Towards safe use of general controls in cars

A real-world driving study assessing internal HMI task frequencies and influencing factors

By

Daniël Auke Maria Auerbach

in partial fulfilment of the requirements for the degree of

Master of Science
in Transport, Infrastructure and Logistics

at the Delft University of Technology,
to be defended publicly on Tuesday, April 11, 2024 at 14:00 PM.

Thesis committee:	Prof. dr. M.P. Hagenzieker,	Chair
	Dr. I.M. Harms,	RDW
	Dr. E. Papadimitriou,	TU Delft
	Dr. J.A. Annema,	TU Delft

TU Delft RDW

Preface

This thesis was written in collaboration with the TU Delft and RDW to obtain a Master of Science degree in Transport, Infrastructure and Logistics. I would like to thank my supervisors including Marjan Hagenzieker, Jan Anne Annema, Ilse Harms and Eleonora Papadimitriou for providing me with great feedback on my work, which helped me take it to a higher level. I would also like to thank RDW for providing me with the Seat Toledo that was used during the observations. Working on my thesis at RDW helped me gain more experience in how these types of companies operate and allowed me to be a part of it. Writing this thesis and performing the research, I have also learned a lot about how to conduct such a study and how to report on it. By reading many studies about passenger car HMIs and how people interact with them, I have learned a lot about this topic. I would also like to express my gratitude to all the people who have participated in this study and to the people who shared my research with others. Without them, it would have been impossible for me to carry out this research. I would like to dedicate this Master thesis to my Uncle Walter and my Grandmother, who sadly passed away during the time of writing this thesis. Unfortunately, they were not able to participate in this study anymore. I also want to dedicate this thesis to my parents who have allowed me to do this study. After studying for my whole life up until this point, it is finally time to use the knowledge that I have gained and do something good with it.

Summary

In the literature, it is estimated that 5-25% of car crashes occur because of some form of distracted driving. In the Netherlands, estimates of 24-31% were found. This shows that distraction is an important factor in crash risk. Operating internal human-machine interfaces (HMIs) in passenger cars could be a form of distraction. In recent years, these HMIs have become more complex, with touch screens replacing physical buttons and more options becoming available. These minimalistic touchscreen designs often result in multiple layers of menus that need to be navigated to perform certain tasks. As shown in the literature, this can increase total task durations and distract people for longer periods. It would be beneficial to have the most performed tasks easily available to reduce the time that people are distracted. Introducing regulations could be a way to reach this goal. As this topic is new in regulations, a first step to pave the way towards regulations could be including the topic in the vehicle assessments of a renowned consumer organisation such as Euro NCAP. This organisation together with RDW want to address this issue and set new standards. Euro NCAP wants to do this by using its star rating system. This system rates the level of safety of different aspects of cars and per 2026 this rating is expected to also include the safety of car HMIs. These safety ratings are then used to influence car manufacturers.

However, knowledge is missing to be able to address this issue. It is not clear which tasks are the most important to assess, how often different tasks are performed during driving and what factors (socio-demographic, contextual, car familiarity, car characteristics) could influence this. It is therefore not possible to accurately assess the safety level of different HMIs in cars. This thesis aims to study how often people perform different tasks while driving and what factors might influence this. In particular, the impact of car familiarity was looked at. This aids in understanding drivers' behaviour with HMI during real-world driving. For this study, the choice is made to only look at tasks that are performed using the internal HMI of a passenger car. Other distracting activities such as using a phone were not studied. This thesis thus aims to answer the following research question: "What are the most performed internal HMI tasks by drivers in passenger cars and what factors (socio-demographic, contextual, car familiarity, car characteristics) influence the frequency with which tasks are performed"?

In this thesis, multiple methods, of which observing drivers in real-life circumstances was the key method, are used to fill this knowledge gap. First, literature research was performed to study which tasks are relevant to assess. Based on this, the decision has been made to only include tactical tasks. In this thesis, tactical tasks are defined as tasks that are not performed fully automatically by the participants and take some seconds to perform. Performing these tasks can be distracting. Examples of these tasks are changing the radio volume, adjusting the climate control temperature and changing the windshield wiper interval. Pilots were included to check if all relevant tasks were noted. Also, based on the literature, some factors have been identified that may influence how often tasks are performed. These factors include age, gender, driving experience, weather, traffic density and car familiarity. In the end, this resulted in an observation checklist that was brought along by the researcher when doing observations.

The number of times tasks were performed during driving was counted manually. To address car familiarity, the participants drove two trips: 1) with their own car and 2) with an unfamiliar car provided by the researcher. Since the cars that the participants brought varied, the characteristics of the cars were noted. The observations resulted in a good variety of cars with different characteristics. The car that was provided was a Seat Toledo from 2014 and acted as the unfamiliar car for 28 out of the 30 participants. All participants started with the same settings for the unfamiliar car to eliminate any effects on the task frequencies due to different initial settings. Due to logistical issues, two participants drove another car for their unfamiliar car trip. Trips were counterbalanced: about half of the participants started by driving their own car while the other half started with the unfamiliar car. The observations were performed from the 29th of November 2023 until the 8th of January 2024 in the Netherlands.

The participants were not informed about the real reason for observation and they were asked about their age, gender and driving experience. They were also asked to come up with a route that was familiar to them. They were told that the total trip time would have to be less than 30 minutes but more than 15

minutes. A trip was considered familiar when a participant rated the familiarity as 8 or higher on a 10-point scale.

During the trips, the researcher was sitting in the passenger seat noting down every time that a task was performed, as well as on what road type and in what situation the task was performed. As soon as the car moved for the first time, the counting started. The counting was performed following a specific method. First of all, a cooldown timer was used. This means that when a task was performed twice within 5 seconds, it was counted as only completing one task. The idea behind this is that the participant did probably not complete the task the first time. However, turning something on and off was still regarded as two separate tasks and was therefore counted as completing two tasks. During the observations, the participants were asked to do everything they normally do and to refrain from engaging in any conversations with the researcher. All trips lasted between 15 and 32 minutes and took 22.4 minutes on average. Most of the driving time was spent on roads in the city.

After driving both trips, the participants were asked a few more questions. First, they were asked if they could describe what the researcher was looking at. This was done to find out whether the participants already had an idea of what the study was about. Then, they were asked if they felt that the researcher had influenced their behaviour during the trip. Most people could not correctly describe what the researcher was observing while they drove their routes and did not feel like their behaviour was influenced. Lastly, their familiarity with both cars was measured. This was done by reading several statements to the participants. The participants stated whether they agreed or disagreed with each statement on a scale of 1-10. One statement aimed to measure overall car familiarity while four other statements aimed to measure different aspects of car familiarity. After the questions, some characteristics of the participant's car were noted, such as the transmission type and whether or not the HMI of the car was able to perform all tasks included in the study. Data analysis consisted of descriptive data analysis, paired t-tests, ANOVA's and Poisson regression analyses.

To assess which statements contributed to the construct of car familiarity, a factor analysis was also performed. This showed that three out of the four statements measure the same underlying factor. It seems that the statements "I know the features/options that the car has available", "I know where the buttons are located" and "I understand the dashboard and the things it displays" indeed measure one underlying factor, which was called perceived car HMI knowledge and was used in the Poisson regression models. The statement "I drive this car a lot" was found to measure another factor.

The main result of the study is that the most frequently used tasks fall within the following 4 categories: Lights, windshield, radio and media and climate control. Specifically, these most used tasks include the use of the indicator light, using the front windshield wipers, adjusting the volume, moving the sun visor, using the rear windshield wiper, adjusting the temperature of the climate control system and adjusting the fan speed. Car familiarity seems to have an impact on the average frequency with which different tasks are performed. People used more climate control tasks in familiar cars, while in unfamiliar cars, people more often adjusted the setup. This included adjusting the mirrors as well as adjusting the seat position and these tasks were even observed while driving. In addition to this, it was revealed that in the unfamiliar car - due to an unintuitive design from the Seat Toledo - the front windshield wiper was accidentally used excessively. When correcting the data for the aforementioned outlier by excluding the front windshield wiper task, it showed that drivers overall used more tasks in familiar cars compared to unfamiliar cars.

Perceived familiarity with a car HMI seems to increase the total frequency with which tasks are performed. This factor seems to increase the frequency of radio and media tasks the most. People who are more familiar with the specific HMI seem to perform more tasks. Male drivers also seem to perform more radio and media tasks, whilst female drivers seem to perform more windshield and climate control related tasks. Furthermore, age seems to influence the frequency of radio and media tasks. Older people in general performed fewer tasks in this category. Sunny and rainy conditions increase the frequency with which windshield tasks are performed due to higher use of the windshield wipers and sun visor respectively. Lastly, people seem to perform less climate control and windshield tasks when driving in more modern cars.

By using the method that was proposed in this study, it was possible to measure the prevalence of different internal HMI tasks. The prevalence of different forms of distraction is not well-studied in the current literature. This is especially true for internal HMI task prevalence. This thesis contributes to the literature by identifying the frequencies with which different tasks are performed and identifying which factors could influence this. It thereby increases the understanding of internal HMI use in cars. This study also provides a methodology that can be used by future studies on the topic of distraction prevalence. The differences in frequencies between different tasks are relatively large and even with such a small sample size, statistically significant differences were found between the top 10 most used tasks. In general, people seem to perform around 12 tasks per hour in familiar cars versus 9 tasks per hour in unfamiliar cars when excluding the front windshield wiper and indicator light tasks. This shows that there are a significant number of distraction moments per hour due to internal HMI use. Future research on the safety impact of performing internal HMI tasks can use the results of this study to better understand the relative risks of the different tasks in real-world driving.

The results of this study can be used as input for assessment protocols targeting the HMI of passenger cars, such as Euro NCAP's draft assessment protocol for the safe use of general controls, and eventually vehicle regulations. The recommendation is to encourage manufacturers to have the most frequently used tasks, which are the indicator light task, tasks in the windshield category and tasks in the radio and media and climate control categories easily available since this could reduce distraction. These tasks should therefore not be several levels deep in menus and should be able to be performed within a short time frame. Also, it would be a good idea to standardize how these tasks are performed, as this study shows that unintuitive designs can create extra distractions. This need for standardization has been mentioned in multiple other studies as well. Another recommendation is that the most frequently used tasks should be weighed as more important for the safety level and should be taken into account when assessing the safety ratings. Lastly, it is recommended that these safety ratings are shared with the consumers and car industry, such that car manufacturers have a greater stimulus to improve the safety of their HMIs.

Future research could focus on extending this study by using a bigger sample and by tracking people for longer periods. This can result in more accurate data for the population and for the individual participants. Also, cameras might be used to eliminate counting errors. It could create a more comfortable situation for the participants as well since some participants found it scary that someone was observing them and found it uncomfortable to ignore the researcher fully. Further research could also focus on the impact of other factors, such as psychological factors. This could be factors such as people's perceived aggressiveness during driving. The impact of these kinds of factors is also still missing in the literature. Lastly, more research is needed on the car familiarity scale, particularly about what is measured exactly by different types of questions and if these measurements are accurate.

Contents

1. Introduction

1.1 Distraction due to car HMIs

The human-machine interfaces (HMIs) in passenger cars have become more complex over the years, with touch screens replacing physical buttons and more options becoming available. These recent HMI designs might look clean and provide more options, but they also take longer to operate and can result in navigating through multiple layers of menus to perform simple tasks (ViBilägare 2022). The distraction that occurs while operating these more complex interfaces could impose problems regarding road traffic safety (Blanco et al., 2006), as a significant part of road accidents are caused by distracted driving. Research by the United States Department of Transportation (2023) has found that about 8% of traffic fatalities occur because of distracted driving. When looking specifically at the Netherlands, it is estimated that 24-31% of car crashes occur because of some form of distracted driving (Davidse et al., 2011). The number of traffic deaths in the Netherlands due to some form of distraction is estimated to range from about 20 to over 100 per year (Stelling & Hagenzieker, 2015).

One of the causes of distraction can be operating car HMIs. At the moment, there is no standardization for HMI design across cars (Ruzic, 2022), even though practice with an HMI reduces off-road glance time and therefore distraction (Broström et al., 2013). This lack of standardization is also mentioned as a problem in a recent study by Uhlving et al. (2023), as it can lead to safety hazards, particularly when driving in unfamiliar cars. This is backed up by current research, which finds that crash risk increases when driving unfamiliar cars (Lee et al., 2005; Dingus et al., 2016). This problem could increase due to a growing amount of people who are using services such as car sharing, where they are driving unfamiliar cars (Autodelen, 2022).

