comes first and they are very aware of what could be considered as unsafe while driving. If they are to perform a task it should be intuitive, easy and done in a safe way.

Behaviour 3: The Discoverer

Users considered to be in this group are interested in exploring new music and want to easily be able to find relevant content. They are not always sure of what to listen to but are curious about new and popular music and happily take suggestions from friends.

Behaviour 4: The Casual Listener

These users listen to music casually and do not care very much about what they are listening to. The Casual Listener is comparable to the listening profile **Casual** in *Engaging with Mobile Music Retrieval*, where users of this profile is content with little investment in music retrieval [7].

4.2.2 Analyzing findings

Based on the insights from the user testing, cognitive mapping was used to structure the found problems and understand them more in detail. As this method allows for more than one central keyword to be in the focus of the map, it was more suitable to use than similar methods, such as mind mapping [30], because there might exist several usability issues causing the user not to reach her goal. The cognitive mapping helped to understand what user interactions caused users to perform poorly in their primary task of driving, thus contributing to unsafe driving behavior. A number of general problem areas were identified using this method.

Moreover, task analysis complemented the cognitive map and the other summarized findings, as this method helped concretize each step the user actually had to go through when carrying out a task. Such steps included any interaction the user had to make to complete the task; however, system responses were left out. By defining this, navigational excise and problematic interactions were easy to spot, by cross-running the task analysis with the findings from the user testing.

Furthermore, based on the found behaviors from the impact mapping, 1-2 scenarios were created for each behaviour. The purpose of these were to define how a person with a specific behavior would interact with a possible future solution, helping the design team to better focus the work on designing for each behavior. These scenarios were made rather high level, leaving out specific aspects of look-and-feel.

In order to gain further insights, user journey maps were created for each of the found behaviors. These included preparation of using the music streaming service, realization of need, getting started with using the service, discovering features and content, playing content and managing the content (including repeated use, use over a long time, return to product after not using it). For each stage, a number of touch points were defined where possible interactions between the driver and the application were stated for the specific behaviour, such as "Play music" and "Add song to playlist". Furthermore, for each touch point, possible opportunities and problems were stated as well. The user journey maps where also used later during the brainstorming sessions to help focus them.

4.3 Developing concepts

With the problem definition from the define phase in mind, the design phase was entered, with purpose of ideating, prototyping and testing a number of concrete design solutions satisfying the needs of each found user behaviour. The aim of this phase was to ultimately derive a number of hypotheses regarding how to design interactions for music streaming applications. These hypotheses would later be tested and compared to existing research and guidelines to eventually be turned into the final interaction design guidelines.

The design stage (referred to as the *Finding the right solution*-stage in Figure 3.1), included an ideate, prototype and test phase and was iterated upon three times for each of the three prioritized user behaviours - **The Curator**, **The Safety Conscious** and **The Casual Listener**.

4.3.1 Ideate

The purpose of the ideation phase was to ideate upon concepts that would satisfy the needs of the specific user behaviour focused on during respective iteration. This helped to narrow the scope of ideas and, not having to consider the other behaviours or user needs, helped to avoid scope creep.

4.3.1.1 Brainstorming workshops

For each of the prioritized user behaviours, a brainstorming workshop was held. During these workshops, external participants were invited, the majority of which were experienced designers, to ideate upon possible solutions to the needs of the behaviour that was focused on during that iteration of the design phase. The group of participants were presented with an initial problem statement, consisting of the user behaviour to focus on, their needs and their user journey. To start with, an individual brainstorming session was held, followed by a group discussion of all ideas, voting and prioritizing of the ideas. After this, yet another individual brainstorming session was performed, giving the participants room to improve the top ideas further by sketching or developing them more. In order to get as much out of the brainstorming session as possible, inspiration was drawn from Seven Secrets for Better Brainstorming [55], where each brainstorming session had a clear focus, playful rules (where ideas were not criticized) and numbering of ideas. Mobile phones were also banned during these workshops to help creative thinking by avoiding external distractions.

4.3.1.2 Prioritizing ideas

After each brainstorming session the design team engaged in affinity diagramming to be able to organize all generated ideas. During this process, the top voted ideas were organized into larger overarching groups based on what user need they satisfied. Based on their relevance to in-car usage, the ideas were prioritized, and connected to findings from the literature study.

4.3.2 Prototype

With a number of ideas of how to satisfy the needs for each user behavior, the design team entered the prototyping phase. During this phase, the design team diverged the ideas into a set of design concepts of varying fidelity that were iterated upon and refined engaging in formative evaluation to be able to further improve the concepts.

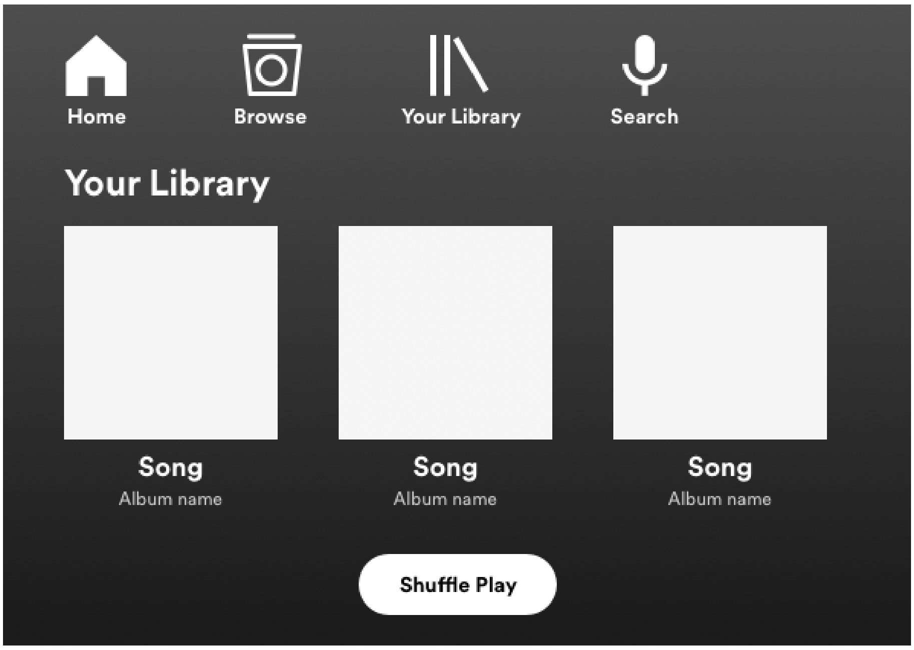
During the first iteration of the design phase the focus was set on **The Curator**. A large number of low-fidelity paper prototypes was created (see Figure 4.2) based on the set of larger overarching groups defined during the affinity diagramming process.

During the second and third iteration of the design phase, when the design team focused on **The Safety Conscious** and **The Casual Listener** respectively, a set of downscaled, digital hi-fidelity prototypes was created using the design toolkit Sketch[1] (see Figure 4.3). The decision to create digital, hi-fidelity prototypes instead of paper-and-pen prototypes during the second and third iteration was made due to the considerable confusion that arouse among test users during the cognitive walkthroughs of the first iteration, based on varying look-and-feel and inherit (sometimes incorrect) affordance of the material properties of paper and the large set of prototypes.

[1] Sketch, https://www.sketchapp.com/

Figure 4.2: One of the low-fidelity paper prototypes that was created during the first iretation of the design phase.

Figure 4.3: One of the high-fidelity paper prototypes that was created during the third iteration of the design phase.



4.3.3 Test

The goal of testing the prototypes was to evaluate the ideas generated during the ideate phase and converge them into a smaller set of more concrete design solutions. These would be used to derive hypotheses regarding how to design safe interactions between drivers and music streaming platforms on in-car head units, which eventually would be turned into concrete interaction design guidelines. The prototyping phase was iterated upon after going through a set of tests, to refine the prototypes and test again. The evaluation was done by carrying out cognitive walkthroughs of the prototypes.

4.3.3.1 Cognitive walkthroughs

Cognitive walkthroughs were carried out to evaluate the large amount of design concepts and draw conclusions regarding usability and user expectations. The purpose of these walkthroughs were to both get an understanding of what generally seemed most intuitive for users and to some degree filter out concepts that did not match the mental model of users very well.

Before hosting the cognitive walkthroughs, a pilot session was held on one test user, as suggested by Rubin [44]. This allowed the design team to test out the method and find flaws in questions and the overall execution of the walkthrough. After the pilot session, all questions were revisited and refined based on how hard they were to understand and similar questions were reworked or removed to avoid answers that was clearly influenced by the thought process and, to some degree, the outcome of previous questions.

Finally, with the goal of being able to answer the four questions of Cognitive Walkthroughs (introduced in the Methodology chapter), a number of questions were formulated for the test users to answer. While most questions concerned the users' expectations regarding the behaviour and actions of certain UI elements, some concerned navigation to certain parts of the UI, while others concerned the users' preferences of placement of certain UI elements.

For the whole cognitive walkthrough, one member of the design team acted as a narrator, interacting with the test user, while the other member engaged in observing and note-taking. As stressed in *Do It Yourself Social Research* [58], it was very important that everything that was said during these sessions was documented, as the results of these tests was essential for the end goal of the project when drawing conclusions and to motivate research decisions.

4.3.3.2 Test participants

In total, for the three iterations, 16 cognitive walkthroughs were held on the various design concepts that was evaluated, where the test participants represented an equal split between the genders.

To cover most usability aspects and users expectations regarding interactions, a wide range of test users were recruited, consisting both of domain experts (namely people with work experience within in-car streaming applications) and people outside of research and development completely. To match the demographic of users of music streaming applications, the age of the test participant matched that of the users of music streaming applications.

4.3.3.3 Overall findings

As suggested in *Do It Yourself Social Research* [58], it is important to not save the reflection of the process until the very end of the research process and therefore, as a part of every iteration, documentation and summarizing of all findings were done. This allowed for reflection upon the result and methods engaged in, which allowed the design team to improve their process for upcoming iterations.

Results from the cognitive walkthroughs indicated that consistency between car UI, smartphone UI, and desktop UI for an application is very important - users very often motivated their expectations based on how applications worked on other devices.

Users reacted differently when familiar with the content and tended to be okay with being shown more content at the same time, compared to unfamiliar content which they did not want to see a lot of at the same time. Many best practices within interaction design were confirmed, such as that hinting is important and should be made clear and consistent. Compared to UI outside of the car, hinting is even more important because if done in the wrong way it will increase the cognitive load of the driver. It was also seen that hierarchical levels should be used to as small extent as possible, as it both adds to more interactions, but also adds to the users' cognitive load, as they need to keep track of where they are in the hierarchies.

4.3.4 Summarizing hypotheses

The summarized insights from the cognitive walkthroughs, along with results from related research and already existing design guidelines and best practises within HCI, were used to form a number of hypotheses regarding how to design interactions for head units in cars. These hypotheses can be seen as early drafts of what would be the final interaction design guidelines, but at this stage the hypotheses had not been tested for visual attention, cognitive load and usability - i.e., how they affect driver safety. It should also be said that some hypotheses even contradict one another. Below follows a bullet list of the hypotheses derived from the design phase.

- Use grids to present content in an easy understandable way
- Use discrete scrolling, not continuous
- Use horizontal scrolling, not vertical
- If the application has an established navigational model, use this, as users expects consistency over platforms
- If the application has an established visual language, use this, as users expects consistency over platforms
- Place playback controls at the bottom of the UI
- Clearly show the currently playing song for the user
- Always show the most wanted and used playback controls
- In a Now-Playing view, center all content.
- Display most relevant content first
- Hint about possible interactions and the outcome of them
- Keep all content on a top level - limit the use of hierarchies in the UI

- Allow users to continue their music from their last sessions, allowing for a seamless experience over platforms, few and simple choices and get on with their driving as quickly as possible
- Allow users control core functions using big swipe gestures, thus removing the need of buttons that require higher precision
- Always show playback controls
- Hide unavailable features when it is unsafe to interact with them
- Don't display all available content, only show a subset of around 6 top items
- Avoid scrolling to as large extent as possible, only show as much content as can fit on one screen
- Adapt content based on context/length drive - some content may be more suitable for longer drives (e.g., podcasts)
- Autostart music from last session after a number of seconds - but allow users to start directly or cancel easily
- Surface new releases of artists the user like

4.4 Turning hypotheses into guidelines

During this last phase of the project, the design team turned the found hypotheses into concrete interaction design guidelines. This was done by comparing said hypotheses with already existing research and by evaluating them for safety in various design concepts, using the five attention measures mentioned earlier.

4.4.1 Prioritizing hypotheses

Due to the vast number of hypotheses, there was a need to prioritize these to focus the work on what was most important and what needed especially to be tested for safety. To do this, the design team arranged discussions with a number of internal stakeholders and designers with insights into in-car HCI. The end goal of the prioritizing of the hypotheses was to understand which hypotheses could be turned into guidelines directly, having enough support from literature and existing research or the cognitive walkthroughs and which needed further testing by the design team.

4.4.2 Prototyping

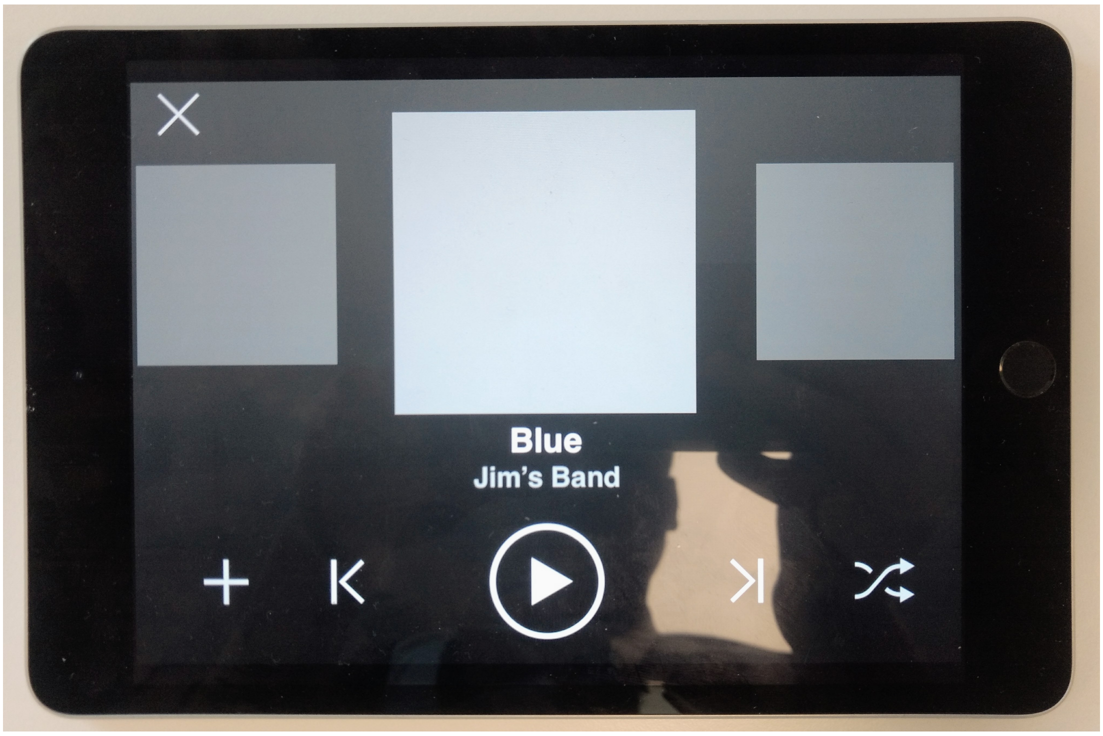
A prototype following the suggested hypotheses was created to be ale to carry out realistic testing. It was designed to cover the tasks to be tested and did therefore not include all possible functionality of a music application. It had several tabs that the user could navigate through, consisting of a home screen, a library (gathering Albums, Artists, Playlists, etc), a search screen and screen for the currently playing song (see Figure 4.10). The interactive prototype was created using the applications Sketch and Principle[2]. An iPad mini placed in landscape in the car simulator was used to simulate the head unit (see Figure 4.4) during the test. Before engaging in user testing with external test participants,

[2]Principle - http://principleformac.com/

the prototype was internally evaluated using heuristic evaluation by following **10 Usability Heuristics for User Interface Design** by Nielsen [35] as described in the Methodology chapter. Internal user testing was also done, where the design team by themselves carried out the different tasks in the created prototype.

To try out different design solutions that sometimes followed contradictory hypotheses, shortcuts were added to the prototype that could be used to switch between design solutions (this can be seen as a colored box in the UI in Figure 4.10).

Figure 4.4: Now Playing View of the prototype created using Sketch and Principle on an iPad mini in landscape mode.



4.4.3 Verifying hypotheses through user testing

To turn the hypotheses into guidelines, user testing was carried out. The testing centered around evaluating the attention measures of *primary task performance*, *secondary task performance*, *eye glance behaviour*, *physiological measures* and *subjective assessments*, as described in the Theory chapter. This was done by using an interactive high-fidelity prototype in a simulated driving environment, where test users were given appropriate tasks that would test the hypotheses.

4.4.3.1 Evaluating attention measures

For each task, the attention measures presented in the Theory chapter were evaluated. Tasks were evaluated by looking at the the average values of the *primary task performance*, *secondary task performance*, *physiological measures*, *subjective* *assessments* and some metrics of the *eye glance behaviour*. The results from the *subjective assessments* using the NASA-TLX can be found in Appendix C, the results of the *primary task performance*, *secondary task performance* and *physiological measures*, as well as comments from users can be found in Appendix D, while results of the *eye glance behavior* can be found in Appendix E. The testing process has been summarized below and visualized in Figure 4.5.

Primary task performance, concerning how well the user drove, was measured by looking at the users lane keeping and speed maintenance abilities, while carrying out a task. A scale of 1-3 was used to determine the user's overall primary task performance, where 1 indicated a good performance, while 3 indicated a bad performance.

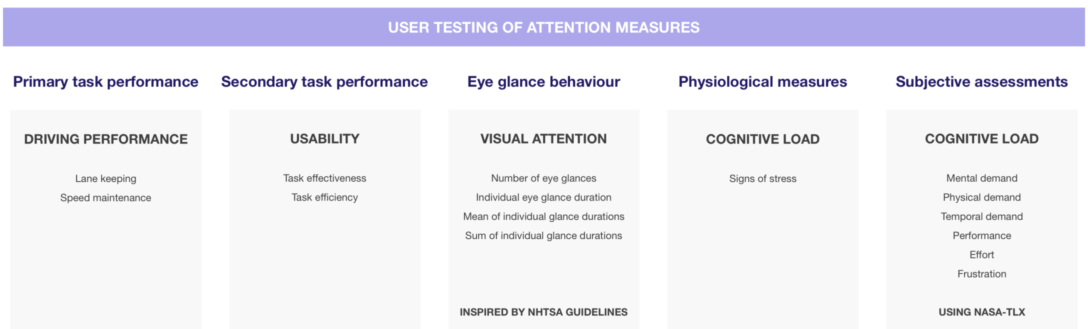
Secondary task performance, concerning the overall usability of a task, was measured by observing task effectiveness and task efficiency - i.e., how easily the task was carried out. Here, the same scale was used, as for the primary task performance.

Eye glance behaviour concerns the driver's visual attention while carrying out a task. Studying this measure, the number of glances, individual glance duration, mean of individual glance durations and the sum of individual glance durations were observed. These metrics were inspired by the driver distraction guidelines for aftermarket devices and built-in devices by NHTSA. More specifically, the Eye Glance Testing Using a Driving Simulator (EGDS) protocol as explained in the Theory chapter, was used as inspiration, but not followed. This was decided as the specific driving scenario required for this protocol could not be reproduced with the equipment available and it was not feasible for the scope of this thesis to carry out 24 user tests, as required in the protocol [33].

Physiological measures, concerning the cognitive load of the user, were measured using the same scale as for the primary and secondary task performance. This was measured by looking for signs of stress or frustration of the user.

Subjective assessments, also concerning the cognitive load of the user, were measured by using the subjective assessment protocol NASA-TLX, as explaied in the Theory chapter.

Figure 4.5: Process of user testing of attention measures.



4.4.3.2 Simulated driving environment

The simulated driving environment used for the user testing of the attention measures can be seen in Figure 4.6, 4.7, 4.8 and 4.9. It consisted of a realistic steering wheel, rear-view mirrors, a gas- and a brake pedal, a head unit simulated with an 7.9 inch iPad mini running the interactive high-fidelity prototype, three high-resolution screens covering a 180 degree angle of the driver's vision, and two cameras for recording interactions. The user was sitting in the driver's seat inside the front part of a car, creating a more realistic feeling for them. The test participant could only drive forward and was not able use the gear, but could accelerate and brake. The driving view presented to the test participant was projected to the three screens put up in front of the car, being able to show different driving scenarios.

The driving scenario used during the tests was a country road included other vehicles, and some common situations you could encounter while driving, like a tailback or standstill vehicles on the road. The simulated driving environment allowed for synced, real-time video recording from multiple sources of the whole test, as well as data recordings of desired metrics, such as *lateral control* (lane keeping) and *longitudinal control* (speed maintenance). Two camera sources were used during the tests to capture video footage from inside of the "car". One of the cameras recorded video from in-front of the driver, showing their face to more accurately determine *eye glance behaviour* and *physiological measures*. The other camera recorded video from the back, showing the UI of the prototype, as well as the road, which helped to evaluate *primary task performance* and *secondary task performance*.

4.4.3.3 Tasks

When deciding upon what task to give the participants of the user tests, the focus was set on interactions that would test the hypotheses. The tasks were created with the found user behaviours in mind, as it was important to ensure representation of these in the user tests. Some behaviours though, like The Safety Conscious where the main focus was to have as little interaction with the UI as possible, were harder to create tasks around but were more used for evaluating the designs. Some tasks were also harder to test in a simulated environment, and would need to be tested in more day-to-day situations. The tasks given to the test participants to carry out, along with their results, can be found in Appendix D and E.

4.4.3.4 Selecting test participants

Since all behaviours were represented in the tasks, it was decided that no screening of test participants, regarding what user behaviour they identified with, would be needed. It would not matter if a user identified with one behaviour or the other when they were given a predetermined task of a specific behaviour.

Eight people were recruited for the user tests. Just as in the early user tests, all test participants were selected according to the demographic of users of music streaming applications. Of these eight people, one participant was used for the

Figure 4.6: The simulated driving environment used during the user testing.

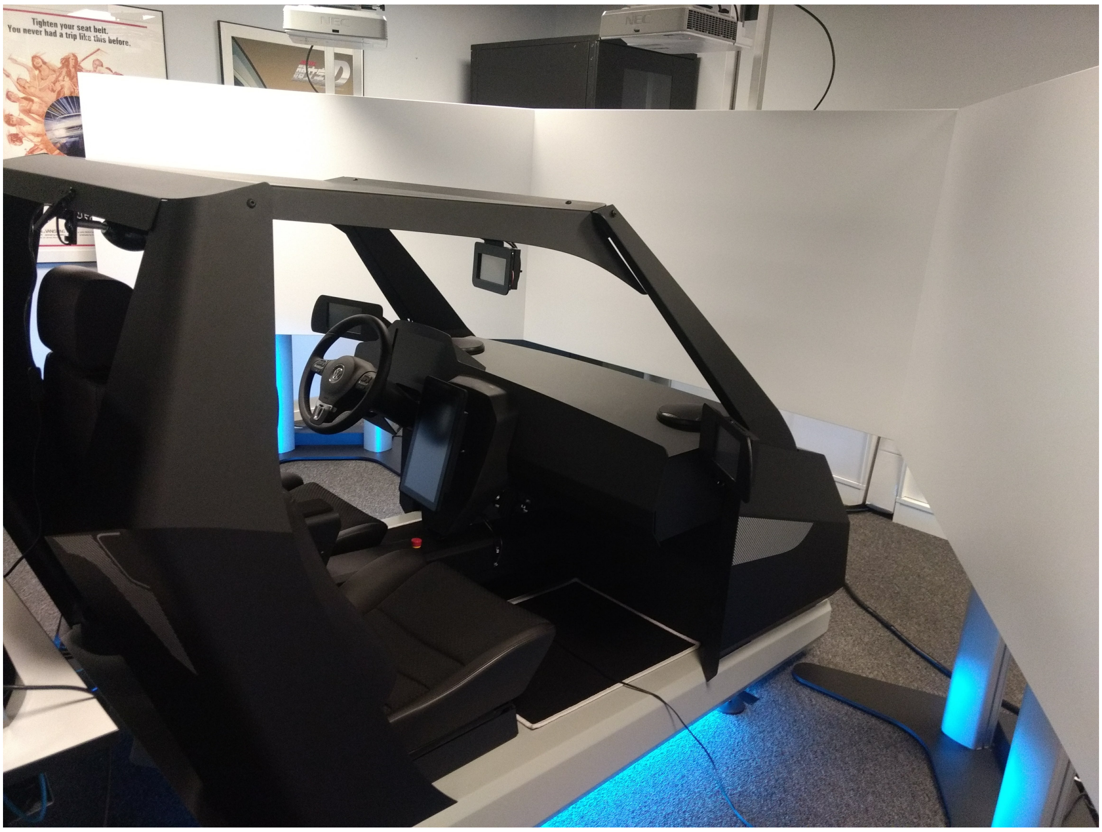
pilot study and not included in the final result. Another participant had to call off the test due to nausea and did therefore not give a complete test result. As motivated earlier, five participants would have been enough; however, to make sure getting enough complete test results, three more participants were recruited.

4.4.3.5 Carrying out user tests using a simulator

As suggested by Rubin [44], a pilot session was held at the beginning of the user tests. This proved useful as some questions needed slight rephrasing, while some were removed completely because they did not add value to the overall test. During the test, users engaged in think-aloud technique and, as the prototype were fully interactive, very little hands-on help was needed from the design team. However, some parts of the prototype had to be manually simulated by the member of the design team that acted as the narrator - e.g., simulating a voice search feature in the UI as well as switching views between different design alternatives.

Although the same attention measures as during the initial user tests was observed, due to an upgrade in equipment, it was now possible to record metrics and evaluate them more precisely. The *primary task performance* was recorded directly through data output from the simulator, while *secondary task performance*, *eye glance behavior* and *physiological measures* was recorded using two cameras, one capturing the face of the test participant, while the other captured

Figure 4.7: The simulated driving environment used during the user testing.



their interactions with the head unit and the road. The *subjective assessments* was also improved upon compared to during the initial user tests, as the NASA-TLX (see Figure 3.2) [21] was now instead verbally carried out after each task during the test instead of in writing after the entire test. In carrying out the NASA-TLX, users were asked to reflect on their performance by referring to a visual scale placed in front of them (see Figure 4.9), which solved the problem found during the initial user tests, where test participants had to remember their performance after completing the whole test.

Figure 4.8: Test participant inside the simulated driving environment, during a user test.

Figure 4.9: Close up of the simulated driving environment used for the user tests. Here, the actual prototype on the iPad mini can be seen, as well as the visual scale on which test participants subjectively graded the *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort* and *frustration level* of carrying out a task.

Figure 4.10: Prototype following design hypotheses - View of all artists with a visible (green) switch to manually change layout of artists in purposes of comparing design concepts.

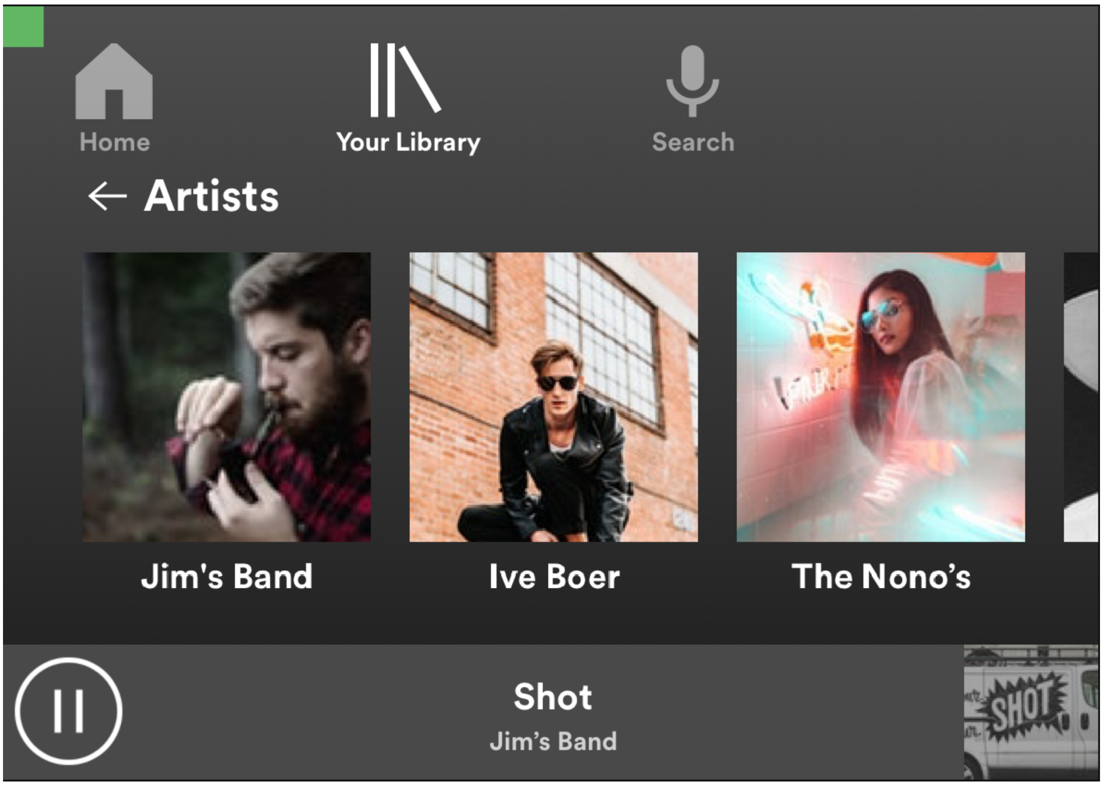


Figure 4.11: Prototype following design hypotheses - Artist page with a horizontally list of image and text content combined.