To face these problems, it is important to constantly improve the designs of these HMIs, such that they can be used safely. One way to do this could be to design cars in such a way that the most performed tasks are easily accessible, as this would reduce the time that people are distracted. Standardizing car HMIs could also be an option, such that people are more comfortable when operating HMIs in unfamiliar cars. Introducing regulations could be a way to reach this goal. RDW and Euro NCAP want to address this issue and set new standards (Euro NCAP, 2022). Euro NCAP intends to do this by using its star rating system (Euro NCAP, 2023). This system rates the level of safety of different aspects of cars and the goal is to also rate the safety of car HMIs. These safety ratings can then be used to influence car manufacturers.

1.2 Knowledge gap

However, as can be concluded from Chapter 2, knowledge to be able to address this issue is missing. Most studies in the current literature are focused on the impact of distraction on different performance measures. The impact of distraction on driving performance is therefore well studied. There is also a lot of research about innovative HMI technologies and how they might be used to reduce driver distraction. To the author's knowledge, no studies have looked at how often different HMI tasks are performed during real-world driving using the internal HMI of a car, which is in line with the findings by Stelling & Hagenzieker (2015), who also found that there is a lack of research on the prevalence of different types of distraction. Only two studies were found that looked into the frequency with which different tasks occur, but these studies did not make a distinction between different internal HMI tasks (Metz et al., 2014; Dingus et al., 2016). This poses problems for the goals set by RDW and Euro NCAP (Euro NCAP, 2022). It is not clear which tasks are most important to assess, how often different tasks are performed during driving and what factors could influence this. In particular, it could be interesting to look at the influence of car familiarity as car-sharing services are becoming more common.

1.3 Study objective

This thesis aims to contribute to the mentioned knowledge gap by investigating how often different tasks are performed using the internal HMI of a car and what factors might influence this. In particular, this study especially looks at the influence of car familiarity. To achieve this, manual observations are performed during real-world naturalistic driving. The results of this study can be used by Euro NCAP and RDW and can be used as input for new regulations. To reduce driver distraction, this new regulation could encourage manufacturers to create designs in which the most performed tasks are easily available and standardized across cars. The results of this study can also be used as input for Euro NCAP's star rating system to assess the safety of car HMIs. Tasks that are performed relatively often could be seen as more important when determining the safety ratings of interfaces. By publishing these ratings, manufacturers can be influenced to improve the design of their car HMIs, such that they can be used safely. These ratings can also help consumers to make better buying decisions. Ultimately, this should lead to safer car HMIs that reduce driver distraction in both familiar and unfamiliar cars. To reach this goal, multiple research questions have been constructed. These are discussed in the next section.

1.4 Main research question and sub-questions

To complete the objective of this study, the following research question is answered:

Main research question: What are the most performed internal HMI tasks by drivers in passenger cars and what factors (socio-demographic, contextual, car familiarity, car characteristics) influence the frequency with which tasks are performed?

To answer the main research question, it is decomposed into smaller sub-questions. The sub-questions that need to be answered are the following:

Sub-questions:

- What are relevant internal HMI tasks that should be considered?
- Which factors (socio-demographic, contextual, car characteristics) could be relevant in influencing the frequency with which internal HMI tasks are performed?
- Which internal HMI tasks are performed most frequently?
- What influence does car familiarity have on the frequency with which people perform internal HMI tasks?
- Which other factors (socio-demographic, contextual, car characteristics) influence the frequency with which people perform internal HMI tasks?

The first step to reach the objective of this study was to identify the relevant user interface tasks for which the frequencies had to be counted. The next step was to identify the factors that could be important in influencing the frequencies. Then, manual observations of real-world driving were performed. The number of times that different tasks were performed during driving was recorded. Based on this data, the frequencies were calculated and studied to find the most often performed tasks during driving. After studying the frequencies, the influence of different factors was investigated. These factors included socio-demographic factors, contextual factors, car familiarity, and car characteristics. Contextual factors are factors such as the weather during a trip or the time of day.

1.5 Research approach

To answer the research questions, different methods are combined. In Table 1, an overview can be found of the methods that are used. Figure 1 shows how the different methods are combined to answer the research questions. The first research question is answered by means of a literature study, performing a pilot and identifying common tasks that can be performed in popular passenger cars in the Netherlands. The pilot consisted of two test drives with the same participant. Based on the results acquired from these different study parts, an observation checklist has been created. This checklist contains a template for all tasks that need to be counted. It also contains all the questions that have been asked to the participants and the factors that were noted. The second research question was also answered by doing literature research. The third research question was answered by using the data from the manual observations of real-world driving. Each participant was asked to drive two trips. The participants drove one trip in their car and one trip in an unfamiliar car. With this data, a descriptive data analysis was performed. The fourth and fifth research questions are answered by using multiple methods to analyse the data. These methods include Poisson regression analysis, paired t-tests and repeated measure ANOVA's. Also, since multiple indicators for car familiarity were included, a factor analysis was performed. This analysis was performed to explore the underlying structure of the multiple indicators and will be used as input for the Poisson regressions. In Chapter 3, a more detailed description of the methodology can be found.

Table 1: Research methods

Main research question: What are the most performed internal HMI tasks in passenger cars and what factors (Socio-demographic, contextual, car familiarity, car characteristics) influence the frequency with which tasks are performed?

Main method: Performing manual observations during real-world driving

Sub-question	**Method**
What are relevant internal HMI tasks that should be considered?	Literature research, piloting and identification of common tasks in popular passenger cars
Which factors (socio-demographic, contextual, car characteristics) could be relevant in influencing the frequency with which internal HMI tasks are performed?	Literature research
Which internal HMI tasks are performed most frequently?	Descriptive data analysis
What influence does car familiarity have on the frequency with which people perform internal HMI tasks?	Poisson Regression Analysis, paired t-tests, repeated measure ANOVA's and factor analysis
Which other factors (socio-demographic, contextual, car characteristics) influence the frequency with which people perform internal HMI tasks?	Poisson Regression Analysis

The figure below shows the study methodology in a way that visualises the input and output of each step. By combining the identification of common tasks in popular passenger car HMIs, conducting literature research and piloting, the relevant tasks that need to be considered are identified. Literature research is also used to identify the relevant factors. The identified relevant tasks and factors are the foundation for the observation checklist that is used to perform the manual observations. The data that was gathered by performing the manual observations is then used as input for the descriptive data analysis, the Poisson regression models, paired t-tests, repeated measure ANOVA's and a factor analysis. These methods are used to analyse the data and draw conclusions.

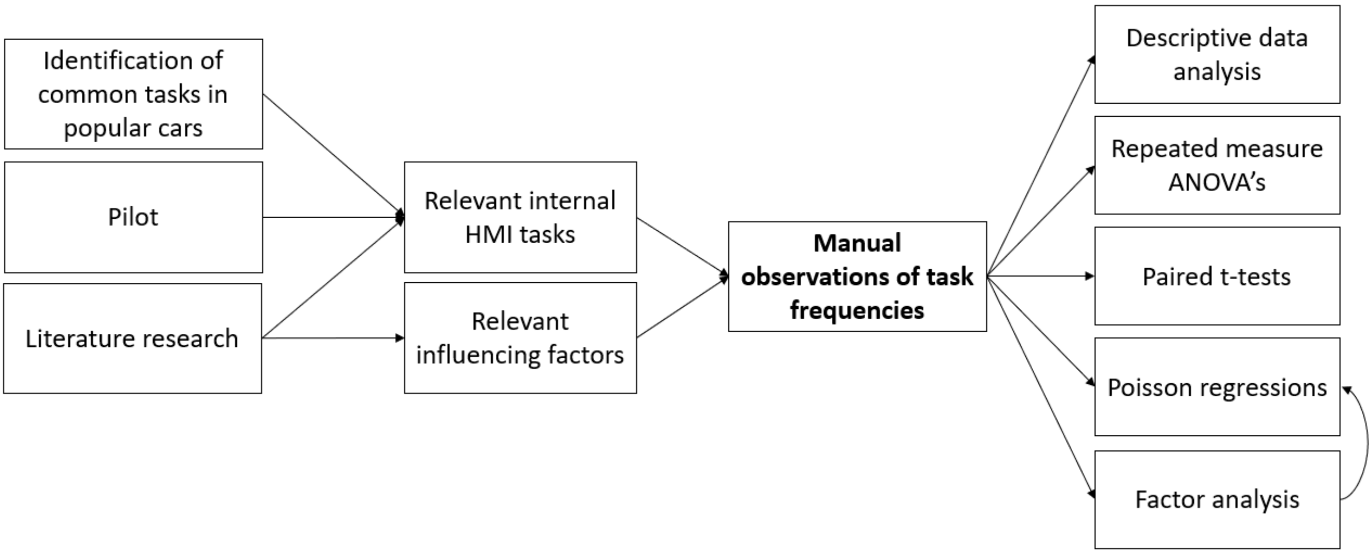


Figure 1: Study methodology

In the next chapter, the current state of the literature about car HMIs and distraction is studied. In Chapter 3, the methodology that is used for the study is presented in detail. After this, the results of the manual observations of task frequencies and the statistical analyses are presented in Chapter 4. Lastly, the results are discussed in Chapter 5 and conclusions and recommendations are presented in chapter 6.

2. Literature review

Due to the increasing complexity and importance of passenger car HMIs, an increasing amount of interest has emerged in this topic (Schmidt et al., 2010). Bad designs of automotive interfaces could have catastrophic consequences that will only be revealed once it is too late. This literature review focuses on manually driven passenger cars, as this form of driving is still the most prevalent with no imminent change in sight. (Kun, 2018). Also, most literature in the field of passenger car HMIs is regarding manual cars. This research defines manually driven as SAE level 2 or lower (SAE, 2021).

In this Chapter, a literature review is performed on the current state-of-the-art research regarding passenger car HMIs. This literature review was performed to find knowledge gaps. The first part explains the methodology that was used to find the relevant literature. The second part discusses the current literature regarding innovative technologies to improve the HMIs in passenger cars. After this, different performance measures used in the literature to assess driving safety are discussed. Then, the impact of different HMI tasks on safety is discussed and several knowledge gaps are described.

2.1 Literature review methodology

To find the relevant literature, a specific search strategy has been used. Google Scholar was used to find literature and a sorting based on relevance was used, which takes into account the number of times the search terms occur within the article, the area in the article where the search terms occur and the amount of citations (Universiteit Utrecht, 2022). This is one of the sorting options that can be selected in Google Scholar. At first, the focus was mostly on literature regarding the safety of human-machine interfaces in the automotive industry. Therefore, the keywords regarding the concept groups “Automotive industry”, “safety” and “Human-machine interface” were used. After this, literature regarding innovative technologies in human-machine interfaces was searched for. Therefore, the keywords in the concept groups “Automotive industry”, “Human-machine interface” and “Technology” were used. At last, forward and backward snowballing was used to complete the set of relevant literature. This resulted in the relevant studies that are discussed in the next section. Using the truncation with all the keywords results in 44.200 hits on Google Scholar. This search was performed in May of 2023.

No exclusion criteria were used for finding the literature. More recent research was prioritized if possible but some older studies were also still found relevant. The literature ranges from 1998 to 2022. Also, research that focussed on internal HMI tasks was prioritized during the search. However, some other studies about distraction and performance measures of other tasks are also relevant to understanding internal HMI task behaviour and possible consequences.

Table 2: Conceptual and methodological framework for literature review

Concept groups: Automotive industry; Safety; Human-machine interface; Technology		
Keywords	Automotive industry	Passenger cars, Road vehicles, Cars
	Safety	Performance measures, Assessment, Driver distraction, Driving performance
	Human-machine interface	User interface, Infotainment system, Information system, In-vehicle tasks
	Technology	Innovation, Touch screen
Truncation total		(Automotive industry) AND (Human-machine interface) AND (Safety) OR (Technology) *(44.200 hits)*
Truncation HMI tech		(Automotive industry) AND (Safety) AND (Human-machine interface) *(40.000 hits)*
Truncation HMI safety		(Automotive industry) AND (Technology) AND (Human-machine interface) *(30.000 hits)*

2.2 HMI technologies

With the increased functionality of passenger car HMIs, it becomes increasingly important to come up with intuitive designs that allow for fast and safe operations. To be able to improve these designs in order to enhance safety, it is important to look at the impacts of different technologies on driving performance. Table 3 shows literature that researched various innovative technologies that aim to reduce driver distraction during HMI use. Only technologies that are related to internal HMI are discussed.

Table 3: Studies regarding innovative HMI systems

Source	Task researched	Technology researched	Method
Lo & Green, 2013	Mix of user interface tasks	Speech interface	Literature research
Prabhakar & Biswas, 2021	Mix of user interface tasks	Virtual touch interfaces, wearable devices, speech recognition and non-visual interfaces and eye gaze-controlled systems	Literature research
Breitschaft et al., 2021	Target selection task	Haptic buttons	Driving simulator
Iqbal et al., 2011	Calling	Notifications and mediation	Driving simulator
Ma et al., 2022	Mix of user interface tasks	Centre stack buttons, touch screen, steering wheel buttons, and voice control	Driving simulator
Uhlving et al., 2023	Turning off lane keeping assist, fan speed to max, adaptive cruise control, changing temperature	Integrated screen	Real-world driving

Several innovative technologies for HMIs in cars have been researched. Iqbal et al. (2011) looked at the influence of alert notifications and mediation on calls while driving. This technology alerts the driver on upcoming critical road conditions and can mediate a call by putting it on hold during these conditions. The results indicate that this alerting technology can improve driving performance and was preferred by most of the participants. Furthermore, Prabhakar and Biswas (2021) looked at some advantages and disadvantages of virtual touch interfaces, wearable devices, speech recognition and non-visual interfaces and eye gaze-controlled systems. The findings indicate that wearable devices are not yet researched well as a way to interact with the user interface of a car. Eye gaze-controlled systems have to possibility to decrease the off-road glance time but at the moment such systems do not comply with automotive regulations. Furthermore, virtual touch systems are only researched using infrared at the moment.

Speech recognition and non-visual interfaces are more likely to have a positive influence on reducing driver distraction, as is also mentioned by Lo & Green (2013). This study performed a literature review and found that speech interfaces consistently improve driving performance across research experiments. However, the results regarding task completion time were mixed. Ma et al. (2022) also looked at voice control and found it to be a suitable technology for basic tasks and tasks of medium complexity. This study also found that touch screens can be more suitable for complex tasks than standard buttons. However, the latter seems to be more suitable for basic tasks.

Haptic feedback is another technology that could be implemented in the HMIs of cars. Breitschaft et al. (2021) looked at the influence of haptic feedback on driving performance. They found that haptic feedback did not have a significant influence on the driving task, but did improve the overall user experience.

Many cars today also include screens in which all or most tasks are integrated. Uhlving et al. (2013) studied the impact of these screens on driving behaviour to assess its potential effect on safety. To measure this, they performed drives on a test track and made use of an eye-tracking system. Results indicate that these integrated touch screens can be potentially dangerous.

2.3 Performance measures

In the current literature, different performance measures are used to evaluate the impact of driver distraction on safety. This section of the literature review looks into the different approaches used to evaluate the performance of HMIs in cars and the performance measures that are frequently used for driving tests. An overview of studies involving driving performance measures can be found in Table 4.

Table 4: Studies involving driving performance measures

Source	Performance measures	Method
Iqbal et al., 2011	Turning errors, collisions, number of missed red lights and stop lights, lane departures, cognitive performance	Driving simulator
Strayer et al., 2017	Visual and cognitive demands	Real-world driving
Green, 1999	Task completion time	Literature research
Green, 1998	Lane departures, task completion time, off-road glance duration	Literature research
Cooper et al., 2020	Visual and cognitive demands and task completion time	Real-world driving
Chisholm et al., 2008	Response time, number of collisions, steering wheel angle variation, off-road glance duration and frequency and task completion time	Driving simulator
Ou et al., 2013	Number of hand movements, Total task time, Number of glances, glance time duration, longitudinal velocity and acceleration, steering wheel angle	Driving simulator
Ružić, 2022	Findability, reachability, identification, interpretability, operability and understandability	Heuristic approach
Broström et al., 2013	Off-road glance duration	Driving simulator
Wang et al., 2010	Standard deviation of velocity, lane position, horizontal gaze centralization, and pupil diameter. Heart rate, skin conductance.	Driving simulator
Lee et al., 2005	Crashes, near-crashes, incidents, longitudinal and lateral kinematic information, headway, lane-keeping behaviour, risk ratio	Real-world driving

2.3.1 Criteria-based approach versus driving tests

One of the studies that investigated HMI systems in cars was performed by Ružić (2022). This study is different from the others mentioned in Table 4 since a heuristic approach was used to assess the performance. This means that the researcher analysed and discussed the internal HMI of a typical passenger car based on 5 criteria that are important from an ergonomic and safety point of view. These criteria include findability, reachability, identification, interpretability, operability and understandability. The focus of this research was also only on the climate control panel. The findings of this study indicate that the cars that were investigated lacked standardization regarding the layout and that there was too much cluttering of the controls. Also, many controls required visual guidance for correct operation. This is one of the few studies that did not include either real-world or simulator driving tests to assess the performance of the HMI. Instead, this study looks at a few different criteria and assesses the HMI based on these criteria.

An advantage of this type of desk research is that the results do not depend on the participant driving the car. In driving tests, the outcomes of the test are very dependent on the person that is driving. Different people might have different competence levels in using the HMI of a certain car. A solution to this problem would be to include a large number of people in the driving tests such that a more representative result for the population is obtained.

Most studies that look into performance measures include either real-world driving or driving in a simulator. Out of the eight studies that included driving, five used driving simulators, whilst the other three were based on real-world driving. The advantage of driving simulators is the fact that they can simulate all kinds of scenarios in the same way for every participant. Also, it is more safe to test in driving simulators than in real cars. In driving simulators, more difficult and demanding situations can be simulated such that collisions are more likely to happen. In real-life tests, this could result in safety issues. However, real-world driving has the advantage that the results are likely to be more ecologically valid and realistic since real-world situations and physics apply. Driving in a driving simulator is still not the same as driving in the real world (Wynne et al., 2019).

2.3.2 Performance measures used for driving tests

One way to measure driving safety is to take into account the number of collisions. However, this is not a very suitable indicator for testing due to the low rate of collisions in both real-life trials and most driving simulator trials (Kun, 2018). Therefore, the common approach is to use performance measures that are known to increase crash risk. A study by Chisholm et al. (2008) is one of the few that took into account the number of collisions. This study made use of a driving simulator and only one study using real-world driving that included the number of collisions was found (Lee et al., 2005).

One of the most frequently mentioned performance measures in the literature is the time to complete a task (Green, 1999; Green, 1998; Chisholm et al., 2008; Ou et al., 2013; Cooper et al., 2020). In the past, studies have suggested a 15-second rule, meaning that the time to complete a task should not exceed 15 seconds (Green, 1999). The time to complete a task is an important indicator of crash risk. It is also relatively easy to measure the time it takes to complete a task. Other measures like cognitive and visual demand are harder to measure. These measures look at the cognitive and visual workload of drivers. A study by Green (1998), showed that total eyes-off-the-road time is a predictor of crash risk and that this measure had a high correlation with task completion time. Ružić (2022) also mentions that HMIs can be tested by analysing the time required for certain tasks by using driving simulators. Chisholm et al. (2008) looked at the time it took to complete different iPod interactions whilst Ou et al. (2013) looked at the time it takes to complete several other types of tasks. Research also suggests that older drivers have increased task completion times (Cooper et al., 2020).

Two other important performance measures that are used to evaluate the safety of HMIs in cars are visual and cognitive demand (Strayer et al., 2011; Iqbal et al., 2011; Cooper et al., 2020). To measure this, participants usually have to perform visually and cognitively demanding tasks while driving. First, a baseline score on these tasks is measured. Then, the participants have to engage in HMI tasks while performing the tasks that are used to measure the demand. By comparing the two scores, it becomes clear how demanding the HMI task was. A study by Strayer et al. (2017) included a real-world driving experiment measuring visual and cognitive demands for different infotainment tasks in selected 2017 passenger cars. These tasks involved audio entertainment, calling, navigation and text messaging. The findings indicate that there were differences in visual and cognitive demands between the selected tasks. The most demanding task was destination entry for navigation. Voice-based commands demanded lower levels of visual demand but increased total interaction time significantly. There was also a significant difference in demand across the different vehicles. The researchers stated that many of the tasks are too distracting and dangerous to use while driving. This was especially the case for navigation entry. Based on their testing regarding visual and cognitive demands an overall demand level per vehicle is determined on a four-point scale ranging from very low to very high demand (*AAA Center for Driving Safety and Technology*, n.d.). In addition to this, research suggests that older drivers have more difficulty operating infotainment systems in cars resulting in a higher cognitive and visual workload (Cooper et al., 2020). In general, ISO standard tasks are used to determine cognitive and visual demand. Iqbal et al. (2011) took a different approach in determining the cognitive load of drivers. They indirectly assessed the cognitive load by looking at the number of missed red lights and stop signs. In addition to this, they also looked at the number of lane departures.

Off-road glance duration is yet another frequently used performance measure (Chisholm et al., 2008; Ou et al., 2013; Broström et al., 2013). A study by Broström et al. (2013) made use of a driving simulator to determine off-road glance duration intervals of 30 participants when performing different tasks within the HMI. The results of this study indicate that off-road glance durations while performing these tasks vary significantly between different people. It also indicates that practice reduces the off-road glance time and that individual glance strategies play a big role in glance duration regardless of in-vehicle task complexity, which is an interesting finding that may contradict earlier beliefs (Broström et al., 2013). Furthermore, Ou et al. (2013) created a risk prediction model. This model tries to predict the risk condition based on different performance measures. Three risk conditions were constructed by the researchers based on the steering wheel angle ranging from safe to hazardous. They found the number of glances to be the most important predictor for the risk condition of in-vehicle information system tasks. The longitudinal velocity was found to be the most important predictor for the risk condition of the traditional in-vehicle tasks. Velocity is also included in a study by Wang et al. (2010). This study indicated that physiological measures such as heart rate and skin conductance are more sensitive to changes in workload than driving performance measures such as the standard deviation of velocity and lane position. Lee et al. (2005) also used velocity as well as lane position as a performance measure. Some other performance measures that can be found in literature include turning errors, response time and number of hand movements (Iqbal et al., 2011; Ou et al., 2013).

As can be seen from the literature, the performance measures used to assess driving performance widely vary between studies. A literature review by Papantoniou et al. (2017) also mentions this and states that there is not a single driving performance measure that can accurately describe all distraction aspects. Most studies do show that distracted driving negatively influences a variety of these performance measures. Furthermore, Chisholm et al. (2008) show that repetition of using an HMI increases driving performance. This should be taken into account when testing vehicles with drivers who are inexperienced with that particular vehicle.

2.4 Safety impact of internal HMI tasks

In the previous section, different methods and performance measures in the literature were investigated. This section focuses on the safety impact of different tasks using these performance measures. Nowadays more internal HMI tasks than ever before can be performed in passenger cars. In literature, the safety impact of performing different tasks has been studied. An overview of some studies that have researched common HMI tasks in passenger cars can be found in Table 5.

Table 5: HMI tasks and safety

Source	Task	Impact
Redelmeier & Tibshirani, 1997	Phone calls	Four times higher collision risk
Strayer et al., 2017	Audio entertainment, calling and dialling, text messaging, and navigation	Increased visual and cognitive demand. Destination entry in navigation biggest increase.
Perez et al., 2013	Radio tuning	Increased total glance time to task, total eyes-off-road time, glance rate and duration of longest glance
Blanco et al., 2006	Different types of navigational tasks	Decreased driving performance when complexity of task increases.
Horberry et al., 2006	Mobile phone conversation, radio tuning	Decreased driving performance
Ma et al., 2022	Music, Air conditioning, Phone call, navigation	Decreased driving performance depending on the type of modality used to perform task
Dingus et al., 2019	Primarily cognitive tasks including radio adjusting	Performing these types of tasks result in low crash risk

Looking at the literature that studies the effect of different tasks on driving performance and cognitive demand, it becomes clear that certain tasks are more often mentioned than others. HMI tasks that are often mentioned are making phone calls, radio tuning and setting up navigation. Of the seven studies that were identified, five mentioned the impact of phone calls on driving performance or driver distraction (Redelmeier & Tibshirani, 1997; Strayer et al., 2017; Horberry et al., 2006; Ma et al., 2022; Dingus et al., 2019). Three studies mentioned the effects of navigation tasks (Strayer et al., 2017; Blanco et al., 2006; Ma et al., 2022) and five studies mentioned the effects of radio or media tasks (Strayer et al., 2017; Perez et al., 2013; Horberry et al., 2006; Ma et al., 2022; Dingus et al., 2019). The consensus found in the literature is that these tasks are found to decrease driving performance and increase cognitive and visual demand. Research regarding phone calls includes both hands-free calling as well as hand-held calling. Horberry et al (2006) found no significant safety advantage of hands-free over hand-held calling. Even though not all studies used the internal in-vehicle HMI for making phone calls, there is still a lot of similarity between making a phone call using the internal HMI and making a phone call using an actual phone. Therefore these studies are still useful to indicate the potential danger that comes along with it. Lastly, the type of demand that is placed on the driver seems to have an impact on the crash risk, as primarily cognitive tasks seem to result in low crash risk (Dingus et al., 2019).