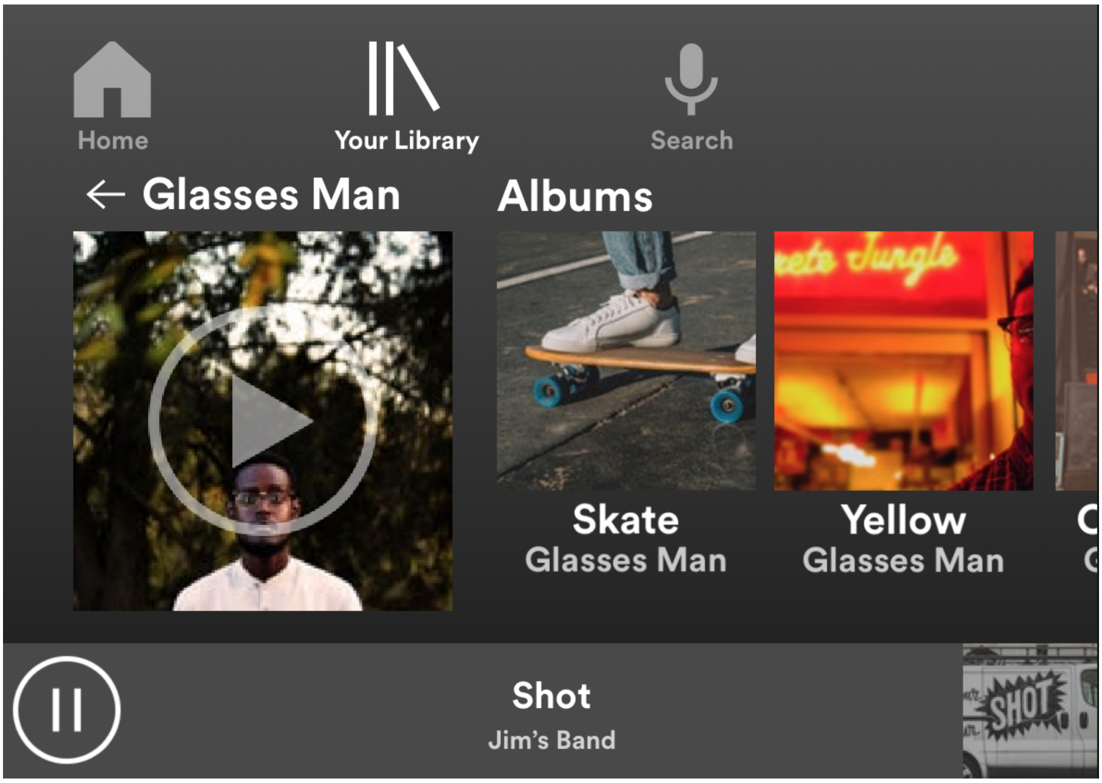


Figure 4.12: Prototype following design hypotheses - Album page with a vertical list of text content, only displaying top 3 songs of the album.

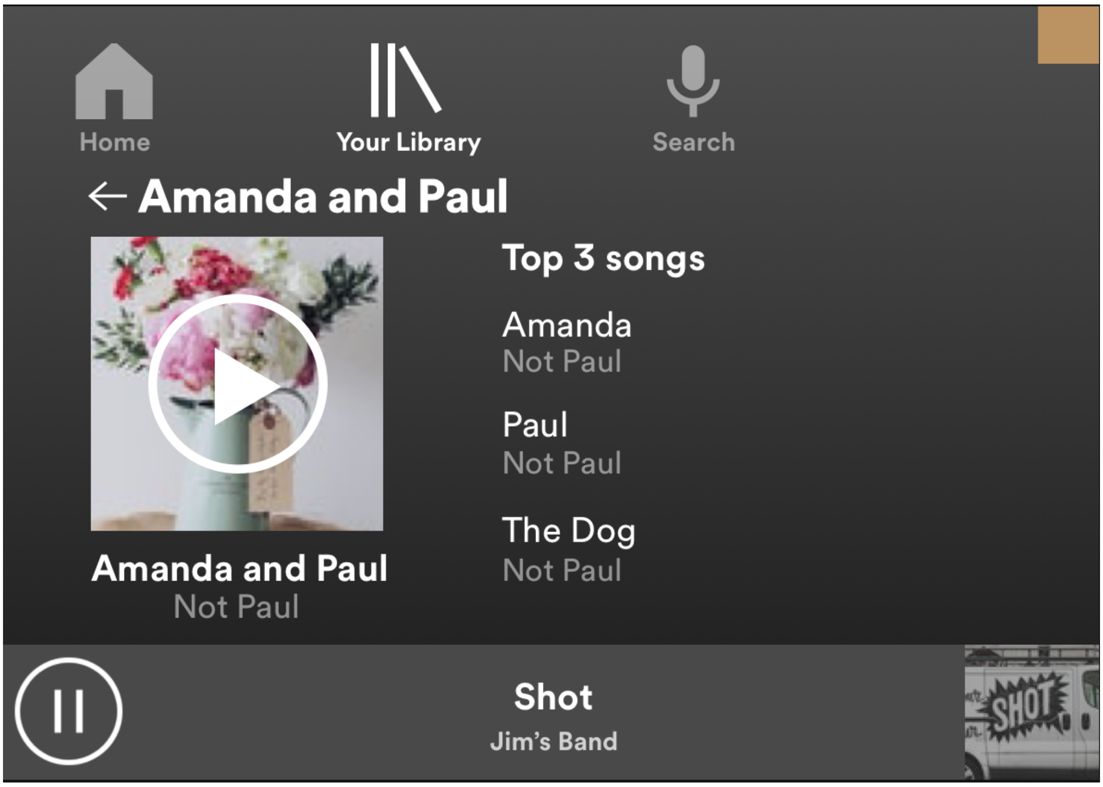
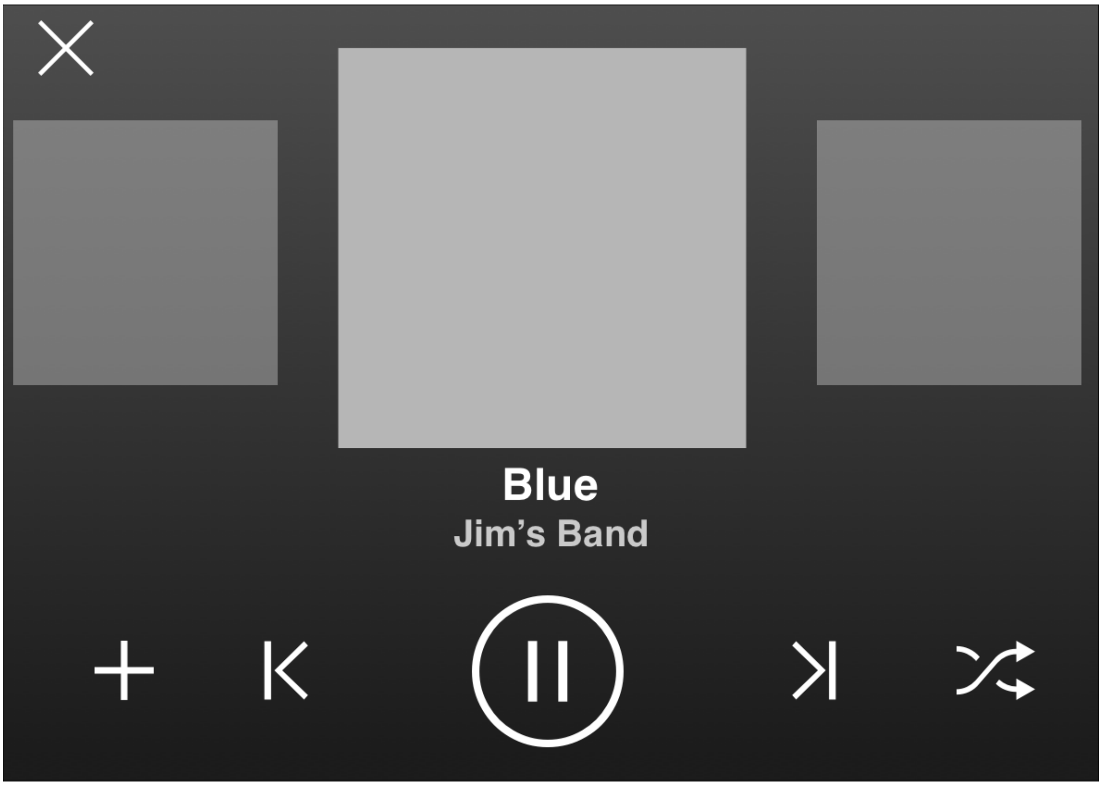


Figure 4.13: Prototype following design hypotheses - Currently playing song view.



5
Results

In the following chapter, the results of this master thesis are presented. As the end goal was to answer the research question of *what guidelines are suitable for designing interactions with music applications found in head units in cars, taking driver distraction and usability into consideration*, these suggested guidelines will be introduced and motivated in this chapter, presented in an order that does not represent priority. The suggested guidelines will act as a complement to general interaction design guidelines and best practises, such as those presented in the Theory chapter.

Comments from the users used in the results were collected during the cognitive walkthroughs and the user testing in the simulated driving environment, see Appendix D.

5.1 Interaction design guidelines for music applications in cars

In this section follows a number of interaction design guidelines for designing interactions with music applications in head units in cars. The design guidelines have been derived from the testing of a number of hypotheses based on literature reviews, user research, as well as a thorough design process evaluating different concepts, as explained in the Process chapter.

The majority of the guidelines have also been evaluated regarding safety while driving, based on five attention measures for measuring driver distraction introduced in the Theory chapter and seen in Figure 4.5 - *primary task performance*, *secondary task performance*, *eye glance behavior* (inspired by the NHTSA guidelines), *physiological measures* and *subjective assessments* (using NASA-TLX). All results of these metrics can be found in Appendix C, D and E.

A number of guidelines rely solely on previous research and results from cognitive walkthroughs, as relevant research has been done on the topic of in-car HCI before this thesis and it would be unwise not to draw inspiration and use insights already gathered from this. When referring to cognitive walkthroughs and user testing in the rest of this chapter, the authors refer to the evaluation methods and tests being carried out by the design team themselves previously during the project, see Process Chapter.

It needs to be said that, even though the results of this master thesis concerns how designers can provide users with safer in-car experiences while interacting with head units in cars, it has been seen that in order to achieve this, interactions need to be limited to as big extent as possible. In line with this, the safest in-car experience is therefore one where the driver simply do not interact with electronic devices, such as head units, at all.

Use vertical list for content that requires users to read

When presenting content for the user that requires them to read text, (e.g., when content can not easily be distinguishable using images or icons) providing them with content in a vertical list is to be preferred. The results of the user tests show that users perform better using vertical lists, compared to using horizontal lists, in tasks that require them to read any form of text (see Appendix C, D and E).

Users felt that they got a better overview of content when being presented with a vertical list and motivated this by being able to read multiple titles at the same time, which was reflected in the results of the tests, due to increased speed of finding specific content using a vertical list.

Looking at *eye glance behaviour*, comparing vertical and horizontal lists, users spend less number of eye glances and shorter total eye glance duration on average when interacting with vertical lists, compared to horizontal ones. This can also be seen when studying the *primary task performance*, *secondary task performance* and *physiological measures.* In these observed metrics, vertical lists outperformed horizontal ones in all cases. The same can be seen when evaluating the *subjective assessments* using NASA-TLX.

Limit hierarchies to 3 levels or less

When ordering content in hierarchies (e.g., Your Library > Artists > Albums > Songs), limit the number of hierarchies to 3 levels or less, as fewer steps means fewer interactions, which ultimately means lower driver distraction.

It has been found in user testing that using more than 3 levels in a hierarchy requires the user to look away from the road for longer time than what is acceptable according to the NHTSA guidelines, and therefore increases the level of driver distraction. This is also supported by Apple's *Human Interface Guidelines* for developing applications on the CarPlay platform - "Limit your content hierarchy to three levels or fewer" [6].

Place playback controls at the bottom of the UI

Allow users to easily access the most used playback controls by placing them in the bottom of the UI. In cognitive walkthroughs and user testing, it has been seen that users expect consistency across platforms. This is also supported by Nielsen's *10 Usability Heuristics for User Interface Design*, as "Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions" [35] and the bottom of the UI is therefore an intuitive placement of the playback controls, based on their placement on other platforms. By doing this, they are easier to reach than if they were placed in other areas of the screen [60] and in user tests, tasks concerning interaction with these controls performed well (see Appendix C and Appendix D).

Hide unavailable features, if there are any

Avoid confusing the user of what features and content are and are not available, by hiding unavailable features and content. Cognitive walkthroughs showed that users were confused by unavailable features that are still visible in the UI, but slightly "grayed out" (as some believed the UI to just dim down). In the case of interacting with in-car UI, it was seen that users prefer not to see such features at all. In the user testing, it was also seen that when browsing for content, users perform better when no unavailable content were present even though this is probably mainly a result of the lists being shorter when unavailable content were taken away.

Support voice interaction for search of specific content

By allowing users to search for specific content using voice interaction, crash risks and driver distractions can be reduced [60] [51]. This is also something that is supported by the research of Tashev *et al.*, where they found it possible to reduce cognitive load and driver distraction by using a limited number of keywords to control the interface by voice [51]. Searching for content using voice allows for less browsing for content and in doing this, unsafe interactions (e.g., scrolling in long lists and reading texts) can be limited. Both during cognitive walkthroughs and user testing, users expressed a desire to use voice control when searching for specific content.

During the user testing, voice search was tested in two different ways, one where pressing a voice search icon (which can be seen in Figure 4.10), which directly made the system start listening for voice commands, and one where the user after pressing the search icon had to perform an extra step to activate the voice command. The overall observations indicated that voice search directly activated performed well on driving and task performance. Therefore, it would be preferred to offer voice search using one action to start listening to voice commands. Users wanted this action to either be an actual button, on the steering wheel or in the interface, or a voice command.

Always show the currently playing song

This could for example be displayed as a horizontal bar in the bottom of the UI with the title of the playing track, the artist and the cover art (see Figure 4.11). By doing this, users are informed regarding where music is coming from, its current playback state (played or paused) and what song is playing. Cognitive walkthroughs showed an expectation from users to always be able to see these things, as they are used to this on other platforms. This goes in line with Nielsen guidelines, stating that "The system should always keep users informed about what is going on.." [35]. Unsurprisingly, user testing proved it to be safer to retrieve information regarding the currently playing song, when the currently playing song is always visible.

Always accessible playback controls

Allow users to easily interact with playback controls by having them quickly accessible at all times. In most cars, the steering wheel provide this already, as it has basic controls for playback, such as skip song and play/pause song. During cognitive walkthroughs, users expressed a desire to be able to access playback controls quickly and easily in the UI and in the user testing, tasks regarding pausing the currently playing songs performed well, when always showing playback controls at the top level of the interface. This can be related to Apple's guidelines for CarPlay, recommending designers to "Include single-touch playback actions at the top level of your hierarchy." [6].

Allow users to control core functionalities using big swipe gestures

Controlling core functionalities[1] should be as easy and intuitive for the user as possible. The ability to perform tasks without having to place the finger at a specific place in the interface allows for users to be less precise in their interaction with the interface and thus makes tasks easier to perform.

According to the results of the NASA-TLX (see Appendix C), when using big swipe gestures to change song, users performed better than when clicking. Users explained that swiping was easier because this could be done with less precision and they did not have to look at the UI as much as when having to click buttons for example. This goes in line with the research carried out by Burnett *et al.*, where they found that users prefer interacting with in-car touch screens using swipe gestures and that controlling music playback using swipe comes naturally for users and it has most likely transferred into the driving environment from other contexts [9]. Even though the number of eye glances were more when users changed song using swipes, compared to clicking, the average sum of all eye glances were lower and the average individual eye glances were shorter (see Appendix D). This indicates a lower driver distraction when using swipe to skip song, compared to clicking.

Use discrete scrolling, not continuous

If scrolling is present in the interface, provide the user with discrete scrolling, rather than continuous scrolling. This makes the action of scrolling easier, as it has been seen to improve usability and scanability of the UI in the user testing (this is essential to consider when providing users with scroll [13]). It has also been shown that pagination (i.e., discrete scrolling) is suitable for long lists of similar elements [13].

During user testing, most users thought discrete scrolling was faster and that continuous scrolling made them have to keep their finger on the screen more, implying taking the hand off the steering wheel. Continuous scrolling required significantly more long eye glances than discrete scrolling and performed worse

[1] What is a core functionality depends on the application

according to both observations and the subjective evaluation (see Appendix C and D), which is why, if scrolling is needed, discrete scrolling is recommended.

Avoid scrolling

To minimize driver distraction and thus increase the safety of drivers, avoid any type of scrolling in the UI if possible. The majority of the tasks that produced the highest driver distraction focused heavily on scrolling. According to the subjective assessments using the NASA-TLX, the tasks that users performed worst at were tasks concerning scrolling (see Appendix C) and generally, users express a desire not to scroll while driving.

Use large images to represent recognizable content

Represent recognizable content by using large images, as users easily can spot these and do not have to read text. Based on user testing, this is a more safe way of presenting content, than by presenting more, but smaller content on the screen. However, it has been seen that if content is unfamiliar, images do not help the user in finding specific content. In such cases, using large text in a vertical list is preferred.

When comparing the task of finding content with a well known cover art to finding content without a well known cover art (where users had to read labels), finding content by cover art performed better when looking at all observed metrics (see Appendix C). This is also connected to Nielsen's guidelines for recognition rather than recall - "Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another." [35], as people will recognize content they know about easier and thus make for a safer experience. It also connects to one of Shneiderman's guidelines, recommending to "Reduce short-term memory load" [47].

Only show a subset of content and enable search for more

Overwhelming users with options, forcing them to explicitly choose between many different alternatives can increase cognitive load, selection time and overall unhappiness [7]. Therefore, instead of showing long lists of content, only show a subset of relevant content and offer the user the ability to search for more specific content using voice. This will reduce driver distraction, as it has been observed in user testing that scrolling is generally an unsafe behaviour when interacting with UIs while driving (see Appendix E). By limiting the amount of content, this can be avoided.

Cognitive walkthroughs showed that when driving, users are fine with being presented with a smaller, yet relevant subset of content. In this way, analysis paralysis can be avoided, where the user, overwhelmed by content, would be unable to make a quick decision. Users admit they would be okay with not being presented with all available content, as long as they can find content in some other way, preferably by voice search.

Limit to 6 items of content on one page

When presenting users with content, do not present more than 6 items on one page. During the user tests, participants said some screens contained too many items and that it was too much information. These screens contained a grid of 8 items, one item consisting of a cover art and a label below. The users then commented on the items being too small and that it was hard finding something. During the cognitive walkthroughs and also during the user test, participants said they were okay with a minimum of 3 items per screen. Therefore, presenting content is preferably done showing 3-6 items per view. Choosing 6 and not 7 as the maximum number comes from preferably presenting that many items in a grid and therefore keeping the number even.

Support direct play of content

Allow users to directly play content on a top-level, instead of forcing them to browse into playlists, albums or artist pages and select specific content to play. This not only reduces the number of levels in hierarchies and thus the number of interactions that users have to perform to get something playing, but also reduces analysis paralysis, as motivated previously. This ultimately creates a safer driver experience, as scrolling in long lists and overall navigational excise can be avoided. During user testing, participants appeared confused when music did not start playing directly when clicking on a cover art and tasks that required users to navigate into albums to select songs performed worse in all aspects, compared to direct play (see Appendix D and E). Moreover, users expressed a preference for direct play, saying they would prefer this to having to click one extra time and select a specific track to get something playing. If they would want to listen to a specific song, they would use voice search.

Provide the user with a personalized in-car experience

By suggesting highly personalized, relevant content[2] for drivers, unsafe behaviours, such as scrolling in lists, can be avoided. It is crucial for drivers to be able to complete secondary tasks as efficiently and effectively as possible and providing them with a personalized user experience can help them do this [38]. An example of how this can be done is by offering drivers the option to continue to listen to music from their last session, directly at start up. It has been found that users prefer to be presented with the ability to quickly get into the music from their last session and that they even prefer it to auto-play after a number of seconds, but with the ability to easily cancel if desired. Another way to create a personalized experience is to present users with a number of personalized playlists, based on previous listening data, recommending them what to play. Turning to recommendations is a common behaviour among users [7] and this goes in line with Apple's *Human Interface Guidelines* for CarPlay - "Show the most relevant content first" [6].

[2] What is and what is not considered as relevant to the user is, however, highly subjective and up to each maker of music applications to decide by themselves.

Place interactable UI-components in easy to reach touch zones

As head units can be divided into three distinct touch zones (see Figure 2.2), it is important to take the aspect of where the user can and can not reach into consideration when designing in-car experiences [60]. Due to form factor and placement of most existing head units today, UI-components that are meant to be interacted with more often than others are more suitable to placed within the Easy (to reach) touch zone, while designers should avoid placing interactable components within the Ok and Hard touch zone. Due to this, the Ok and Hard zones are better suitable for presenting information that are not directly interactable. Interactable UI-components can also be placed at edges and corners of the screen for easy access by the driver, as this allows them to physically feel their way to the component. Users commented on the ability to see the controls while keeping both hands on the steering wheel, which was hard for example when the pause button were placed in the left lower corner.

Provide users with audio feedback when possible

Giving users clear feedback is important, as it informs them about the status of an object or process in the UI. In the context of driving it is crucial to keep the user informed about the systems current state to avoid confusion, which ultimately leads to driver distraction. To allow for the user to keep her eyes on the road as much as possible, provide them with clear, understandable audio feedback when possible [20]. Fortunately, this is is an inherit advantage for music applications, where interactions such as pausing the current song gives direct audio feedback in the sense that the music stops. In general, audio feedback is to be preferred because of the goal to minimize the need to look at the screen while driving.

Allow autostart of music at startup, but allow users to cancel

In order to get the driver into their music as fast as possible, present them with the option to autostart relevant music at startup. It is, however, important to also provide them with the possibility to easily cancel or start the music directly. By doing this, drivers will not have to browse for music while driving. This allows for a safer driving experience, not only as navigational excise decreases, but also as cognitive load and selection times decreases with less options to choose from [7]. As this depends much on what content is relevant, testing this in a simulated environment could not give a completely correct answer. Instead this would need to be tested by users in their everyday life to see if this behaviour would work in the long term.

If the application has an established navigational model and/or visual language, use this, as users expects consistency over platforms

Provide users with the same navigational model and/or visual language they are used to if the application exists on other platforms as consistency helps users to reach their goals more easily [47]. From the cognitive walkthroughs and user testing, it has been seen that users expect consistency over platforms when interacting with an application they are familiar with. Breaking such consistency could lead to confusion, which contributes to driver distraction in the context of driving. Even though Nielsen's *10 Usability Heuristics for User Interface Design* recommend to "Follow platform conventions." [35], it seems to make more sense to follow conventions across platforms.

6

Discussion

This chapter will offer insights and reflections into the work and outcome of this master thesis. There will also be a critical and evaluative discussion regarding the methodological approaches of the project and their effect on the end result. Moreover, in this upcoming Chapter, the use of *we* will refer to the the authors (Jäberg, L and Malmqvist, M).

6.1 The result

6.1.1 Focus on touch interfaces

The work of this master thesis mainly focused on interactions using touch, but as have been seen in research there exists a plethora of other interaction models (such as steering wheel controls or detailed voice control) that can be investigated further to improve driver safety. These other interaction models were not further investigated due to the limited time and already quite large scope of interactions using touch. As been explain in the Theory chapter, some research suggest that other interaction models can provide safer in-car interactions and we therefore believe it would make sense to look deeper into this; however, as has been seen, most current embedded systems in cars today rely on touch as their primary interaction model and as there exist a need for guidance on how to design safe interactions for such systems now, it would be unwise to not set a focus on touch interaction. As also previously argued, makers of music streaming applications have limited power over what in-car hardware is available, which is also a reason why it makes sense to focus on touch interactions that can be performed within the UI, instead of other interaction models that rely on other systems outside of it, such as steering wheel buttons and active microphones for voice control.

6.1.2 Hypotheses

Regarding the design hypotheses that were derived from the three iterations of the design phase, some of these hypotheses did not prove to be as we thought and could therefore not be confirmed and turned into guidelines. Some turned out to be wrong and for some the results of testing was ambiguous and would therefore need more testing. This is interesting, as some hypotheses was derived from user expectations and desires, but turned out not to be as intuitive when being tested for safety in the context of driving a car. One reason for such situations to arise could be users not being used to the thought of the system being used in a car, while driving. Just telling users to imagine driving a car and asking them how they expect something to work and drawing conclusions based on this could be faulty. A better way of carrying out cognitive walkthroughs could therefore have been in the simulated driving environment as well; although, this would have been more time consuming and overall demanding.

Some hypotheses were able to be turned into guidelines, some were proven to be wrong, and some are still in need of more testing to be evaluated correctly. For example, the hypothesis regarding presenting content in a grid rather than in a list could not be confirmed partly due to the fact that the tasks including a grid often were described as more demanding from the user. Even though grids performed slightly better than lists in observations of driving and task performance, due to many comments from users who thought the grid contained too many items, we chose to leave it as a hypothesis to be researched further. It was, however, seen that the main advantage with a grid, over a list, was probably that more content could be shown in one view, which would ultimately help to avoid scrolling, which was shown to be a prominent factor to unsafe user behaviour.

Something else that contradicted the hypotheses was that vertical scrolling performed better in all tests, compared to horizontal scrolling. Moreover, as most tasks containing scrolling in some way was seen as especially problematic regarding eye glance behaviour, we prioritized the guideline to limit or decrease scrolling. We could also see that the tasks testing the scrolling direction was often very much connected to content and how much effort the participant had to put into actually finding specific content. This made the scrolling less prominent and it would have been better to design specific tasks only which test the actual interaction, instead of combining it with finding something in the UI. This would need to be further tested to be able to be turned into a guideline.

6.1.3 Keeping the guidelines general enough

It is debatable how general the suggested design guidelines really are, as much of the work focused around analyzing existing music streaming applications and eventually prototyping one for ourselves. Although this possibly made the guidelines less generally applicable, it enabled us to look into questions such as how much of an applications original functionality and look-and-feel should be brought into car platforms and how much this actually affect the driver performance and the level of distraction. The fact that the prototype was designed in a certain way made it hard for users not to see similarities between the prototype and already existing applications. This can have resulted in simplifications in tasks where users recognize how to carry out certain actions, as well as expectations on the prototype to function in a certain ways. However, doing it the other way around would have been to have the prototype not function in the way that users expected it to and we believe this could have been an even bigger problem, as it would create unnecessary confusion and divert attention from the things that were tested in the user tests.

6.1.4 Autonomy

Even though the results of this master thesis undeniably are of great significance for many stakeholders, such as car manufacturers, makers of music streaming applications present on embedded car platforms, drivers and the area of in-car HCI as a whole, one must not forget that advancements within autonomy and self driven cars could render much of our findings obsolete; or at least out-dated.

This is an interesting subject touched upon in *Vehicle Interaction Tailored for You* - "We must not forget that the task of driving is becoming less important as more research and industry is geared toward autonomously driving cars. If, and when, we get there on a large scale, none of this will matter, except to the car driver/passenger, who will want to make the best out of his or her time in the vehicle." [38]. Given a future where drivers do not have to engage in the primary activity of driving any more, interaction design in cars will not have the same constraints as today and the research on the subject of in-car HCI will have to be updated accordingly. In such a future, drivers can afford to be more engaged in secondary activities. Users will probably not be satisfied with UI's tailored for a safe, yet limited experience, if they at the same time have the ability to interact with other devices, such as their hand held smart phones, where they are offered the richer experiences they are used to. Even though there are plenty of research within HCI focusing on the transition to autonomous driving, there is a lack of research on the everyday experience with autonomous driving [32]. Therefore, designers of in-car HCI will have to explore the area further, as constraints constantly change. Even though safety and driver distraction is of great importance today, it is a constraint that might not be as important in the future. Just as Cooper concludes in *About Face* - No matter how skillful the designer, she will not be successful if she does not have clear and detailed knowledge about the problems constraints [13]. Even though the findings of this master thesis is very relevant for the market of today, it is the job of a designer to adapt to constraints and just as within all disciplines of design, it is the responsibility of the in-car HCI designer to adapt their work to development within relevant fields of technology and hopefully, in-car HCI will grow alongside autonomous driving, instead of falling behind it and rendering itself as useless.

6.2 Methodological discussion

6.2.1 Empathize

The existing insights into needs and goals of users that we based the user behaviours on was both based on the overall user experience of music streaming applications on platforms such as smartphone and desktop, as well as the in-car platforms. Consequently, some user needs and goals were less applicable in the context of driving than others. It was therefor considered important to take this into consideration when engaging in impact mapping and setting priorities. Therefore, user needs that seemed less applicable to the context of driving was prioritized lower than other more important needs. To ensure that these priorities were realistic, this was discussed with our supervisor and designers with experience on the subject of in-car HCI.