The effects of more traditional in-vehicle tasks such as the operation of the air-conditioner and opening or closing windows are less common in literature. Ma et al., 2022 performed the only study that was found for this review in which the effects of operating the air conditioning were studied. This study found that different types of tasks are more or less suited to certain types of modalities, which were classified as knobs and buttons on the centre stack, touch screen, steering wheel buttons, and voice control. Knobs and buttons were found to be most suitable for basic tasks whilst the other modalities could be better for more advanced tasks.

Furthermore, Metz et al. (2014) performed a study investigating the frequencies of secondary tasks during naturalistic driving. They found that passenger-related distraction is a very big part of the total amount of distraction during driving. The driver was distracted in some form during 25% to 40% of the driving time. The most frequently used secondary tasks were found to be telephoning, internal HMI inputs and using the mobile phone. Dingus et al. (2016) performed a similar study and found that drivers engaged in distracting activities 52% of the time and engaged in internal HMI tasks 3.5% of the time.

2.5 Conclusion and discussion

In this literature review, multiple aspects of HMIs in cars have been discussed. The following distinction between three different types of research regarding this topic has been made: 1) Research based on innovative technologies that are used in car HMIs; 2) Research based on the methods and performance measures that are currently used to evaluate driving performance; 3) Research on the impact of different tasks on driving performance.

Most studies in the current literature are focused on the impact of distraction on different performance measures. The impact of distraction on driving performance is therefore well studied. There is also a lot of research about innovative HMI technologies and how they might be used to reduce driver distraction. However, there are no studies yet that have created an assessment methodology to assess the overall safety levels of a complete HMI. The current literature includes a low number of tasks and studies their impact on driving performance measures. However, no effort has been made to evaluate complete HMIs and to be able to compare them on the same scale. This makes it difficult to compare different HMIs and to assess potential safety effects when using them, which poses problems for creating regulations.

The foundation to create such an assessment method is still missing in the literature, as there is a lack of research on the prevalence of different internal HMI tasks (Stelling & Hagenzieker., 2015). Obtaining this information will help to identify which tasks are most important to take into account and to find the relative importance of each task for the assessment. Tasks that are performed more often can be seen as more important when assessing the safety of a car HMI. Creating this kind of assessment methodology could help evaluate the relative safety levels of HMIs in passenger cars and establish a ranking. It could also provide a stimulus for car manufacturers to create better HMI designs that are safer to use. Creating safety rankings of car HMIs could also influence the buying decisions of individuals. To get a better understanding of the effectiveness of such a ranking system on buying decisions, logit models can be used in a willingness-to-pay experiment.

Hardly any studies were found to investigate the frequencies of secondary tasks in driving except for two studies by Metz et al. (2014) and Dingus et al. (2016). However, these studies looked at all sorts of secondary tasks in vehicles including non-internal HMI tasks. Furthermore, they grouped internal HMI tasks under a limited amount of categories and in the study by Metz et al. (2014), all participants drove a car brand that they were familiar with. To the author’s knowledge, there has not been a study that has looked at the secondary task frequencies of the individual internal HMIs in cars and at the effect of car familiarity. Also, more effort is needed to understand if there are other determinants of HMI behaviour during real-world driving. To name an example, cars with touchscreens could inherently be less safe because they might invite more interactions. Understanding these kinds of relationships is important to be able to create well-considered regulations.

When looking at the literature, it can also be observed that some studies used older driving simulators. Over time, driving simulators have become increasingly realistic and can maybe produce more ecologically valid results since it is closer to real life. Therefore, it is important to keep studying the effects of interacting with HMIs in these newer driving simulators, particularly for scenarios that are difficult to find in real life or scenarios that are viewed as too unsafe for real-world testing.

The literature furthermore indicates that repeated use of an HMI enhances driving performance during HMI tasks. However, this research field is still lacking. Chisholm et al. (2008) found that repetition in performing tasks with an iPod resulted in better driving performance when performing these tasks. Broström et al. (2013) found that repetition with an HMI reduces off-road glance time during the performance of HMI tasks. However, no other studies were found that looked at the effect of repetition with an HMI on other performance measures. The literature also made clear that no single driving performance measure can capture all the effects of distraction, as also mentioned by Papantoniou et al. (2017). Therefore, it is not sufficient to use a single indicator for evaluating the safety of HMIs.

Geographical areas could also influence the results of the different studies mentioned. People in different areas of the world might have different levels of familiarity with technological systems. These kinds of geographical differences could change the results of the individual studies mentioned significantly. No studies conducted in the Netherlands were found that examined the impact of performing various HMI tasks on driving performance.

The following knowledge gaps have been identified:

- No assessment methodology to assess the overall safety levels of a complete HMI has been created.
- There is a lack of research on the prevalence of different tasks that are performed during driving with internal HMI tasks in particular.
- More studies on the effect of performing HMI tasks on driving performance could be performed using newer driving simulators.
- More studies on the effect of repetition with an HMI on driving performance can be performed using different performance measures.
- A study on the effect of performing HMI tasks on driving performance in the Netherlands is seemingly missing.

This thesis aims to contribute to the knowledge gap regarding the frequencies with which different internal HMI tasks are performed during driving in passenger cars. This gap has been chosen due to its immediate practical relevance. The frequencies can be used to create regulations and also to create an assessment methodology. The results of this kind of research could therefore aid Euro NCAP and RDW in reaching their goals (Euro NCAP, 2022). This thesis also looks into the factors that might influence the frequency with which tasks are performed. As car sharing is becoming more common (Autodelen, 2022), it is particularly interesting to look at the impact of car familiarity. Studying the impact of different factors improves the interpretation of the results and could be input for creating measures to reduce the frequency of HMI use.

3. Research methodology

In this chapter, the methodology that was used for this study is explained. First, the main aspects of the experimental design are discussed. After this, the procedures are explained. Then, information about the materials used for this study is presented and explained. Hereafter, the measures that are used to answer the research questions are discussed and information about the participants is presented. Lastly, the results of the pilot are discussed and the data analysis methods that were used to analyze the data are explained.

3.1 Experimental design

In a 2x1 within-subjects design, the number of times that drivers performed specific tasks during a naturalistic driving experiment was counted by driving along. To be able to quickly observe and count, an observation checklist was created that structured the different types of tasks intuitively into categories. The participants had to drive two trips: 1) In their own car and 2) in an unfamiliar car. About half of the participants started by driving their own car while the other half started with the unfamiliar car. The initial settings of the unfamiliar car were the same for all participants. The choice has been made to use grouped observations since the goal was to find the effect of car familiarity on the task frequencies. Keeping the personal characteristics constant therefore provides better results than having two different groups of people. A convenience sample was used for this study, which should suffice, due to the explorative nature of the study. Ethical approval was granted on November 16th 2023 by the Human Research Ethics Committee (HREC) of the TU Delft with application number 3679.

The observations during the trips were performed manually using the observation checklist. Performing observations during naturalistic driving is a good method for gathering data, as it can be seen as an ecologically valid method. Real behaviour is studied in settings that are as close to real life as possible. A downside of observations relative to other methods like surveys is the fact that it is more difficult to get a big sample size due to the extra time it takes to observe. It can also result in more errors. No video footage has been used for the observations. The advantage of video footage is that more observations can be done and that they can be rewatched. The downside is that factors regarding the context and weather are harder to observe. For the video footage to be efficient, multiple cameras would have been necessary. This would have resulted in high expenses which was unfeasible for this research.

Only familiar routes were driven during the trips since these are the routes that are mostly driven by people and therefore most representable for the average trip. Being familiar with a route means that participants perform their user interface tasks as they usually do. Using unfamiliar routes may introduce different results. A study by Harms et al. (2021) found that the familiarity of a route has a big impact on driving behaviour. One of the findings in this study is that increased familiarity resulted in lower task difficulty and increased engagement in secondary tasks. Since unfamiliar routes require more attention, fewer user interface tasks are likely performed compared to familiar routes. The route familiarity was self-reported by the participants on a scale from 1 to 10. A route is considered familiar when the participants score the familiarity higher than or equal to 8, which is a similar approach to Harms et al. (2019) and Burdett et al. (2019). The fact that different routes are driven increases the noise in the data but has the advantage that real-world trips are observed. Therefore, the choice has been made to use different familiar routes that are driven by the participants instead of experimental or controlled routes. However, to still be able to understand the contexts around the trip, some route characteristics are noted. The route length for each trip was less than 30 minutes to reach a maximum of 60 minutes per participant. This maximum was set since the budget for this study was limited and participants might not have been willing to drive longer than 60 minutes. These trip durations have resulted in distances travelled that are reasonably close to the average distance per car trip in the Netherlands, which is found to be 17.44 kilometres (van Heukelingen et al., 2023). However, this does depend on the type of road that was used, as the speed on highways is much higher than the speed on roads in the city.

Table 6 shows an overview of the experimental setup of the study, which is further elaborated on in the next sections.

Table 6: Overview of the experimental setup

Setup factors	**Choice**
Route choice	Familiar route (Harms et al, 2019; Burdett et al., 2019)
Route length	Maximum of 30 minutes
Seat choice for researcher	Front seat
Observation type	Manual observation
Conversation during trip	No speaking (Metz et al., 2014)
Car type	All types of cars
Type of tasks	Internal Human-machine interface tasks that are on the tactical level (Michon, 1985)
Groups for familiar/unfamiliar car	Within subjects
Counting method	Cooldown timer of five seconds for counting a new occurrence (Metz et al., 2014)
Frequency data depth	HMI Task Frequency per road type and whether standing still or not

3.2 Procedure

3.2.1 Recruitment procedure

The participants were asked if they wanted to participate in the study directly. In addition, they received a flyer which described some aspects of the study. This flyer also contains a QR code that can be scanned to book a time slot. For everyone willing to participate, one date was scheduled to drive the two trips. Only the first seven participants had a separate appointment for the second trip. This was due to some delay with acquiring the unfamiliar car.

When people scanned the QR code on the flyer, they were directed to a Google calendar webpage where they could book a time slot and enter their name and e-mail address. The researcher then contacted these individuals through their e-mail addresses to provide more information and arrange a location. The information that the participants received can be found in Appendix F. Participants could also contact the researcher right away by using the contact information on the flyer and sending an e-mail themselves. Before the observations began, the participants got some more instructions. These instructions can be found in the observation checklist in appendix C.

This way of recruiting participants meant that the researcher had access to people's e-mail addresses and names. These e-mail addresses are not shared with anyone, including RDW or TU Delft. Before the trips were driven, participants were shown a consent form, which they had to fill in. These consent forms were stored in a secure location and will not be shared. They contain only the name of the participant and a signature. The data is fully anonymised, meaning that the names cannot be linked to any data points. All 30 participants gave consent. The informed consent form can be found in Appendix F.

When contacting the participants, they were asked what car they would drive during the observations, so that the researcher was able to look at the tasks that the car in question can perform. This information was deleted as soon as the trips with that participant were finished and can also not be linked with the data points in any way. Participants were recruited before but also during the time in which the observations took place.

3.2.2 Experimental procedure

For the observations. the unfamiliar car was driven to the starting location by the researcher, which was usually the home address of the participants. For most participants, the two trips were driven during the same appointment. This was done to minimize weather and context differences. However, due to time constraints, 7 out of the 30 participants had to drive the two trips on separate days.