Moreover, the fact that we did not carry out our own interviews or in another way collected our own data about the use of music streaming applications while driving was a choice we made early in the process. A lot of research had already been done in the field and due to the vast amount of data already collected, collecting an equivalent amount of qualitative data on our own would have taken too much time and would not add enough value to the project, although a "purer" data set of only user needs from an in-car context would be welcome. One way of doing this could have been to conduct a quantitative user survey of what actions and features are most sought-after in the in-car experience. In such a case we would have to make sure such a survey only targeted actual users of music streaming applications in cars; however, these users would already have an existing preconception of what an in-car experience includes.

6.2.2 Define

The define phase was focused on summarizing and analyzing findings from the literature review, initial user tests of existing platforms, and the qualitative data already collected. Some different methods were used and the findings were very valuable. In hindsight though, the work could have been more focused on this phase and we could have given it some more time because it felt like we rushed through this phase a bit too fast. We believe it would have improved the work of the upcoming phases if the define phase was given more time and had been structured better. In hindsight, some methods were rushed through and did therefore not add as much value as hoped to the project, which made the process a bit unfocused. It would probably have been better to focus on fewer methods carrying out these more rigorously.

The user behaviours found in this stage was of great importance for the work and made us able to make sure to design for the intended users. Although, the behaviours were based on both non-driving users as well as driving users which might have given another result than if only focusing on drivers. It can also be discussed that because we did not conduct the qualitative interviews from where the data was extracted ourselves, we had no influence on what questions were asked and how they were asked.

6.2.3 Ideate

To focus ideation, user journey maps were developed and even though this proved to be useful, we believe that we could have focused more on this work, as the context were very similar for each behaviour. As motivated, we saw it as beneficial to create behaviours rather than personas; however, an advantage with personas is that you can explore the whole picture of a user rather than just a kind of behaviour when interacting with a design. Working with behaviours, it was easy to focus on the right thing, but we experienced that it also was hard to ideate on solutions for these behaviours, as they provided very limited amount of information about a certain user and we did not really know anything regarding the user outside of context of its behaviour. Exploring this further, creating more comprehensive user journeys could therefore have helped ideation to move a little further outside the box.

The three brainstorming workshops that were held to ideate on solutions for the user behaviours worked out well, but as the behaviours were all set in the context of driving and interacting with the head unit, the workshops turned out quite similar. Using the same participants for several workshops also could have affected the outcome of this phase as they had prior knowledge about ideas that came up during the earlier brainstorming workshops and took this with them into the next one. This might have made limited their creativity and due to this it might have been good to not use the same participants twice. On the other hand, we think that when using the same participant twice, they were used to the concept, they felt more engaged and confident in brainstorming and could thus contribute more to the workshop.

The design concepts that were developed were not always focused on interaction, but sometimes more on content and look-and-feel. This came natural, as we aimed for concepts being safe and more user interaction implies more driver distraction. Therefore, interactions sometimes came second, which might have caused the design concepts to be less about how interaction would be made and more about design choices limiting interactions. It was interesting to explore such design concepts where the focus lay more on limiting interactions and not about how they would be performed.

6.2.4 Prototype

By the second iteration of the behaviours, we changed from low-fidelity paper prototypes to high-fidelity prototypes created with Sketch, which made testing easier even though it narrowed down the number of design solutions. This was proven to be a success, as users generally seemed less confused regarding pliancy, hinting, and affordance, when trying out the high-fidelity prototypes. The process of sketching was also more structured after the first iteration where we planned what to test before deciding on what functionality the prototype should include. This made it easier for us as designers to know we were testing the right things and what the expectations of the results from the tests were. It should though be said that we did spend more time on fine tuning the visuals of these prototypes compared to the paper ones, which in hindsight was unnecessary.

Moreover, an issue that was identified was how test users tried to quickly visually scan the UI when given the task of finding specific content, but failed to do so because of the content being unfamiliar and not recognizable to them. Most users reported that they found it irritating having to read texts, as they could not rely on finding a well known cover art when the content in the prototype was not recognized by them. Interacting with a well-used application in everyday life, users would most likely recognize "their own" content and recall for example album covers of known artists. This was also reflected in their primary and secondary task performance. When looking for a well known album cover (e.g., The Beatles' widely known zebra crossing cover of Abby Road), users commented on easily being able to find what they were looking for because they could quickly scan the large images for the specific album cover.

A possible solution to the problems with unrecognizable content could have been to create a more personalized prototype for the participants. This could, however, be problematic, as to obtain this data, the participants would have to report what they usually listen to before the tests. It would also mean that we would have to redesign the prototype between every user test, which would not be feasible. Therefore, instead of real content, only made-up content was used in the prototype, to avoid users partially recognizing artists, albums, songs, etc. We think that recognizing some of the content could lead to misleading results, as this could arbitrarily make some tasks harder than others, based on what the users recognize. Using content well known to most users could have been a solution; however, figuring out what is generally well known by most users would be a challenge in itself. By only using made-up content, all the participants had the same starting-point, with no prior knowledge about any content found in the prototype.

6.2.5 Test

6.2.5.1 User testing

During the initial user tests of existing in-car applications, which we carried during the pre-study, we solely used a manual assessment model for determining distraction levels for tasks, regarding eye glance duration and eye glance frequency, presented in the Process chapter. An inherit problem with this is the subjective judgment of the observer during the test and the possibility of missing out on important behaviour and events; however, given that the same person observed all tests (and in the same way), this would at least indicate on overall themes and thus help us prioritize usability issues to focus on.

For the user testing of the design hypotheses (at the end of the design process), a recording was done of the test participant to be able to correctly evaluate the eye glances and overall behaviour. As done in many other studies on the topic of in-car HCI, we could have used eye tracking equipment to track eye glance behaviors [29], instead of manually analyzing video footage of the test participants gazing, implying a certain subjective judgment of the footage. If eye tracking equipment was used we would have been able to generate heat maps of where users focus their gaze. This would probably have helped to improve the overall result of the user tests; however, subjective judgment would still exist, as in such a scenario, we would still have to examine the heat maps and make decisions based on them. Although, as a report produced by Norman Nielsen Group [40] concluded, an eye-tracking study would require a larger set of test participants (around 40) if the data produced by it should be considered meaningful and this would frankly not be realistic for the scope of this thesis.

Something else that was observed as an issue during the user tests was when participants carried out comparative tasks. These tasks were functionally similar by design and it was found that some users recognized playlist and album covers by their second task. This inevitably had an affect on the cognitive load of the participant. This was also something observed in a related study [17]. The issues were alleviated by asking the participants to find other content than they first did, minimizing the risk of having them remember what cover art to look for; however, one could argue that they still had an advantage when carrying out the second task as they had scanned the UI once before.

Furthermore, in many of our cognitive walkthroughs, which were carried out as a part of each iteration of the design phase, test users were subjected to several design proposals. This was done as we wanted to evaluate expectations of various alternatives. We believe this raised the overall quality of our findings, as research suggest that test users presented with several design solutions give more honest critique towards them, as opposed to being presented with only one design solution [54].

6.2.5.2 Evaluation methods

During the course of the project, we engaged in various evaluation methods with different purposes. In the beginning of the project our evaluation of existing in-car platforms served as benchmarking, in the middle of the project, we used formative evaluation of our various prototypes, and in the end we evaluated our design hypotheses using summative evaluation.

The initial user tests that were carried out during the pre-study provided us with valuable insights into usability issues of existing in-car applications and can be seen as a type of benchmarking of existing design solutions.

The cognitive walkthroughs that was carried out during each iteration of the design phase was carried out in a formative way to evaluate our created prototypes. This helped us to quickly get feedback on our ideas and to gain insights on users' expectations and what they found intuitive. Not only did this help us to improve the quality of our ideas and eventually create new prototypes, but it also helped us to improve on our method, as we became better at carrying out the walkthroughs with the users.

At the end of the project, the design hypotheses were evaluated in a summative way, in order to see if we were able to verify them or not. As mentioned, this was done by giving the test user a number of tasks to complete in the simulated driving environment while measuring driver distraction. As we wanted to validate the derived design hypotheses, it made sense for us to engage in user testing where we evaluated the hypotheses by creating a prototype following them. Even though this evaluation was primarily summative, as we tried to verify hypotheses and draw result-based conclusions, we used some formative methods to do it as well. An examples of this was how we engaged in the thinking-aloud technique. We found this to be helpful as it made it easier for us to understand why certain hypotheses were verified or not. With this said, our suggested guidelines are just that - *suggested*. The guidelines would therefore need to be more thoroughly evaluated using more summative methods.

6.2.5.3 Simulated test environment

Measuring driver distraction in a simulated environment cannot at all be compared to real driving, which needs to be considered for the result of this thesis. It is very hard, probably impossible, to simulate a real driving experience where the driver acts exactly like they would have done if being on a real road. The stakes would of course be higher if driving a real car with other real cars around, carrying real people, and the driver would therefore probably act in a safer way than when driving in the car simulator. Therefore the interaction design guidelines derived form this thesis would preferably be further tested in a real driving environment to get a more accurate result. Some features, such as having a personalization in-car experience would need to be tested in users daily life for a longer period of time to be proven as valuable. On the same note, user might express a desire for specific features or expect them to function in certain ways during a test, but in real life, is this still true? Some things that seem useful might turn out to be annoying in daily use - such as being presented with the option to autoplay music at startup.

While carrying out the user tests in the simulated driving environment, a number of issues with setup and the prototype arose. Most issues were minor, while some were big enough that the test had to be aborted altogether. As mentioned, the driving simulator used for the user tests of the design hypotheses were not the same as the driving simulator used for the initial user testing. One thing that was not anticipated from the new simulated driving environment was the higher level of nausea it caused among some participants, where they felt sick from driving. The majority of test participants experienced some degree of nausea, whereas one participant had to abort the test 10 minutes into it and one had to take break and lower the speed in the middle of the test. To solve this issue, the speed limit which the participants were told to keep were lowered from 80 km/h to 60 km/h for the rest of the upcoming tests and more thorough instructions and warnings regarding turning and breaking were given, as this seemed to be the main cause of nausea. This was undeniably problematic as changing the speed to maintain during the test most probably affected the participant's performance of the secondary tasks given to them, thus also the overall outcome of the tests.

Another issue that ties together with testing in everyday situations concerns the realism of the driving simulator and the fact that it did not simulate non-optimal road conditions that in many cases would affect the precision of the users interactions. It is not hard to believe that driving on any other road surface than an ideal one would cause secondary tasks to be much harder, and the fact that the driving simulator did not simulate realistic road surfaces allowed test participants to keep their hands perfectly still and interact with better precision than in reality, while driving.

6.2.5.4 Test participants

As mentioned, one of the requirements for our test participants was to be an active user of a music streaming application. Due to this, there could be an impact on the results as we did not take into account first time users and users that generally are not that familiar with existing designs or have any preconceptions regarding how the applications work. A more diverse user group might have been beneficial for the project as for example older people might have different knowledge of digital applications and it is not hard to believe that many owners of cars are older than the major user group of music streaming applications. Even though the limited variation of test participants that were used most probably affected the result, it made sense to focus on users from the main user group of music streaming applications due to the time limit of this thesis.

6.2.5.5 Use of NASA-TLX

Our early approaches of applying the NASA-TLX protocol might be flawed as test users were given the self-assessment of filling out the NASA-TLX at the very end of the initial user tests, and not at the end of every task itself. It is not hard to imagine that users forgot how well they performed a certain task and might also have mixed up tasks trying to remember them in hindsight. In the user test of the design hypotheses this was instead done right after each task to get a more correct answer. The users verbally answered the NASA-TLX questions asked by the narrator, while still driving. Even though this seemed like a better option than to do it afterwards, it is fully possible that users answered differently when answering to us, rather than filling out their answers on a paper with out answering to us. The NASA-TLX is very self assessing and admitting to having difficulties performing tasks may be harder to admit verbally, than by writing this down individually. As proposed by NASA Research Group, another possible solution to solving the issue of filling out the assessment afterwards, might have been to provide the test user with a videotaped playback of themselves performing the task evaluated for them to recall [21].

6.2.5.6 NHTSA guidelines

As explained in the Theory chapter, to test a design against the NHTSA guidelines one could follow several distraction test protocols with different testing purposes and acceptance criteria. During this thesis these guidelines helped in guiding the work and testing design solutions for safety regarding driver distraction. We focused on the test protocol called EGDS (Eye Glance Testing Using a Driving Simulator) [33], comparing the results of tasks to its criteria; however, as our test procedure and setup did not follow that of EGDS, the results could not be evaluated against its acceptance criteria. Even though this means that we can not present concrete examples of UIs that have passed the NHTSA guidelines, we do not see this as a very big problem, as the results of this thesis are guidelines and not concrete design solutions. The NHTSA guidelines were also mainly used to compare various interaction models in order to derive the guidelines about which interactions were safer than others and which interactions should definitely be avoided.

It has been seen that the safest in-car experience is one where the driver does not interact with the head unit, at all. This is well reflected in the results of the user tests, as it was nearly impossible for tasks to pass the acceptance criteria set up by the NHTSA guidelines (namely EGDS8.a). Given a task with 10 participants that approximately spends 6 glances each finishing a task, it is enough with 1 glance slightly above 2 seconds, out of all 60 glances (as 1 is 16.76 % of 6) to fail this criteria. Although, as stated we did not let these results solemnly influence the final result. This proves that there needs to be more research carried out and that new design solutions needs to be thoroughly tested before actually releasing something to be used by drivers around the world.

6.2.6 Overall reflection on Methodology

The decision to focus each iteration of the design phase on one of the found user behaviours, instead of trying to put together one concrete design concept in the beginning and evaluating it against the behaviours can be discussed. We believe that going through the design process for each behaviour was of great importance for the outcome of this thesis, as it gave us the opportunity to explore ideas and solutions iteratively, which might not have come up if we directly had focused on finding one concept that would fit all and included all functionality from the beginning. Doing the way we did it, we were able to develop concepts that did not have to include all basic functionality of the original application because of only looking at the specific needs of one specific behaviour at the time. This worked, as the deliverable of the thesis was design guidelines and not a concrete design solution. As behaviours of people can be quite different it makes sense to divide work and focus on one behaviour at the time to explore what solutions would best fit each behaviour, before putting things together, creating a final concept to test more thoroughly.

6.3 Ethical issues

Designing for applications that are to be used in cars by drivers forces you to think twice about what to actually add to an application and what actions are allowed for the user to perform. If an action is available, the users will most likely use it and if it in that case cause unsafe driving, it could result in fatal accidents. This has to be taken into consideration, as we provide guidelines for designers on how to design interactions with music streaming applications for drivers. If we provide faulty guidelines and designers follow them, it could lead to unsafe behaviour and in worst case, car crashes. At the same way guidelines needs to be concrete and clear, as ambiguity can cause designers to interpret guidelines in wrong ways. Even though we believe that we suggest clear and valid guidelines, it is of great importance for everyone working within the field of in-car HCI to test and evaluate their applications thoroughly to make sure they are safe to use while driving.

7 Conclusion

This master thesis looked into the user experience of music applications found in head units in modern cars. More specifically, the focus was set on how designers can create safer interactions for users by taking driver distraction and usability, aiming to answer the research question of

What guidelines are suitable for designing interactions with music streaming applications found in Head units in cars, taking driver distractions and usability into consideration?

To be able to answer this question, the thesis work used a research through design-, and goal-directed design approach to understand the user, her goals, and behaviours in the context of driving and interacting with a music streaming application. A number of user behaviours and needs were found and a music streaming application on existing in-car platforms were evaluated to understand what usability and safety issues existed for users trying to reach their goals. To find possible solutions to these issues, a number of brainstorming workshops focusing on the different user behaviours were arranged, which lead to various design concepts that were evaluated using cognitive walkthroughs. Insights gained from this, along with an extensive literature review resulted in a number of hypotheses regarding how to design safer interactions for music applications in the car. These hypotheses were evaluated through user testing in a simulated driving environment that evaluated driving performance, usability, visual attention and cognitive load. As a result, a number of suggested design guidelines with a focus on increased driver safety were derived. These guidelines are considered to be taken into account, together with other already existing guidelines and best practices within interaction design and in-car HCI, to improve the overall experience and interactions with music applications having its interface displayed in the head unit of the car.

As the suggested guidelines were derived, the research question was able to be answered, but as for any wicked problem, there is no way to be certain that they are the best guidelines for designing safe interactions for drivers. What can be said, however, is that the derived guidelines can help designers in creating safer in-car experiences for music streaming applications and to eliminate especially unsafe interactions.

The work of this thesis was limited in some different ways. The testing was only made in a simulated environment and not during real driving. The indirect focus on just one music streaming application can be argued to have affected the result and given a less general outcome of this thesis than hoped for.

7.1 Future work

Due to the wicked problem nature of the research question, the final suggested guidelines would preferably have been iterated upon several times and evaluated more thoroughly with a more diverse user group. This is something that cannot be stressed enough, as in-car HCI has such a big impact on drivers, and can contribute to critical situations for both drivers, passengers and people outside of the car. Creating a safe in-car user experience must be the number one priority of designers. For this, future work would both need to take place in the simulated driving environment and in an ethnographical fashion in cars of users, to disclose how users interact with music streaming applications that follows the proposed guidelines. Even though a driving simulator provide good ways of observing users, it can never accurately reconstruct everyday situations that arise when driving, such as how non-ideal road surfaces affect clicking or how screaming children in the backseat adds to the cognitive load of the user. Moreover, as testing only was done on a 7.9 inch iPad mini in landscape mode, future work also needs to investigate how the derived guidelines translate to other form factors of head units, as it was seen during user testing and cognitive walkthroughs, form factor and screen orientation have a big impact on user expectations and their intuitive interactions with the UI. Moreover, even though the delimitation of focusing on touch interaction proved reasonable for the scope of this thesis, it would also be of interest to investigate other interaction models than touch, such as voice interaction and its implications on driver distraction.

Bibliography

[1] Florian Alt, Dagmar Kern, Fabian Schulte, Bastian Pfleging, Alireza Sahami Shirazi, and Albrecht Schmidt. Enabling micro-entertainment in vehicles based on context information. In *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '10*, page 117, New York, New York, USA, 2010. ACM Press.

[2] Android. Android Auto.

[3] Android Auto. Android Auto - Guidelines.

[4] Android Central. Getting started with Android Auto in your car.

[5] Apple. Apple CarPlay.

[6] Apple CarPlay. Audio Apps - Overview - Human Interface Guidelines for CarPlay Apps.

[7] Daniel Boland, Ross McLachlan, and Roderick Murray-Smith. Engaging with Mobile Music Retrieval. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15*, pages 484–493, New York, New York, USA, 2015. ACM Press.

[8] Warren. Brodsky. *Driving with music : cognitive-behavioural implications.*

[9] Gary Burnett, Elizabeth Crundall, David Large, Glyn Lawson, and Lee Skrypchuk. A study of unidirectional swipe gestures on in-vehicle touch screens. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, pages 22–29, New York, New York, USA, 2013. ACM Press.

[10] G.E. Burnett. On-the-Move and in Your Car. *International Journal of Mobile Human Computer Interaction*, 1(1):60–78, jan 2009.

[11] Centre for Excellence in Universal Design. User testing.

[12] Donald Clark. Types of Evaluations in Instructional Design .

[13] Alan Cooper, Robert Reimann, Dave Cronin, and Christopher Noessel. *About face : the essentials of interaction design.*

[14] Riccardo Coppola and Maurizio Morisio. Connected Car: Technologies, Issues, Future Trends. *ACM Computing Surveys*, 49(3):1–36, oct 2016.

[15] Siang Teo Dam Rikke. 5 Stages in the Design Thinking Process | Interaction Design Foundation, 2018.

[16] DMV. Distracted Driving Laws | DMV.ORG.

[17] Johan Engström, Gustav Markkula, Trent Victor, and Natasha Merat. Effects of Cognitive Load on Driving Performance: The Cognitive Control

Hypothesis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(5):734–764, aug 2017.

[18] Danny Godin. Aspects of Research through Design: A Literature Review.

[19] GROM. The Evolution of the Head Unit.

[20] Erik Hassel. Touch screens in cars Investigating touch gestures and audio feedback in the context of in-vehicle infotainment. 2016.

[21] Human Performance Research Group. NASA Task Load Index (NASA-TLX). Technical report, NASA Ames Research Center, Moffett Field, California.

[22] IDEO U. Effective Brainstorming Techniques.

[23] IFPI. Global music report. Technical report, IFPI, 2018.

[24] Jakob Nielsen. Why You Only Need to Test with 5 Users. 2000.

[25] Jasper Liu. Visualizing the 4 Essentials of Design Thinking – Good Design – Medium, 2016.

[26] By John Chris Jones, John Wiley, and New York. Design Methods: seeds of human futures.

[27] Tuomo Kujala and Pertti Saariluoma. Effects of menu structure and touch screen scrolling style on the variability of glance durations during in-vehicle visual search tasks. *Ergonomics*, 54(8):716–732, aug 2011.

[28] Petri. Lankoski, Staffan Björk, and ETC Press. *Game research methods : an overview*. ETC Press, 2015.

[29] Kenneth Majlund Bach, Mads Gregers Jaeger, Mikael B Skov, and Nils Gram Thomassen. Interacting with In-Vehicle Systems: Understanding, Measuring, and Evaluating Attention.

[30] Bella. Martin and Bruce M. Hanington. *Universal methods of design : 100 ways to research complex problems, develop innovative ideas, and design effective solutions*. Rockport Publishers, 2012.

[31] Kevin McCull. UX for connected cars, 2016.

[32] Alexander Meschtscherjakov, Wendy Ju, Manfred Tscheligi, Dalila Szostak, Sven Krome, Bastian Pfleging, Rabindra Ratan, Ioannis Politis, Sonia Baltodano, and Dave Miller. HCI and Autonomous Vehicles. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '16*, pages 3542–3549, New York, New York, USA, 2016. ACM Press.

[33] Department of Transportation (DOT). National Highway Traffic Safety Administration (NHTSA). Federal Register :: Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. Technical report, 2013.

[34] Department of Transportation (DOT) National Highway Traffic Safety Administration (NHTSA). Visual-Manual NHTSA Driver Distraction Guidelines for Portable and Aftermarket Devices. Technical report, 2016.

[35] Jakob Nielsen. 10 Usability Heuristics for User Interface Design, 1995.

[36] Nielsen Music. U.S. Music 360 - 2017 Report Highlights. Technical report, 2017.

[37] Don Norman. *The Design of Everyday things*. Basic Books, 2002.

[38] Carl Jörgen Normark and Carl Jörgen. Vehicle interaction tailored to you. *interactions*, 22(1):32–36, jan 2015.

[39] Ingrid. Ottersten and Mijo. Balic. *Effect managing IT*. Liber, 2007.
[40] Kara Pernice and Jakob Nielsen. How to Conduct Eyetracking Studies.
[41] Polisen. Mobilanvändning i trafiken | Polismyndigheten.
[42] N Reed and R Robbins. The Effect Of Text Messaging On Driver Behaviour. Technical report, 2008.
[43] Teo Siang Rikke Dam. 5 Stages in the Design Thinking Process | Interaction Design Foundation, 2017.
[44] Jeffrey Rubin. *Handbook of usability testing : how to plan, design, and conduct effective tests*. Wiley, 1994.
[45] Paul Saettler. *The Evolution of American Educational Technology*. Libraries Unlimited, 1990.
[46] Albrecht Schmidt, Anind K. Dey, Andrew L. Kun, and Wolfgang Spiessl. Automotive user interfaces. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA '10*, page 3177, New York, New York, USA, 2010. ACM Press.
[47] Ben Shneiderman. *Designing the user interface : strategies for effective human-computer-interaction*. Addison Wesley Longman, 1998.
[48] Carolyn. Snyder. *Paper prototyping : the fast and easy way to design and refine user interfaces*. Morgan Kaufmann, Elsevier Science, 2003.
[49] David L Strayer, Joel M Cooper, Jonna Turrill, James R Coleman, and Rachel J Hopman. Measuring Cognitive Distraction in the Automobile III: A Comparison of Ten 2015 InVehicle Information Systems. 2015.
[50] Michael K. Svangren, Mikael B. Skov, and Jesper Kjeldskov. The connected car. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17*, pages 1–12, New York, New York, USA, 2017. ACM Press.
[51] Ivan Tashev, Mike Seltzer, Y. C. Ju, Ye-Yi Wang, and Alex Acero. Commute UX: Voice Enabled In-car Infotainment System, sep 2009.
[52] Tesla. Model X Specifications.
[53] The Royal Society for the Prevention of Accidents (ROSPA). Road Safety Factsheet. 2017.
[54] Maryam Tohidi, William Buxton, Ronald Baecker, and Abigail Sellen. Getting the Right Design and the Design Right: Testing Many Is Better Than One.
[55] Tom Kelly. The Perfect Brainstorm. In J Kelly, T &Littman, editor, *The Art of Innovation*, chapter 4. Doubleday, New York, 2000.
[56] Tim Triplett, Rob Santos, Sandra Rosenbloom, and Brian Tefft. American Driving Survey: 2014 – 2015. *American Driving Survey*, 2016.
[57] University of Salzburg. Center for Human-Computer Interaction: Car Interaction Lab.
[58] Yoland. Wadsworth. *Do it yourself social research*. Left Coast Press, 2011.
[59] Ashley Williams and Ashley. User-centered design, activity-centered design, and goal-directed design. In *Proceedings of the 27th ACM international conference on Design of communication - SIGDOC '09*, page 1, New York, New York, USA, 2009. ACM Press.
[60] Richard Young and Jing Zhang. Safe Interaction for Drivers: A Review of Driver Distraction Guidelines and Design Implications. apr 2015.

A

Results of subjective assessments, using NASA-TLX from usability testing of existing platforms

The following Appendix summarize the findings from the NASA-TLX carried out during the initial usability testing of the existing implementation of a music streaming application in CarPlay and Android Auto. S1, S2, S3, S4 and S5 represent the subjective score of test subject 1, 2, 3, 4 and 5 respectively, while Average score is calculated accordingly:

$$Averagescore = \frac{S1 + S2 + S3 + S4 + S5}{5}$$

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	3.0	0.0	1.5	2.5	2.0	1.8
Put on a chart	2.5	6.5	5.5	1.5	1.5	3.5
Save song to library	8.0	9.5	7.0	9.5	8.0	8.4
Put on a personalized playlist	7.0	9.5	5.5	6.5	8.0	7.3

Table A.1: Mental demand, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	2.0	1.0	1.5	3.0	1.0	1.7
Put on a chart	2.5	6.5	4.5	1.5	3.5	3.7
Save song to library	6.5	8.5	5.0	9.5	8.5	7.6
Put on a personalized playlist	7.0	9.5	2.5	7.5	8.5	7.0

Table A.2: Physical demand, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	2.5	4.0	2.5	3.0	4.0	3.2
Put on a chart	2.0	8.0	4.5	1.5	4.5	4.1
Save song to library	8.5	8.5	4.5	9.5	9.0	8.0
Put on a personalized playlist	5.0	8.0	4.0	7.5	8.0	6.5

Table A.3: Temporal demand, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	2.0	0.0	0.0	0.5	6.0	1.7
Put on a chart	1.0	3.5	2.5	1.5	5.5	2.8
Save song to library	9.0	9.5	9.5	9.5	9.5	9.4
Put on a personalized playlist	9.5	9.5	2.0	7.5	7.0	7.1

Table A.4: Performance, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	3.5	1.5	2.0	1.5	5.0	2.7
Put on a chart	1.0	8.0	7.5	1.0	7.0	4.9
Save song to library	7.5	9.0	4.0	9.5	9.5	7.9
Put on a personalized playlist	6.0	9.5	3.0	8.0	8.0	6.9

Table A.5: Effort, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	1.5	0.5	1.0	1.5	5.0	1.9
Put on a chart	1.0	6.0	6.0	1.0	4.0	3.6
Save song to library	7.5	9.5	7.5	9.5	9.5	8.7
Put on a personalized playlist	9.0	8.5	6.5	8.0	8.5	8.1

Table A.6: Frustration level, CarPlay

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	1.5	9.5	6.5	8.5	8.5	6.9
Put on a chart	5.5	7.5	4.0	1.0	6.5	4.9
Save song to library	3.0	0.5	7.0	0.0	2.5	2.6
Put on a personalized playlist	2.0	8.5	3.0	0.5	2.0	3.2

Table A.7: Mental demand, Android Auto

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	2.5	9.5	2.0	8.5	8.0	6.1
Put on a chart	6.5	8.0	3.5	1.0	6.0	5.0
Save song to library	4.0	0.5	4.0	0.0	2.5	2.2
Put on a personalized playlist	4.0	8.5	2.5	0.5	2.0	3.5

Table A.8: Physical demand, Android Auto

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	1.5	9.5	2.5	8.5	7.5	5.9
Put on a chart	3.5	8.0	3.5	1.0	7.0	4.6
Save song to library	4.0	0.5	4.5	0.0	2.5	2.3
Put on a personalized playlist	4.0	8.5	4.0	0.5	3.5	4.1

Table A.9: Temporal demand, Android Auto

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	1.5	9.5	3.5	9.0	7.5	6.2
Put on a chart	3.5	7.0	2.0	1.0	7.0	4.1
Save song to library	2.5	0.0	8.0	0.0	3.0	2.7
Put on a personalized playlist	3.0	6.5	0.5	0.5	2.5	2.6

Table A.10: Performance, Android Auto

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	4.5	9.5	4.0	9.0	8.0	7.0
Put on a chart	4.5	7.5	4.0	1.0	6.5	4.7
Save song to library	3.0	0.5	4.5	0.0	3.0	2.2
Put on a personalized playlist	4.0	8.5	3.0	0.5	2.0	3.6

Table A.11: Effort, Android Auto

Task	S1	S2	S3	S4	S5	Average score
Switch to a podcast app	2.0	9.5	4.5	9.5	8.5	6.8
Put on a chart	5.5	6.0	4.5	0.5	6.0	4.5
Save song to library	4.5	0.5	7.5	0.0	3.0	3.1
Put on a personalized playlist	4.0	6.5	4.5	0.5	1.5	3.4

Table A.12: Frustration level, Android Auto

B

Results of primary and secondary task performance, eye glance behaviour and physiological measures from usability testing of existing platforms

The following appendix contains the results of *primary task performance* (Driving performance), *secondary task performance* (usability), *eye glance behaviour* (driver distraction) and *physiological measures* (cognitive load) from usability testing of existing in-car platforms.