Before the trips, the participants were asked about their age, gender and driving experience. They were also asked to come up with a route that was familiar to them. They were told that the total time per trip would have to be less than 30 minutes but more than 15 minutes. The participants would then drive from the origin to the destination and back again in both cars. During the trips, the researcher was sitting in the passenger seat noting down every time that a task was performed. After driving both trips, the participants were asked a few more questions. Also, their familiarity with both cars was measured. After the questions, the characteristics of the participant's car were noted. It was noted whether or not the HMI of the car was able to perform the tasks included in the study and whether or not the car was equipped with adaptive climate control, adaptive cruise control, a touch screen, automatic lights and an automatic transmission.

The observations took place with the researcher sitting in the passenger seat of the car since this gave the best view for observing the tasks that a driver performed. The choice was also made to not engage in any conversations during the trip, as most trips with a car are made without passengers (CBS, 2022). Engaging in conversations can distract the driver and can result in the driver performing fewer secondary tasks than they would normally perform (Metz et al., 2014). It is still possible that the researcher influenced the results by driving along, as having a passenger might have resulted in different behaviour. However, for this research, it was the only viable way of observing. The observations were performed from the 29th of November 2023 until the 8th of January 2024.

3.3 Materials

3.3.1 Car fleet

The unfamiliar car for the study was provided by the researcher. 28 out of the 30 participants drove a Seat Toledo from 2014 as their unfamiliar car. The climate control temperature of this car was set to 19 degrees and the radio was set to radio 538 on volume 8. Also, all lights and the windshield wipers were turned off such that all participants started with the same settings. Two participants drove a different car as their unfamiliar car. For one of the participants, this was due to time constraints. The other participant who did not drive the Seat was only able to drive in cars with an automatic transmission, which the Seat did not have. The other two cars that were used as unfamiliar cars are a Renault Megane from 2012 and a Renault Modus from 2005. The characteristics of these cars can be found in Appendix E.

Not every car that was used has the same functions in its HMI. More advanced cars had more options than relatively older cars. Therefore, trips with more advanced cars were preferred over trips with older cars that have fewer options. To increase the chance of a larger sample size, the choice has been made to allow the use of all types of cars. The type of car and user interface that people used during the experiments might have influenced the task frequencies. Therefore, the type of car and all the tasks that are possible to be performed with its user interface are noted. Both cars with a manual transmission and automatic transmission are used during the trips. In the Netherlands, manually shifted cars are the most driven type but this is changing since most new cars are automatic (Bovag, 2022).

When looking at the car characteristics, a good variety of different types of cars can be found in the sample. Most cars did include adaptive climate control and automatic lights, whilst most cars did not include adaptive cruise control. About half of the cars were equipped with a manual transmission compared to an automatic transmission. Also, most cars included a touch screen. The Seat Toledo did have adaptive climate control but was not equipped with adaptive cruise control or automatic lights. The transmission of the Seat was not automatic and no touchscreen was present. Both other cars did not include a touch screen or adaptive cruise control. Furthermore, one of these cars was equipped with an automatic transmission whilst the other was equipped with a manual transmission. Information about the car fleet can be found in Table 7.

Table 7: Car fleet

Car fleet characteristics of participants				
Variable	**Category**	**Amount**	**Percentage**	**N**
Adaptive climate control	Yes	19	63.33	30
	No	11	36.67	
Adaptive cruise control	Yes	11	36.67	30
	No	19	63.33	
Automatic lights	Yes	21	70	30
	No	9	30	
Transmission	Manual	17	56.67	30
	Automatic	13	43.33	
Touch screen	Yes	20	66.67	30
	No	10	33.33	

RDW car characteristics (Seat Toledo 2014)	
Ad. Climate	Yes
Ad. Cruise	No
Auto Lights	No
Transmission	Manual
Touch screen	No

3.3.2 Observation checklist

To be able to quickly and accurately make observations, an observation checklist with predefined tasks was created. To construct this checklist, it was important to first identify the relevant tasks that had to be taken into account during the trips. To identify the tasks, passenger car observations have been made. Also, literature research has been performed. To test if most of the relevant tasks were included in the observation checklist also a pilot was performed. Manually observing the frequency of the tasks that are performed using an observation checklist can be seen as a task analysis. Task analysis is used often for studies using observation as a main method and can be widely used. Wu et al. (2012) used a task analysis to study the backing manoeuvre of a car. Here, they used it to decompose the tasks into separate sequences.

To better understand the types of tasks that are documented during this research, it is useful to look at some conceptual frameworks. Rasmussen (1983) constructed a model with three levels of performance. This framework helps to understand different types of behaviour. These levels of behaviour were named skill-based, rule-based and knowledge-based. In this model, skill-based behaviour is seen as behaviour that happens automatically. Rule-based behaviour is less automatic and is more based on rules which can be reported by a person. The highest conceptual level is knowledge-based behaviour, which happens during more unfamiliar situations. Since familiar routes are driven in both familiar and unfamiliar cars, tasks on all three of these levels were performed during the trips. This model also implies that certain tasks can change between the levels based on their familiarity.

A similar model but more focused on the driving tasks is the model of Michon (1985). Only tasks at the tactical level as described by Michon are looked at during the observations. In this model, tasks at the strategic level include the general planning stage of a trip. This includes tasks such as determining the mode and route. These tasks are not relevant to this study as they are mostly performed before a trip. At the tactical level, tasks such as lane changing and deciding gap acceptance are performed. The choices at the tactical level aid in realizing the choices made at the strategic level. Lastly, there is the operational level, where small adjustments are made to follow the tactical choices.

Since the interpretation of these models can vary across researchers and to make the model better fit this research, the choice has been made to slightly adapt the definitions. The definitions used in this research can be seen in Table 8. Following these definitions, during the car trips, only tasks at the tactical level have been documented, since these tasks are performed while driving, still require some decision-making and do not happen fully automatically. Since these tasks can be distracting and can reduce safety they are the focus of this research. Even though an increase in cognitive demand is not the same as a decrease in driving performance, the two are very much related. Wang et al. (2010) found that increased cognitive demand did have a significant negative influence on driving performance, even though changes in physiological measures showed greater sensitivity to increased cognitive demand.

To identify common tasks in modern cars and to construct the observation checklist, the user interface options of two very different car models were observed, which are both quite popular in the Netherlands (Bovag, 2023; Allianz Direct, 2023): the Kia Picanto (Figure 2) and the Tesla Model Y (Figure 3). The Kia Picanto is a more compact city car while the Tesla Model Y is a more advanced electric car.

Figure 2: Kia Picanto (Auto Express, 2023a)

As the Kia Picanto shows, even smaller cars are now equipped with advanced computers and touch screens. These screens change the way that tasks are performed compared to using physical buttons only. However, for this research, it does not matter how a task is performed. The more important goal of this research is to find the frequency with which a task is performed. Still, touch screens do sometimes offer more options than regular buttons due to the multiple menus that can be accessed. For example, more advanced screens make it possible to use navigation, which can also be found in the Kia Picanto.

Tasks like making phone calls are also possible using the interface of the Kia Picanto. Other tasks relate to operating the radio and other media devices. There are also multiple buttons for changing the settings of the system. Behind the shifter, multiple buttons are available to operate climate control functions. On the steering wheel, there are also multiple buttons. The functions of these buttons include changing the volume, answering calls and changing system settings. On the right-hand side of the steering wheel, some buttons to operate the cruise control can be found. Then, there are some buttons for changing the fan speed and temperature of the climate control system. Lastly, there are also buttons for the standard functions such as the horn, the windshield wipers, lights, windows, mirrors and the hazard lights.

Figure 3: Tesla Model Y (Auto Express, 2023b)

When looking at the Tesla Model Y, it becomes clear that there are fewer buttons available. The Tesla follows a more clean and minimalistic design. However, this implies that the driver becomes more dependent on the touchscreen. Another implication that comes with a clean and minimalistic design is the fact that a lot of controls are now hidden behind menus. For example, operating the climate control is now hidden behind a menu and the same goes for operating the radio. This can make it more difficult to operate them quickly while driving. Also, the buttons on the touchscreen are smaller than the physical buttons on the Kia and could be harder to find.

The list of tasks that is presented in Table 8 is constructed by studying the popular passenger cars, performing a pilot test trip and identifying tasks that are studied in the literature. This task list forms the basis for the observation checklist that was constructed and used during the observations. The table shows the different identified tasks and shows whether the tasks are seen as strategical, tactical or operational. The tasks are divided into multiple categories. The safety effects of some tasks and categories are more studied than others. Sources that found negative safety effects of some of the tasks are added to the table. These studies were consulted to better understand the risks of performing different tasks and to make sure these tasks are entered into the observation checklist. The table shows which tasks are seen as safety-critical. The definitions used in this research are given at the bottom of the table. Also, the definitions for the operational, tactical and strategical levels as inspired by Michon (1985) are given. Most of the tasks in the table are seen as tactical. They take several seconds to perform and require some decision-making. The basic tasks while driving are seen as operational since they happen automatically and are constantly performed with small adjustments that take milliseconds. The tasks in “Extra features 1” can be considered both tactical and strategical. Most people set these for the long term before even going for a drive but there can also be instances where these settings are changed during a trip as part of the tactical level.

Table 8: Identified tasks

Category (Sources)	**Tasks** Safety relevant indicated with S	**Strategical, tactical or operational**
Radio and Media (Strayer et al., 2017) (Perez et al., 2013) (Horberry et al., 2006) (Ma et al., 2022) (Gillin & Gillin, 2023)	- Adjusting volume (S) - Switching radio station (S) - Connecting with Bluetooth - Switching media input (S) - Switching song (S)	Tactical
Climate control (Gillin & Gillin, 2023) (Ma et al., 2022)	- Turning on/off air conditioning - Changing temperature (S) - Adjusting fan speed (S) - Adjusting fan layout - Activating/Deactivating heated seat - Adjusting recirculating mode (S) - Opening/Closing windows	Tactical
Phone calls (Redelmeier & Tibshirani, 1997) (Strayer et al., 2017) (Horberry et al., 2006)	- Answering phone call (S) - Calling someone (S)	Tactical
Lights	- Turning on/off headlights - Turning on/off high beam light - Turning on/off mist light - Turning on/off interior light - Changing indicator light direction	Tactical
Adjusting car setup	- Adjusting inside mirror - Adjusting outside mirror	Tactical
Windshield tasks	- Changing front windshield wipers speed - Changing back windshield wipers speed - Using front window fluid - Using back window fluid - Changing sun visor position - Activating/Deactivating front window heater - Activating/Deactivating back window heater	Tactical
Basic tasks while driving	- Applying throttle - Changing gear - Steering - Braking	Operational
Danger Signalling	- Using the horn - Turning on/off hazard lights	
Cruise control	- Turning on/off cruise control system	

	- Cancel/resume cruising - Adjusting cruise control speed	Tactical
Adjusting settings (Horberry et al., 2006)**	- Changing automatic distance control setting - Changing sound mix settings (S)	Strategical/tactical
Using extra features (Strayer et al., 2017)* (Blanco et al., 2006)* (Ma et al., 2022)* (Gillin & Gillin, 2023)*	- Enabling/disabling automatic steering - Enabling/disabling automatic lane changing - Enabling/disabling automatic parking - Enabling/disabling automatic distance control - Enabling/disabling automatic lane-keeping - Setting up navigation (S)	Tactical
Safety-linked - Task is linked in the literature to result in distraction and/or reduced driving performance		
Strategical: Tasks that are usually performed before driving, require the most amount of attention and are usually performed to set something up for a longer time period **Tactical:** Tasks that are not performed fully automatically by the driver and take some seconds to perform **Operational:** Tasks that a driver does not think about and happen more or less automatically and constantly during the drive		

* Sources specifically about navigation
** Sources specifically about sound mix settings

The safety aspect of more conventional tasks, like turning on the headlights is not well studied in literature. A reason for this could be that they seem very simple and could be seen as safe actions. Blanco et al. (2006) showed that less complex tasks are less dangerous. However, even simple tasks like turning on the headlights might in the future be incorporated in more complex touchscreen interfaces. It is therefore important to understand how often they are used and that they are included in this research.

The list of tasks as presented in the task list is not exhaustive, since there can always be one more car with one more option. During the trips, it sometimes happened that a driver performed a task that was not on the list. When this was the case, the task was added to the list and was still documented.

3.4 Measures

3.4.1 Counting method and calculations

To count the number of tasks that were performed during the trips, manual observations were done. For both the trip with the familiar and unfamiliar vehicle, the counts were noted per task for different road types and situations. The tasks have also been divided into different categories for which the total counts per road type and situation have been calculated. Based on these counts and the trip durations, the frequencies have been calculated per task and category. To be able to calculate the frequency per road type and situation, the driving times for each of these except the standing still situation have been documented. The total time while standing still was too hard to measure for this study. The total frequency including all tasks has also been calculated. The distinguished categories can be found in appendix D.