B.1 Tasks

1. Put on some music
2. Just put on some new music
3. Skip 3 songs ahead
4. Skip 20 seconds of this song
5. What artist makes this song?
6. Put on specified artist
7. Go to the album of this song
8. Add song to playlist
9. What song is next in queue?
10. Put on any other content than music
11. Put on a user made playlist by you
12. Put on a song you listened to recently during this test
13. Switch to podcast app
14. Put on a podcast
15. Add this podcast to your favourite podcasts
16. Put on one of your favourite podcasts
17. See if you can find a podcast about cats
18. Go back to the music streaming application
19. Put on specified playlist
20. Save song to library (not available in Carplay)
21. Put on specified playlist

22. Stop the music

B.2 Individual scores

EGF = Eye glance frequency
EGD = Eye glance duration
FL = Frustration level
DP = Driver performance
TP = Task performance

Task	EGF	EGD	FL	DP	TP
0	2	2	2	1	1
1	1	1	1	1	1
2	1	2	1	1	2
3	2	3	3	2	3
4	2	2	1	2	1
5	3	3	3	3	3
6	3	3	2	3	3
7	2	3	3	3	3
8	3	3	2	2	3
9	2	3	2	3	3
10	1	3	1	1	1
11	2	2	2	1	2
12	1	2	2	2	1
13	1	1	1	1	1
14	3	3	3	1	2
15	1	2	1	1	1
16	3	3	3	3	2
17	1	1	1	1	1
18	1	2	1	1	2
19	3	2	2	2	2
20	3	3	3	3	3
21	1	1	1	1	1

Table B.1: Test participant 1

Task	EGF	EGD	FL	DP	TP
0	1	1	1	1	1
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	3	2	2	2	2
6	3	2	1	2	1
7	2	1	2	1	1
8	3	2	2	3	3
9	2	2	1	1	1
10	1	2	1	1	1
11	1	1	1	1	1
12	2	3	2	3	2
13	1	1	1	1	1
14	3	2	2	2	3
15	2	1	1	2	2
16	3	3	2	3	2
17	2	2	2	1	1
18	2	1	1	1	1
19	3	3	3	1	1
20	3	2	2	2	3
21	1	1	1	1	1

Table B.2: Test participant 2

Task	EGF	EGD	FL	DP	TP
0	1	1	1	1	1
1	1	1	1	1	1
2	1	1	1	1	1
3	1	1	1	1	2
4	1	1	1	1	1
5	3	2	2	2	2
6	2	2	1	2	2
7	3	2	2	2	3
8	3	1	1	1	2
9	2	2	1	1	2
10	3	2	2	2	2
11	1	1	1	1	1
12	2	1	1	1	1
13	1	1	1	1	1
14	3	2	2	1	2
15	1	1	1	1	1
16	3	3	3	2	2
17	1	2	1	1	1
18	1	1	1	1	1
19	3	1	2	2	2
20	2	2	1	2	2
21	1	1	1	1	1

Table B.3: Test participant 3

Task	EGF	EGD	FL	DP	TP
0	2	2	1	3	2
1	2	1	1	2	1
2	1	1	1	1	1
3	2	2	1	3	3
4	1	1	1	1	1
5	3	2	2	3	3
6	2	2	2	3	3
7	2	2	1	3	3
8	1	1	1	1	1
9	2	1	1	3	2
10	1	1	1	2	2
11	1	2	1	3	2
12	2	2	1	2	1
13	1	1	1	1	1
14	2	3	2	2	3
15	3	2	2	3	1
16	3	3	2	3	3
17	1	1	1	3	1
18	2	1	1	3	2
19	2	2	1	3	3
20	3	3	2	3	2
21	1	1	1	3	2

Table B.4: Test participant 4

Task	EGF	EGD	FL	DP	TP
0	2	1	1	1	1
1	1	1	1	1	1
2	1	1	1	1	1
3	3	2	2	2	2
4	1	1	1	1	1
5	3	3	3	3	2
6	2	2	2	3	2
7	3	3	2	2	2
8	3	2	1	2	1
9	2	2	1	1	1
10	2	3	2	2	2
11	2	2	1	1	1
12	2	2	1	1	1
13	1	1	1	1	1
14	3	3	2	2	2
15	3	3	2	1	2
16	3	2	3	2	2
17	2	1	1	1	1
18	1	1	1	1	1
19	3	2	2	2	2
20	2	3	2	2	2
21	1	1	1	1	1

Table B.5: Test participant 5

C

Results of subjective assessments using NASA-TLX from user testing of hypotheses

The following chapter contains the raw data from the NASA-TLX carried out to test of the attention measure *subjective assessments* (cognitive load) during evaluation of the design hypotheses.

S1, S2, S3, S4, S5 and S6 represent the subjective score of test subject 1, 2, 3, 4, 5 and 6 respectively, while Average score is calculated accordingly:

$$Averagescore = \frac{S1 + S2 + S3 + S4 + S5 + S6}{6}$$

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	1	1	3	1	1	1	1.50
Big swipe gesture CLICK	1	1	1	6	3	4	2.67
Big swipe gesture SWIPE	3	1	1	3	2	1	1.83
Scroll content - Horizontal scroll	7	7	8	4	3	3	5.33
Scroll content - Vertical scroll	4	3	4	3	4	3	3.50
Grid content Horizontal grid	8	7	6	4	6	6	6.17
Grid content Vertical grid	5	2	7	5	6	5	5.00
Limit to X levels in hierarchies	7	3	9	6	8	7	6.67
Discrete scrolling	6	3	6	4	3	3	4.17
Continuous scrolling	6	5	7	3	5	6	5.33

Table C.1: Mental demand

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	3	1	2	2	1	3	2.00
Big swipe gesture CLICK	3	2	1	3	3	4	2.67
Big swipe gesture SWIPE	1	1	1	3	2	1	1.50
Scroll content - Horizontal scroll	5	8	8	3	3	2	4.83
Scroll content - Vertical scroll	4	2	4	3	4	3	3.33
Grid content Horizontal grid	7	3	6	3	4	2	4.17
Grid content Vertical grid	6	2	6	3	7	2	4.33
Limit to X levels in hierarchies	8	2	8	3	8	6	5.83
Discrete scrolling	5	3	6	2	4	3	3.83
Continuous scrolling	6	5	6	2	7	3	4.83

Table C.2: Physical demand

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	2	1	3	2	1	1	1.67
Big swipe gesture CLICK	2	2	2	3	4	4	2.83
Big swipe gesture SWIPE	2	1	1	3	4	1	2.00
Scroll content - Horizontal scroll	8	8	8	3	2	4	5.50
Scroll content - Vertical scroll	5	2	4	3	2	3	3.17
Grid content Horizontal grid	9	5	5	4	5	2	5.00
Grid content Vertical grid	4	2	4	4	6	2	3.67
Limit to X levels in hierarchies	6	3	8	3	9	6	5.83
Discrete scrolling	6	2	4	2	5	3	3.67
Continuous scrolling	5	5	8	3	6	5	5.33

Table C.3: Temporal demand

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	5	10	3	10	10	3	6.83
Big swipe gesture CLICK	7	10	7	6	7	1	6.33
Big swipe gesture SWIPE	8	10	8	6	8	8	8.00
Scroll content - Horizontal scroll	2	7	2	6	7	7	5.17
Scroll content - Vertical scroll	7	9	5	6	6	7	6.67
Grid content Horizontal grid	3	7	4	7	5	7	5.50
Grid content Vertical grid	7	9	4	6	4	7	6.17
Limit to X levels in hierarchies	7	9	4	7	3	2	5.33
Discrete scrolling	7	7	4	8	7	3	6.00
Continuous scrolling	7	7	3	8	5	4	5.67

Table C.4: Performance

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	4	1	3	5	2	1	2.67
Big swipe gesture CLICK	3	1	1	4	4	3	2.67
Big swipe gesture SWIPE	3	1	1	4	3	2	2.33
Scroll content - Horizontal scroll	6	6	8	4	4	4	5.33
Scroll content - Vertical scroll	4	2	3	3	5	2	3.17
Grid content Horizontal grid	7	5	6	3	6	6	5.50
Grid content Vertical grid	5	2	6	4	7	4	4.67
Limit to X levels in hierarchies	7	2	8	3	8	6	5.67
Discrete scrolling	5	3	7	4	3	2	4.00
Continuous scrolling	6	6	7	3	5	5	5.33

Table C.5: Effort

Task	S1	S2	S3	S4	S5	S6	Average score
Just put on some music	2	1	2	0	2	1	1.33
Big swipe gesture CLICK	2	1	1	3	3	0	1.67
Big swipe gesture SWIPE	1	1	0	3	2	0	1.17
Scroll content - Horizontal scroll	7	3	5	3	2	1	3.50
Scroll content - Vertical scroll	3	1	3	2	4	1	2.33
Grid content Horizontal grid	6	4	7	2	7	3	4.83
Grid content Vertical grid	3	2	6	4	7	2	4.00
Limit to X levels in hierarchies	4	2	9	3	8	5	5.17
Discrete scrolling	4	2	8	2	3	1	3.33
Continuous scrolling	4	4	7	2	5	3	4.17

Table C.6: Frustration level

D

Results from primary and secondary task performance and physiological measures from user testing of hypotheses

This appendix present the results of *primary task performance* (driver performance), *secondary task performance* (usability) and *physiological measures* (cognitive load) of user testing the design hypotheses of the project. It also contains comments given by users during the testing.

D.1 Tasks

0 Just put on some music
1 Put on some new music
2.1 Always show playback controls - pause
2.2 Don't always show playback controls - pause
2.1 Always show playing song - artist name
2.2 Don't always show playing song - artist name
3.1 Skip 3 songs ahead - CLICK
3.2 Skip 3 songs ahead - SWIPE
4 Find a user made playlist
5.1 Find playlist - horizontal list (SCROLL direction)
5.2 Find playlist - vertical list (SCROLL direction)
6.1 Find playlist - horizontal grid (GRID direction)
6.2 Find playlist - vertical grid (GRID direction)
7.1 Find playlist - Unavailable content
7.2 Find playlist - No unavailable content
8.1 Voice search (direct)
8.2 Voice search (one extra step)
9.1 Put on an album (Not direct play)
9.2 Put on an album (Direct play)
10.1 Your Library Grid (Grid vs list)
10.2 Your Library List (Grid vs list)
12 Put on song in an artist's album (full path)
13.1 Find song - discrete scroll

13.2 Find song - continuous scroll
14 Put on personalized playlist

D.2 Comments and observations on tasks

0 Just put on some music
- "A bit hard to reach button at the bottom maybe."
- "Want to have the play button in the middle, cannot see the icon because of my arm (when hands on the steering wheel)."
- "Took some time to find the button because I hadn't seen it before."
- "Some thought on where to find the button, but easy to see."

1 Put on some new music
- "It's easy to reach the big images of playlists."
- "There was no skip button in the home screen, that would have been nice."

2.1 Always show playback controls and now playing or not
- "Expected the pause button to be in the middle."
- "I like to NPV, it's more clear then the home screen."
- "I would prefer the first one (always show). But I'm also confused about the image and title in the middle of this view."
- "I Would prefer to be in the NPV."
- "Closer to have it (the controls and music information) in the home screen. First time it (pause button) was in the corner, so I knew where to look. NPV looks very familiar."

3.1 Skip 3 songs ahead - CLICK vs SWIPE
- Uses the buttons. Easy to reach. Small buttons. A bit hesitant driving Driving a bit bad in the first swipe. Then better.
- Says it was a bit hard to press the buttons three times Didn't see that something was happening when pressing the buttons
-
- "Swiping was easier than finding the button."
- "Better to swipe than to click, no need to find something on the screen"
- "Would prefer swiping, feels more natural and I have a bigger area to perform the action in"
- "It was a bit hard to press the button three times and I didn't see was happening when pressing it."
- "Swiping was a bit easier. Larger. Better feedback this time than the previous (click)."
- "Swipe is nice, you don't have to press anywhere special."

4 Find a user made playlist
- "Easy clicking images (of playlists)."
- "Not hard to find but had to give it some thought. Didn't really read anything but was easy to find."
- "I had a better overview when vertical"

5.1 Find playlist - horizontal list (SCROLL direction)
- "Easy to find."

- "It was pretty far away. Would not want to do that (scroll that much)"
- "I had to read a lot but could only read one at the time"
-

5.2 Find playlist - vertical list (SCROLL direction)
- "Small. Harder to read when vertical."
- "Easier to read the list, multiple titles at the same screen."
- "I could read more than one title at the time which was good. More natural to scroll vertically. Alphabetical order would be nice."
- "This was easier for reading text, had to read because I didn't recognize the images. Faster to scroll because you were able to see several playlists at the same time "
- "A little more challenging than the previous task (Horizontal). I needed to look more at the screen. Smaller items, afraid of missing something. "
- "You can see more with the vertical scroll and find what you're looking for faster. I would prefer vertical I think. Less scrolling. Usually a mobile user is used to vertical lists."

6.1 Find playlist - horizontal grid (GRID direction)
- "Hard to find playlist. Text small. Images small. This is the most confusing one (view). Too much stuff on the screen. No structure and I tried to read everything."
- "Too many things on the screen."
- "Too much information on the screen. "
- "Scrolling horizontally was easier but I don't know why"
- "Horizontal was the easiest."
-

6.2 Find playlist - vertical grid (GRID direction)
- "Easier to find."
- "Didn't have to scroll as much as the previous task (horizontal). "
- "Knew which color of the cover now since I had seen it before. Looked more for color this time"
- "Liked the vertical scroll better"
- "The grid makes it hard to find something, I had to look through a lot of things. I think (scrolling) vertically is easier for me. I would prefer the list as to the grid. You don't want a grid. Prefer the list over the grid - grid was too much at once and it was hard to look across. She was not sure where to look when she scrolls."

7.1 Find playlist - Unavailable content vs No unavailable content
- "irritating to have to look down many times"
- "Shorter way to find it (no unavailable content)."