The distinguished road types include city roads, rural roads and highways. City roads are defined as roads inside city limits and rural roads as roads outside city limits. This can show the effect of different road types and contexts on the frequencies with which tasks are performed. Highways in the Netherlands are recognized by their special traffic sign. This is a similar approach to Metz et al. (2014) and could help to identify which tasks are performed relatively often on which type of road. Also, it can provide insight for determining on what type of road people perform more HMI tasks in general. The two situations that were included are standing still and driving in a traffic jam. Tasks performed while standing still could be seen as less dangerous. For each of these road types and situations, the frequencies per task are calculated.

Since the execution of the same task can differ between vehicles it was important to define the method that was going to be used for counting. There are multiple options for how to do this. The first option would be to count every single button press as performing the task once. However, some buttons are used by turning the button. This makes it very hard to count. Also, it is not effective for the goal of this research. When a person wants to perform the task “Adjusting volume”, it should only be accounted for once and not for every single button turn or press to change the volume by 1 point. A second option for counting the occurrences of the tasks is to use a cooldown timer for which the same task is not counted multiple times. Metz et al. (2014) did something similar. In this study, if there were button presses within two and a half seconds of the last button press, the task was seen as a continuous task and therefore one period of distraction. Such a cooldown timer has the advantage that single button presses are not measured as single tasks. A downside is that such a cooldown timer will always be relatively arbitrary. Setting the limit to two seconds could provide significantly different results from setting the limit to five seconds. The third option for counting the occurrence of tasks is to count a task every time the driver initiates the task again. The definition for this would be that every time the driver pulls their hand away from the button, the next button press will be counted as a new task. The advantage of this is that no arbitrary cooldown time has to be set but the disadvantage could be that people sometimes interrupt their tasks to continue with it slightly later. This should still be seen as the same task but will now be counted as two separate tasks.

Therefore, it was decided to use the cooldown timer to count the occurrence of the tasks. An educated guess was made on the number of seconds this should be set to, which was decided to be 5 seconds. The time should not be set too low since this could result in overcounting and not too high since this can result in undercounting. In this research, the assumption was made that when a driver performs the same task again within five seconds or continues with a task within five seconds, it is still part of the first task occurrence. The task is seen as being continued and not as a new task occurrence. Turning something on is seen as a separate task from turning something off. For the indicator light, turning the right indicator light on is seen as a separate task from turning the left indicator light on. Also, turning the indicator light off is counted as a separate task, but only when the participant has to move his hand towards the indicator light again. If the indicator light is turned right and then in the same movement turned off again it is only counted as one indicator light task. Also, when a participant switches the indicator light from left to right or right to left in one go, only one task occurrence has been counted since the overarching goal was to change the indicator direction once.

The counting began when the driver moved the vehicle for the first time. All tasks that were performed before this moment were not documented, since they were not seen as tasks during the trip. Only tasks that are part of the original internal HMI have been looked at during the observations. This research does not focus on tasks that could be distracting outside of the internal HMI, since the overarching goal of this study is to be able to assess the internal HMIs of passenger cars. Tasks that are performed with a smartphone are therefore not documented. When a driver touched a button or control but did not change anything, this was also not counted. The way of performing the task was also irrelevant for this research. Performing a task with speech is counted the same as performing that task with a button.

3.4.2 Measured factors (socio-demographic, contextual, car familiarity, car characteristics)

Table 9 shows the different factors that were documented and added to the observation checklist. The factors are divided into three categories. These categories are socio-demographic factors, contextual factors and car familiarity.

Table 9: Identified factors

Category (Sources)	Factors	Unit	Hypothesis
Socio-demographics *(Burdett et al., 2019)* *(Arca et al., 2022)* *(Klauer et al., 2014)* *(Cooper et al., 2020)*	Age	Years	Negative correlation
	Gender	Male/Female	Nominal variable
	Time with a driving license	Years	Positive correlation
	Frequency of driving	Average times per week last year	Positive correlation
Contextual factors *(Cuentas-Hernandez et al., 2023)*	Outside temperature	Degrees Celsius	Positive/negative correlation
	Weather type	Sunny/Rainy/Cloudy/Snowing/Dark	Nominal variable
	Peak hour	Yes/No	Nominal variable
	Day of the week	Mon/Tue/Wed/Thu/Fri/Sat/Sun	Nominal variable
Car familiarity *(Harms et al., 2021)* *(Harms et al., 2019)* *(Chisholm et al., 2008)* *(Broström et al., 2013)*	Car familiarity	1-10	Positive correlation
Socio-demographic: Characteristics of the participants **Contextual:** Factors that are related to the context of a particular trip **Car familiarity:** How familiar a driver is with the car that they are driving			

* A car is seen as familiar when the car familiarity score is higher than or equal to 8 on a scale from 1-10 (Harms et al., 2019).

For the socio-demographic variables, the decision was made to include age, gender and driving experience. These variables are often included in literature when performing data analysis. Burdett et al. (2019) also included the participants' age and driving experience as variables in a study where they drove along with participants. The expectation was that age has a significant effect on the frequency with which people perform tasks. More specifically, the relation with age was expected to show a negative correlation with the task frequencies. Older people are often less tech-savvy and therefore struggle more to perform secondary tasks (Cooper et al., 2020). The expectation was therefore that older people show lower frequencies of performing tasks. Metz et al. (2014) also included age and driving experience in their study which investigated secondary task frequencies. Gender was not expected to have a significant correlation with the frequency. Arca et al. (2022) did find that women tend to display worse driving performance when distracted in comparison to men, which could result in men having higher frequencies, but this is questionable. It was expected that the number of years with a driving license would have a positive correlation with the frequencies since Klauer et al. (2014) found that novice drivers engaged more frequently in distracting tasks as they gained more driving experience. Also, Harms et al. (2021) found that lower task difficulty has a relation with increased engagement in secondary tasks. Since more experienced drivers have less difficulty with the driving task, it was expected that more experienced drivers have higher frequencies with which they perform tasks. Driving experience was measured in this study by using two variables, which are years with a driving license and frequency of driving. Only using one variable would not be sufficient in measuring driving experience because there might be participants who have had their driving license for many years but never drive.

Four contextual factors have been identified as independent or controlling variables for this study. These are outside temperature, weather type, peak hour and day of the week. The weather type and outside temperature are included since they could have a big influence on the use of climate control options and windshield wipers. The correlation between the temperature and frequencies was expected to be negative as the temperature increased to 20 degrees and positive as the temperature increased from 20 degrees onwards. The reason for this is that people are expected to only use climate control options in environments that are either too hot or too cold. Different weather types were furthermore expected to have different correlations with the frequencies. The hypothesis was that rainy weather increases the amount of windshield tasks that are being used. Lastly, it was noted whether the trip was driven during peak hours or not and on what day of the week. The inspiration for this variable was found in a study by Cuentas-Hernandez et al. (2023). This study looked at the influence of context factors on distracted driving. They found that traffic density had an impact on secondary task engagement. In this study, the traffic density was measured by a pre-recorded variable from a database that contained levels of service. Since this type of data was not available for this research, the choice was made to note whether or not the drive was during peak hours and on what day of the week, since this can give good indications of traffic density (de Haas, 2020).

To measure car familiarity, the participants had to state whether they agreed or disagreed with different statements on a scale of 1-10. One statement attempted to measure overall car familiarity while 4 other statements tried to measure car familiarity from different angles. This variable was included since the literature suggests that familiarity can have a big impact on driving behaviour. There are studies indicating that repetition of a task influences driving performance (Chisholm et al., 2008; Broström et al., 2013). Furthermore, the study by Harms et al. (2021) showed that route familiarity influenced driving behaviour and secondary task behaviour. Also, similar scales for measuring familiarity have been used in the past (Harms et al., 2019). The participants were also asked if they could describe what the researcher was looking at. This was done to find out whether the participants already had an idea of what the study was about. Then, they were asked if they felt that the researcher had influenced their behaviour during the trip. Most people could not correctly describe what the researcher was observing while they drove their routes and did not feel like their behaviour was influenced.

Since the routes and cars in this experiment were not controlled, also some characteristics were documented. For the route characteristics, the driving time per road type and situation was noted, as well as the total trip duration. These characteristics can be seen in Table 10. The car characteristics were used to look at the effect of the type of car on the task frequencies and can be found in Table 7 in section 3.3.1.

Table 10: Characteristics to be noted

Category (Sources)	Variables	Unit
Route characteristics (Metz et al., 2014)	Time per road type	Minutes
	Time in traffic jam	Minutes
	Trip duration	Minutes

3.5 Participants

Most participants who took part in the study were friends and family of the researcher. In the end, 30 participants were recruited. To be eligible for the study, participants were required to be above the age of 18 and had to own a driver's license. Also, as far as this was possible, participants were selected in a way that a variety of different ages and genders participated. All 30 participants successfully completed the experiment. In return for their participation, they received a bol.com gift card worth 10 euros.

The average age in the study was found to be 41.1 years. A good range of people of different ages participated in the study as the youngest person was aged 20 and the oldest person was 85. Since people were asked to drive a familiar route, the minimum route familiarity was scored as 8 with an average route familiarity of 9.47. Two-thirds of the participants were male drivers. Most participants did not feel influenced by the researcher nor did they know what the researcher was looking at during the trips. The information about the participants can be found in Table 11.

Table 11: Socio-demographic variables

Numerical sociodemographic variables					
Variable	**Average**	**Min**	**Max**	**Std dev**	**N**
Age (years)	41.10	20	85	16.59	30
Time with a driving license (years)	22.03	3	65	16.03	30
Driving frequency (times per week)	7.90	1	28	5.44	30
Route familiarity	9.47	8	10	0.78	30
Categorical sociodemographic variables					
Variable	**Category**	**Amount**	**Percentage**		**N**
Gender	Male	20	66.67		30
	Female	10	33.33		
Q1: Did researcher influence participant?	Yes	4	13.33		30
	No	26	86.67		
Q2: Did the participant know what the researcher was observing while driving their route	Yes	1	3.33		30
	No	29	96.67		

3.6 Pilot

The study setup and the relevant tasks that are discussed in the previous sections have been fine-tuned by performing a pilot involving two test trips with the same participant. The pilot showed that manual observations are possible in a fairly accurate manner using the methodology described above. However, this does require a bit of practice. During the pilots, repetition of performing them increased the confidence of the researchers in accurately describing the tasks. The pilots revealed that the hardest part of doing the observations was to make a distinction between tasks that are performed while driving or whilst standing still. They also revealed that sitting in the back of the car was not possible due to the limited view of the user interface and observations in the dark were more difficult due to the lack of visibility.

The pilot revealed some tasks that were still missing from the first list which was constructed by looking at popular passenger cars and literature research. New tasks were added each time that a pilot revealed that it was still missing. Also, the pilots were input for changing some of the task definitions. This is a similar approach to Metz et al. (2014), who also performed a study looking at different tasks. To come up with the task list for this study, the researchers looked at video footage to make a distinction between different tasks.

During the pilots, the participants were asked two questions. They were asked if they felt the researcher had influenced their general behaviour or driving behaviour in any way by sitting next to them during the drive and if they already had an idea of what the researcher would be looking at after hearing the introduction. The consensus among the participants from the pilots was that they did not feel influenced by the researcher in their behaviour and also did not know what the researcher was looking for during the drive after hearing the introduction.

Another discovery that was made during the pilots was that people did not always know how to answer the car familiarity question. When asking someone how familiar they are with a certain car, different people think of different things. Some people might think about how often they drive it while others think about how many of the options of the car they know. Therefore, overall car familiarity can be seen as a latent variable. It might not be possible to accurately measure overall car familiarity with one question. To still be able to measure overall car familiarity and make the scale more valid, multiple statements are presented to the participants which should cover every angle of car familiarity. The four statements that are used to cover these different angles can be found in the observation checklist in Appendix C. People have also still been asked about their overall car familiarity. The scores on this variable can later be compared to the scores on the factor to see if they measure the same. This research could then be the first step to creating an overall car familiarity scale.

The information from the previous sections resulted in the initial observation checklist. This can be found in Appendix C. Over time, some tasks were added to the observation checklist and the final results can be found in Appendix D.

3.7 Data analysis methods

The data gathered from the observations was entered in SPSS and the statistical tests were performed using this software package.

3.7.1 Repeated measure ANOVA's and paired t-tests

To test if the differences in the average frequencies of the different tasks that were performed are statistically significant, repeated measure ANOVA's were performed. This method was chosen because the frequencies of all tasks are measured with the same individuals. Performing a one-way ANOVA does therefore not work, as this method requires different groups for each task. Alternatively, paired t-tests could have been performed. However, this would have to be done for each pair of tasks. Without adjustment for the fact that multiple paired tests are performed, the chance of at least one type 1 error would be too high. An assumption of repeated measures ANOVA's is sphericity. This assumption states that the variance of the differences between all tasks is the same. Therefore, Mauchly's test of sphericity was used. When the assumption is violated corrections have to be made. One of these corrections is made by the Greenhouse-Geisser coefficient.

Paired t-tests are performed to test for differences in average frequencies between the two trips that are driven by the participants. Repeated measure ANOVA's were not applicable here, since not all pairs are tested for differences. Only the two averages for the same tasks are compared.

3.7.2 Factor analysis

To find if overall car familiarity is an underlying factor of different indicators, a factor analysis was performed. This factor analysis shows whether the variables partly measure the same or not. If the statements had a high load on the same factor, they were used to create an overall familiarity scale. A factor analysis is very useful to find and measure underlying factors, which are also called latent variables (Yong & Pearce, 2013). A standard approach for factor analysis has been used. Statements that load lower than 0.3 on the factor have been deleted for the overall familiarity scale. The overall scores on the familiarity scale have been calculated with a sum of scores. This sum of scores has been divided by the amount of statements for the factor. This means that the new overall car familiarity scale can be interpreted in the same way as the four statements.

To test whether the resulting factor and the overall car familiarity statement measure the same thing, a linear regression analysis was performed. This method can be used to measure the impact of independent variables of interval, ratio, nominal and ordinal scale on a dependent variable that has an interval or ratio scale (Aiken et al., 2012). The dependent variable in this case was the score on the factor that was created in the factor analysis. The independent variable was the score on the overall car familiarity statement. Factor analysis and linear regression have been combined in literature before. A study by Hamari et al. (2015) also used a factor analysis with a multiple linear regression model. In this study, they tried to find the influence of different factors on behavioural intention.

3.7.3 Chi-square test

To test if the distribution of tasks that are performed while driving versus standing still is the same for different categories a chi-square test was performed. Tasks that are more typically performed while standing still could be seen as less dangerous than tasks that are typically performed while driving. To meet the requirements of a chi-square test, only tasks or task categories that have enough observations are tested for differences in distribution. The requirements of a chi-square test state that a maximum of 20% of the expected cell counts are allowed to be smaller than 5. Also, all expected cell counts have to be bigger than 1.

3.7.4 Poisson regression

The dependent variable in this study is the number of times that a task has been performed adjusted by the driving time for each participant. This type of data can be seen as count data. Performing a linear regression on count data can result in biased results (Coxe et al., 2009). This is because count data can only be positive whilst linear regression can also result in negative numbers. Also, count data is usually poison-distributed and not linear. Therefore, Poisson regressions were performed. An assumption of the Poisson model is equidispersion. This infers that the mean and variance of the model are identical. This can be checked by looking at the Deviance and Pearson chi-square statistics. When the values of these statistics divided by their degrees of freedom are close to one, the assumption is met. A value greater than one implies overdispersion, which means that the variance is greater than the mean and a value less than one implies underdispersion, which means that the variance is smaller than the mean. Also, the likelihood ratio chi-square test has been used to compare the fitted model to the intercept-only model. When the coefficient for this test is significant, it indicates that including the factors results in a statistically significantly better fit. The natural log of the trip durations has been used as the offset variable to take into account the fact that the trip durations varied across participants. The formula for the Poisson regressions in which x_i stands for independent variable i can be found below:

- ***ln(number of times) = $\beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots. \beta_i x_i$ + ln(trip duration)***

In Figure 4, the conceptual model for the Poisson regressions is presented. Multiple models have been constructed with two different types of dependent variables. These are the total number of tasks that have been performed per hour and the total number of tasks per category that have been performed per hour. The Poisson models have been constructed using the data from the trips where the participants drove their own car. The reason for this is that the unfamiliar car data has almost no variety in car types, as 28 out of the 30 participants drove the same unfamiliar car. The Poisson regression analyses reveal which factors influence the overall frequency with which tasks are performed and which factors influence different types of factors. The operationalization of all the variables can be found in Appendix B.

A specific method has been followed to get to the final Poisson models. First, separate models have been made for all factors. Only one factor at a time was entered per dependent variable. The most interpretable factors with the highest significance values were then combined in one model. The experimenting stops when a final model has been constructed that is interpretable whilst also having (near) significant variables. The correlation table has been used to better understand which factors should be entered for the final model. Since the sample of this study is quite small, adding highly positively correlated variables might reduce the significance by a lot.

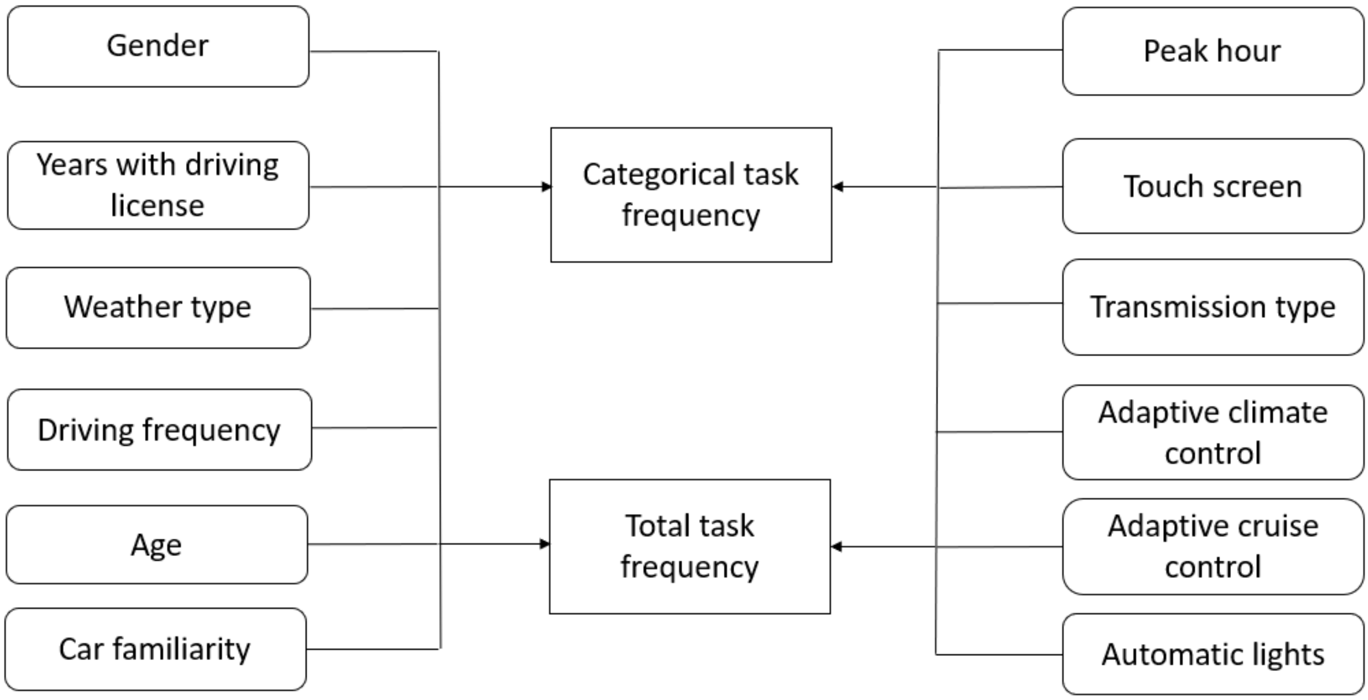


Figure 4: Full conceptual model for Poisson regression

4. Results

In this chapter, the results of these analyses are presented. To understand the context of the results, It is important to note that the most common weather type during the trips was found to be cloudy weather followed by rainy weather and that most trips were performed outside of peak traffic hours. More information on this can be found in Appendix E. In the next section, the most frequently used tasks are presented and differences between the average frequencies have been tested for statistical significance. The detailed results for this can also be found in Appendix E. Also, the aggregated data is presented to get a more compact overview of the frequencies with which tasks are performed. Afterwards, the results of the factor analysis that has been performed for the car familiarity statements are presented. These results are used to study the impact of car familiarity on the frequencies. The influence of all factors on the frequencies is looked at by performing Poisson regressions. During the analyses, an alpha of 0.05 was targeted for statistical significance.

4.1 Most frequently performed tasks

The most used task during the 60 trips was found to be the indicator light, followed by adjusting the front windshield wipers' speed and adjusting the volume. On average, the frequency of using the indicator light is found to be 66.4 per hour. The front windshield wipers' speed was adjusted with a frequency of 4.22 times per hour whilst the volume was adjusted with a frequency of 2.15 times per hour. The windshield wiper task is an interesting one, as it displays a relatively high maximum compared to some of the other tasks. The reason for this is that some participants were unable to figure out how to set the windshield wipers to an interval. Therefore, some participants had to constantly manually trigger the windshield wipers. This explains the high maximum of 29 occurrences on a single trip for this task. The figure below shows the average frequency per trip and the maximum number of occurrences in a single trip of the 10 most frequently performed tasks during the 60 trips, which includes both the trips with the familiar and unfamiliar car. The figure is ordered from the least used tasks to the most used tasks. In Appendix D, a more complete overview of the data can be found for all tasks. However, this appendix shows the data separately for the familiar car trips and unfamiliar car trips. More information about these 10 tasks can be found in Appendix D.

To test if the differences in average frequency for the different tasks are statistically significant, a repeated measures ANOVA was performed. This analysis can be found in Appendix E. The repeated measures ANOVA indicates that at least one average frequency is statistically significantly different from one of the others. In particular, the results show that the average frequency of the indicator light is statistically different from the average frequency of all other tasks. The average frequency of the volume task is statistically different from the average of switching radio stations and adjusting the inside mirror. All the other tasks only have statistically significantly different average frequencies from using the indicator light and/or adjusting the volume.

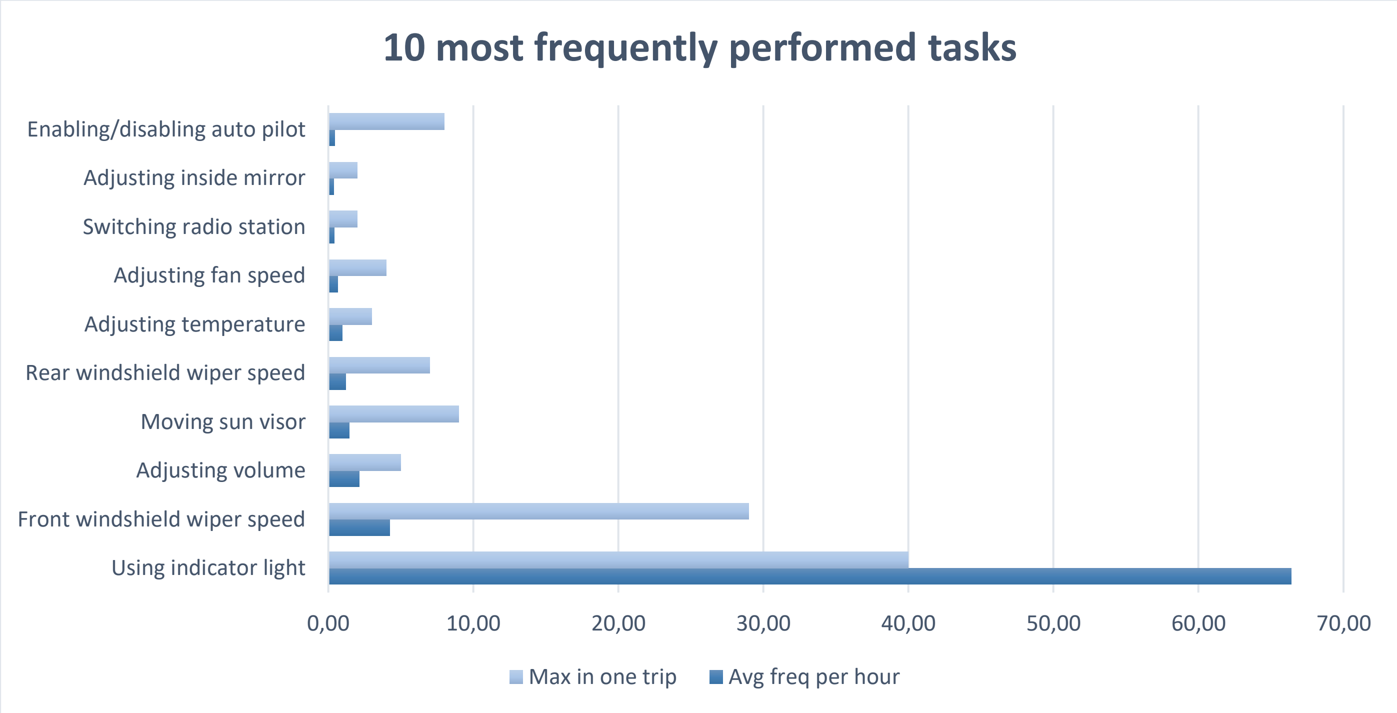


Figure 5: Average frequency and maximum per trip for the 10 most often performed tasks