8.1 Voice search (direct) vs (one extra step)
- "Prefer this one not having to click twice"
- "Would prefer the first one, it required one less click, the button in voice search in the middle was too small."
- "I would prefer the first one (direct play)"
- "Why would I need to click twice? Really irritating!"

- "Quite easy to understand functionality. Would prefer second one (one extra click). More clear when I should start talking. "
- "If I had used it once before it's probably easier. I clicked on both but would prefer if I didn't have to."
- "Prefer first one (direct). If I tap a microphone I don't expect I have to tap another button. As long as it's limited to one tap (it's okay), otherwise start it with voice"

9.1 Put on an album (Not direct play) vs (Direct play)

- "A bit confused about that it doesn't start directly. Prefer the later one (direct play). Do not have to see album information. I just want something playing. It's hard to read while driving."
- "Task took a long time to perform. Didn't expect to have to click two times. When clicking album cover it should start playing directly. Would never search for song by going into an album but I may want to listen to the whole album. Direct play is better."
- "Preferred when you just had to click once (direct play) to start the music."
- "Too many steps." (not direct play)
- "If I wanted a special song I would prefer the first alternative (not direct play). Although I kind of assumed it should start playing directly. I would like the second alternative actually (direct play)."
- "Would be frustrated with this (not direct play). Looks like I would go back if I press the title."
- "Prefer to go straight to playing. Maybe if I was looking for a specific song the other way would be nice (not direct play).

10.1 Your Library (Grid vs list)

- "Too much, don't know how to find anything here in My Library, there's no natural order to find something" (grid)
- "Prefer the grid, because there is a limited number of options and it helps to see them all at the same time."
- "This (list) is better, not too cluttered. Better with less to choose from"
- "Easier to find something in the second alternative (list). Could focus my eyes on one part of the screen. Although once you've learned what the icons look like and where they are placed it might be easier with the first alternative (grid)."
- "This was nice, I had a good overview." (grid)

11 Unlimited content vs subset (this question is not in the usability tests)

- "Recently played all artists would be nice."
- "Would prefer subset, with recently played. No structure, not an efficient way of finding an artist anyway. Easier to use voice command to search for it. Would like the subset to be Most played artists."
- "Would prefer all artists. Or if there were more in the same view."
- "Would prefer Recently played if I also have the search function. Wouldn't want to swipe a lot, and Less things to choose from is good."
- "Would prefer recently played. I often listen to the same music for a while and then recently played would be the best for me."

- "I want to be able to find all artist I have saved. Would not like a long list though, would maybe prefer using voice to search in that case. Recently played would be nice as the subset."
- "The thing with voice is it might be hard to find specific (content). The worst thing would be to be stuck in a view. As long as I knew I could search for stuff using voice I would be okay (with only being shown a subset)."

12 Put on song in an artist's album (full path)

- "Most things being found in the first screen made it easier. If you would have to scroll a lot in the different screens it would be harder."
- "Hard to drive and too many steps."
- "Because a lot happened on the road it was harder to perform."
- "Confusing that in one view I could play the artist and in one view I was browsing to the artist. Would rather use voice search."
- "Missed the image twice. I didn't recognize the cover art, maybe that was why I missed it. "

13.1 Find song - discrete scroll vs continuous scroll

- "An album cover you recognize was easier to find, didn't have to read the text. Preferred the continuous scroll, feels more “natural”. If there's a lot of content you could quickly scroll to the end."
- "The song was further away, First scrolling was better (discrete)"
- "Took more time (continuous). I liked the previous one (discrete) better because it took less time. This (continuous) looks better but is not good in the car"
- "Would prefer the second alternative (continuous). Easier to get an overview. Felt faster."
- "This was harder. I want to keep track of where I am so I don't miss it. Now I have to keep my finger in the display all the time. Definitely prefer pagination. Can swipe without looking."
- "Definitely made it easier that I recognized the cover art (discrete). Oh, I don't like this one. (continuous)"

14 Put on personalized playlist

- "Hard to find."
- "I want to find Discover weekly, but I don't know where they are. They move around kind of. "
- "Good with direct play!"

FL = Frustration level
DP = Driver performance
TP = Task performance

D.3 Individual scores

Task	FL	DP	TP
0	1	1	1
1	2	2	1
2.1	2	2	2
pause 2.2	1	2	2
pause 2.1	1	2	1
artist 2.2	2	2	2
artist 3.1	1	1	1
3.2	1	1	1
4	2	3	2
5.1	1	3	2
5.2	1	2	1
6.1	1	2	2
6.2	1	2	1
7.1	1	1	2
7.2	1	1	1
8.1	1	1	1
8.2	1	1	2
9.1	1	2	3
9.2	1	1	1
10.1	1	1	1
10.2	1	2	2
12	2	2	2
13.1	1	1	1
13.2	1	1	1
14	1	1	2

Table D.1: Test participant 1

Task	FL	DP	TP
0	1	1	1
1	1	1	1
2.1	1	1	1
pause 2.2	1	1	1
pause 2.1	1	1	2
artist 2.2	1	1	3
artist 3.1	1	1	1
3.2	1	1	1
4	1	1	2
5.1	1	1	1
5.2	1	1	1
6.1	1	1	2
6.2	1	1	3
7.1	1	2	1
7.2	1	1	1
8.1	2	1	2
8.2	2	1	2
9.1	2	2	3
9.2	1	1	1
10.1	1	1	1
10.2	1	1	1
12	1	2	2
13.1	1	1	2
13.2	2	1	3
14	1	2	2

Table D.2: Test participant 2

Task	FL	DP	TP
0	1	1	1
1	1	1	2
2.1	1	1	1
pause 2.2	1	1	1
pause 2.1	1	2	2
artist 2.2	2	2	2
artist 3.1	1	2	1
3.2	1	1	1
4	1	1	1
5.1	1	2	1
5.2	1	1	1
6.1	1	1	1
6.2	1	1	1
7.1	1	2	1
7.2	1	1	1
8.1	1	1	1
8.2	1	2	2
9.1	1	1	2
9.2	1	1	1
10.1	1	1	1
10.2	1	1	1
12	2	1	1
13.1	1	3	1
13.2	1	2	1
14	1	1	1

Table D.3: Test participant 3

Task	FL	DP	TP
0	1	1	1
1	1	1	1
2.1	1	1	1
pause 2.2	1	2	1
pause 2.1	1	1	2
artist 2.2	2	2	1
artist 3.1	1	1	1
3.2	1	1	1
4	1	2	1
5.1	3	3	2
5.2	2	1	1
6.1	2	2	1
6.2	1	1	1
7.1	1	1	2
7.2	1	1	1
8.1	1	1	1
8.2	2	1	2
9.1	3	2	3
9.2	1	1	1
10.1	1	3	1
10.2	2	1	1
12	2	3	3
13.1	2	3	2
13.2	3	2	2
14	1	1	1

Table D.4: Test participant 4

Task	FL	DP	TP
0	2	2	1
1	1	1	1
2.1	1	2	1
pause 2.2	1	2	1
pause 2.1	1	2	1
artist 2.2	2	2	1
artist 3.1	1	3	1
3.2	1	1	1
4	1	1	1
5.1	1	1	1
5.2	1	1	1
6.1	1	2	2
6.2	1	1	1
7.1	2	2	2
7.2	1	1	1
8.1	1	1	1
8.2	2	1	2
9.1	1	2	2
9.2	1	2	1
10.1	1	1	1
10.2	1	2	1
12	3	1	2
13.1	1	2	1
13.2	2	3	2
14	1	1	1

Table D.5: Test participant 5

Task	FL	DP	TP
0	1	1	1
1	1	2	1
2.1	1	1	2
pause 2.2	1	2	1
pause 2.1	1	1	1
artist 2.2	1	2	2
artist 3.1	1	1	1
3.2	1	2	1
4	1	2	1
5.1	1	2	2
5.2	1	1	1
6.1	1	2	1
6.2	1	2	2
7.1	1	3	2
7.2	1	1	1
8.1	1	1	2
8.2	1	1	2
9.1	1	3	2
9.2	1	1	1
10.1	1	1	1
10.2	1	1	1
12	1	2	2
13.1	1	1	1
13.2	1	1	1
14	1	1	2

Table D.6: Test participant 6

Task	FL	DP	TP
0	1.17	1.17	1
1	1.17	1.33	1.17
2.1	1.17	1.33	1.33
pause 2.2	1	1.67	1.17
pause 2.1	1	1.5	1.5
artist 2.2	1.67	1.83	1.83
artist 3.1	1	1.5	1
3.2	1	1.17	1
4	1.17	1.67	1.33
5.1	1.33	2	1.67
5.2	1.17	1.17	1
6.1	1.17	1.67	1.5
6.2	1	1.33	1.5
7.1	1.17	1.83	1.67
7.2	1	1	1
8.1	1.17	1	1.33
8.2	1.5	1.17	2
9.1	1.5	2	2.5
9.2	1	1.17	1
10.1	1	1.33	1
10.2	1.17	1.33	1.17
12	1.83	1.83	1.33
13.1	1.17	1.83	1.33
13.2	1.67	1.67	1.67
14	1	1.17	1.5

Table D.7: Summary average

E

Results of eye glance behaviour from user testing of hypotheses

The following appendix contain the results of users' *eye glance behaviour* (driver distraction) during the user testing of the design hypotheses.

E.1 Individual scores

The tasks presented in the following appendix are the same as found in Appendix D, section **D.1 Tasks**.

EG = Eye glance
EGD = Eye glance duration

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	2.32	2	0	1.16
1	2.48	3	0	0.83
2.1	2.12	2	0	1.06
2.2	6.6	6	0	1.10
3.1	1.32	1	0	1.32
3.2	1.96	3	0	0.65
4	6.92	8	0	0.87
5.1	6.72	5	0	1.34
5.2	3.16	3	0	1.05
6.1	9.28	9	0	1.03
6.2	2.88	3	0	0.96
7.1	3.32	3	0	1.11
7.2	1.92	2	0	0.96
8.1	0.92	1	0	0.92
8.2	2.2	2	0	1.10
9.1	9.32	10	0	0.93
9.2	2.64	3	0	0.88
10.1	2.36	2	0	1.18
10.2	9.2	6	1	1.53
12	8.84	8	0	1.10
13.1	4.36	4	1	1.09
13.2	6.32	7	0	0.90

Table E.1: Test participant 1

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	1.76	1	0	1.76
1	2.6	3	0	0.87
2.1	3.72	3	0	1.24
2.2	5.44	5	0	1.09
3.1	2.28	3	0	0.76
3.2	2.52	5	0	0.50
4	7.57	11	0	0.69
5.1	4.315	5	0	0.86
5.2	3.68	4	0	0.92
6.1	6.111	5	0	1.22
6.2	5	4	0	1.25
7.1	6.2	6	0	1.03
7.2	5.8	5	0	1.16
8.1	5.4	5	0	1.08
8.2	4.2	5	0	0.84
9.1	5.28	5	0	1.06
9.2	4.12	3	0	1.37
10.1	1.68	1	0	1.68
10.2	2.44	2	0	1.22
12	13.76	12	1	1.15
13.1	5.4	6	0	0.90
13.2	10.44	10	0	1.04
14	6.56	6	0	1.09

Table E.2: Test participant 2

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	1.48	1	0	1.48
1	2.68	3	0	0.89
2.1	3.04	4	0	0.76
2.2	10.88	13	0	0.84
3.1	2.24	3	0	0.75
3.2	3.24	3	0	1.08
4	3.76	4	0	0.94
5.1	8.28	6	0	1.38
5.2	4.12	3	0	1.37
6.1	3.92	4	0	0.98
6.2	3.2	2	0	1.60
7.1	4.4	4	0	1.10
7.2	2.6	3	0	0.87
8.1	6.32	5	0	1.26
8.2	6.24	4	1	1.56
9.1	5.2	3	0	1.73
9.2	3.76	3	0	1.25
10.1	2.4	2	0	1.20
10.2	4.16	3	1	1.39
12	10.12	8	1	1.27
13.1	4.84	12	2	0.40
13.2	9.56	7	2	1.37
14	4.76	3	0	1.59

Table E.3: Test participant 3

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	1.6	2	0	0.80
1	1.08	1	0	1.08
2.1	3.32	3	0	1.11
2.2	5.08	3	1	1.69
3.1	0.68	1	0	0.68
3.2	0	0	0	0.00
4	3.78	3	0	1.26
5.1	5.24	3	1	1.75
5.2	4.4	3	1	1.47
6.1	4.68	3	1	1.56
6.2	5.48	3	2	1.83
7.1	4.32	3	0	1.44
7.2	3.44	2	1	1.72
8.1	2.06	1	1	2.06
8.2	3.48	2	1	1.74
9.1	7.36	4	1	1.84
9.2	2.76	2	0	1.38
10.1	2.88	2	0	1.44
10.2	3.36	2	0	1.68
12	13.2	12	0	1.10
13.1	7.12	4	2	1.78
13.2	10.2	5	2	2.04
14	2.52	3	0	0.84

Table E.4: Test participant 4

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	2.08	3	0	0.69
1	3.88	2	1	1.94
2.1	2.04	2	0	1.02
2.2	4.28	4	0	1.07
3.1	2	1	1	2.00
3.2	1.32	5	0	0.26
4	3.08	3	0	1.03
5.1	4.88	3	0	1.63
5.2	4.04	4	0	1.01
6.1	4.52	3	1	1.51
6.2	3.32	3	0	1.11
7.1	4.6	3	1	1.53
7.2	3.32	3	0	1.11
8.1	1.8	2	0	0.90
8.2	2.28	3	0	0.76
9.1	7.96	7	0	1.14
9.2	2.92	3	0	0.97
10.1	2.2	2	0	1.10
10.2	2.4	1	1	2.40
12	16.92	15	1	1.13
13.1	3.92	4	0	0.98
13.2	10.64	10	0	1.06
14	3.24	3	0	1.08

Table E.5: Test participant 5

Task	Total EGD	#EGs	#EGs >= 2 sec	Mean EGD
0	1.2	1	0	1.20
1	3.76	3	1	1.25
2.1	4.8	4	0	1.20
2.2	5.946	7	0	0.85
3.1	3.08	2	0	1.54
3.2	1.6	1	0	1.60
4	4.16	5	0	0.83
5.1	12.04	8	2	1.51
5.2	3.56	3	1	1.19
6.1	3.76	3	0	1.25
6.2	6.36	5	0	1.27
7.1	5.24	4	0	1.31
7.2	2.44	2	0	1.22
8.1	3.44	3	0	1.15
8.2	2.96	3	0	0.99
9.1	11.32	8	1	1.42
9.2	3.08	3	0	1.03
10.1	2.68	2	0	1.34
10.2	3.28	2	1	1.64
12	8.84	7	1	1.26
13.1	4.2	4	0	1.05
13.2	6.88	4	1	1.72
14	9.36	6	0	1.56

Table E.6: Test participant 6