On average, participants performed a total of 29.93 tasks per trip and performed 81.35 tasks per hour of driving. In total, participants performed 1796 tasks during all trips combined. The average amount of tasks performed per trip is reduced a lot when the indicator light is excluded. In this situation, the average amount of tasks performed per trip goes from 29.93 to 5.38. Since indicator light is part of the lights category, it is not surprising that this category has the highest average amount of tasks performed per trip with 24.62. The categories with the second and third most tasks completed during the trips are the windshield and radio and media categories with 2.68 and 1.17 tasks completed per trip on average. The categories with the lowest averages per trip are calling, changing settings and danger signalling. Figure 6 shows the average frequency per trip and the maximum number of occurrences in a single trip for the different categories. Only the categories that have more than 10 total tasks performed are presented. In Appendix D, it is made clear which tasks are included in which category. An overview of the data for all categories can be found in Table 36 in appendix E.

The repeated measures ANOVA in Appendix E shows that the average frequency of the lights category is statistically different from all other categories. The average frequency of the windshield task category is only statistically significantly different from the lights category, the setup category and the “other” category. The average frequency of the radio and media category is also different from these categories. The average frequencies of the other categories in Figure 6 are only statistically significantly different from one or more of these three categories.

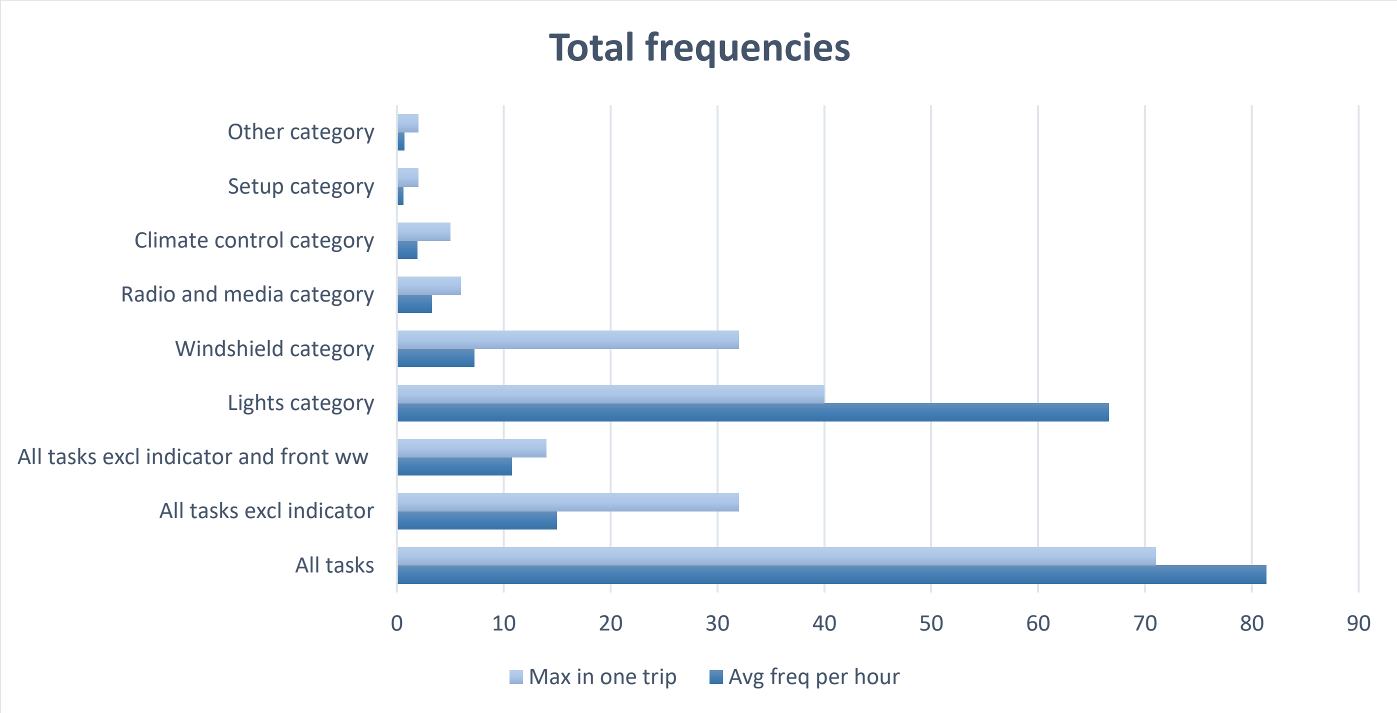


Figure 6: Aggregated average frequencies and maxima per trip

Table 12 shows the descriptive data for the different road types and situations. The indicator light task has been excluded since it is such an outlier. The data including the indicator light task can be found in Appendix E. The windshield wiper speed task was kept in since this task is less of an outlier than the indicator light. All of the 60 trips had at least some driving time on roads within city limits, whilst only 8 of the trips had driving time on the highway or encountered a traffic jam. This means that the averages for the city road type have 60 data points compared to only 8 data points for the highway road type. Therefore, the averages for the city road type are more accurate. On average, participants drove 18.22 minutes per trip on roads in the city, which makes it the road type with the highest average. It is therefore no surprise that this road type also has the highest average amount of tasks performed per trip. However, when we look at the average frequency which takes into account this difference in driving time, the numbers are a bit different. With the indicator light excluded, the highest average frequency can be found during traffic jams and the lowest average frequency with the rural road type. On average, a trip during the study took 22.35 minutes.

Table 12: Tasks performed per road type using the data from both trips combined

Number of tasks performed per road type and situation excluding the indicator light						
Variable	**Average number of tasks**	**Total number of tasks**	**% of total**	**Avg freq per hour**	**Average driving time in minutes**	**# of trips**
City	3.78	227	70.28	12.58	18.22	60
Rural	0.3	18	5.57	6.54	2.98	34
Highway	0.23	14	4.33	17.5	0.85	8
Standing still	0.93	56	17.34	NA	NA	NA
Traffic jam	0.13	8	2.48	26.29	0.30	8

4.2 Factor analysis of car familiarity

To test if the four car familiarity statements in the study measure the same underlying factor, a factor analysis has been performed. The hypothesis is that the four statements together measure the underlying factor of overall car familiarity. Participants were also asked about their overall car familiarity directly. This makes it possible to compare the results of the factor analysis to the overall car familiarity statement. It could be a first step to better understand what is measured when asking people to rate their overall familiarity with a car.

The indicators in this study are measured on a 10-point scale. A factor analysis normally requires indicators measured on an interval- or ratio scale. Strictly speaking, the scale used for this study should be seen as ordinal, however, Likert scales are often used as interval scales in this type of research (Wu & Leung, 2017). For this research, the only viable solution was to use a 10-point scale. The factor analysis was performed with 30 data points. Each data point represents one participant. For each participant, the average score on the statements for the two trips was used. During the factor analyses, the statement "I drive this car a lot" was deleted based on its low communality and low factor loading. In the end, one factor was extracted and this result is shown in Table 13.

Table 13: Factor analysis

Perceived HMI car familiarity ($\alpha = 0.846$)	**Factor loading**
I know the features/options that the car has available	0.945
I know where the buttons are located	0.891
I understand the dashboard and the things it displays	0.649

From the factor analysis, it becomes clear that the three statements in Table 13 measure the same underlying factor. This factor has been called the perceived HMI car familiarity, as all the statements are about how familiar the HMI of the car is perceived by the participants. This new perceived HMI car familiarity variable has been created with a sum of scores. To see if this was possible, the reliability of the scale was checked. The scale was found to be reliable (>0.70) with a Cronbach's alpha of 0.846. The sum of scores has the advantage that it can be interpreted in terms of the original scale. The average and standard deviation of the new variable can be found in Table 14 and are compared to the average and standard deviation of the overall familiarity statement.

Table 14: Descriptive statistics of the factor and overall car familiarity question

	Average	**Standard deviation**	**P-value**
Perceived HMI car familiarity (α = 0,846)	7.79	1.18	0.003
I am familiar with this car	7.12	1.52	

The difference between the averages for the two variables is quite small. This could indicate that people do take in mind the different aspects of car familiarity when rating overall car familiarity. However, when testing the averages, they come out as statistically significantly different. To test how well the perceived HMI car familiarity can be predicted by the overall car familiarity statement, a linear regression model has been constructed. The results of this analysis are shown in Appendix E. The results indicate that 46% of the variance in perceived HMI car familiarity is explained by the overall car familiarity statement. Furthermore, it can be seen that the coefficient for the overall car familiarity statement is statistically significant and positive. This indicates that the higher the score on the overall car familiarity statement the higher the score on the perceived HMI car familiarity variable.

4.3 Influence of car familiarity on task frequencies

This section studies the impact of car familiarity on the average frequencies with which different tasks are performed. In Table 15, the average scores on the car familiarity statements are presented for the familiar and unfamiliar car and the differences are tested for statistical significance. Then, the average frequencies with which different tasks are performed are compared for the two trips. Lastly, paired t-tests are performed to test the differences for statistical significance.

Looking at Table 15, it can be seen that the averages for the familiar car are higher across the board and the standard deviations are lower. Interestingly, participants stated that they were relatively familiar with the dashboard of the unfamiliar car, given their average score of 8.23 on that statement. Also, the participants rated the perceived HMI car familiarity of the unfamiliar car at an average of 6.57 and some participants even rated the familiarity of the unfamiliar car with a 9 or a 10. When looking at the paired t-tests, it can be concluded that all the differences between the means are statistically significantly different. This means that indeed the participants' cars were rated more familiar than the car that was brought by the researcher.

Table 15: Car familiarity statements paired t-tests

Variable	Own car	unfamiliar car	T-value	P-value
I drive this car a lot	9.17	1.37	22.06	<0.001
I know the features/options that the car has available	8.6	5.23	9.34	<0.001
I know where the buttons are located	8.97	6.23	9.14	<0.001
I understand the dashboard and the things it displays	9.47	8.23	5.40	<0.001
I am familiar with this car	9.5	4.73	8.80	<0.001
Perceived HMI car familiarity	9.01	6.57	10.70	<0.001

In Figure 7, the 10 most performed tasks for both trips are presented and their average frequencies are compared. In Appendix E more detailed information is presented for these tasks. The most performed task in both the trip with the familiar car, as well as the trip with the unfamiliar car, is changing the indicator light direction. This task was performed with an average frequency of 66.22 and 66.58 respectively. For the familiar car, the next most used tasks were found to be changing the volume, moving the sun visor, changing the temperature of the climate control system and changing the front windshield wiper speed. For the unfamiliar car, the most performed tasks after changing the indicator light direction are found to be changing the front windshield wiper speed, adjusting the volume, changing the speed of the rear windshield wiper and moving the sun visor. For the unfamiliar car, higher frequencies can be found for tasks that are related to setting up the car. Examples of these tasks are adjusting the inside mirror and setting up the seat.

For both trips, repeated measure ANOVA's have been performed to test if the differences in average frequencies between tasks are statistically significantly different. For the familiar car, the results show that the average frequency of the indicator light task is different from all other tasks. The average frequency of the volume task is found to be statistically different from the average frequencies of the indicator light, switching radio and opening/closing windows task. The other tasks for the familiar car in Figure 7 only showed statistically significant differences with either the indicator light task and/or the adjusting volume task. For the unfamiliar car, the results show that only the indicator light task has a statistically significant difference in average frequency from the other tasks.

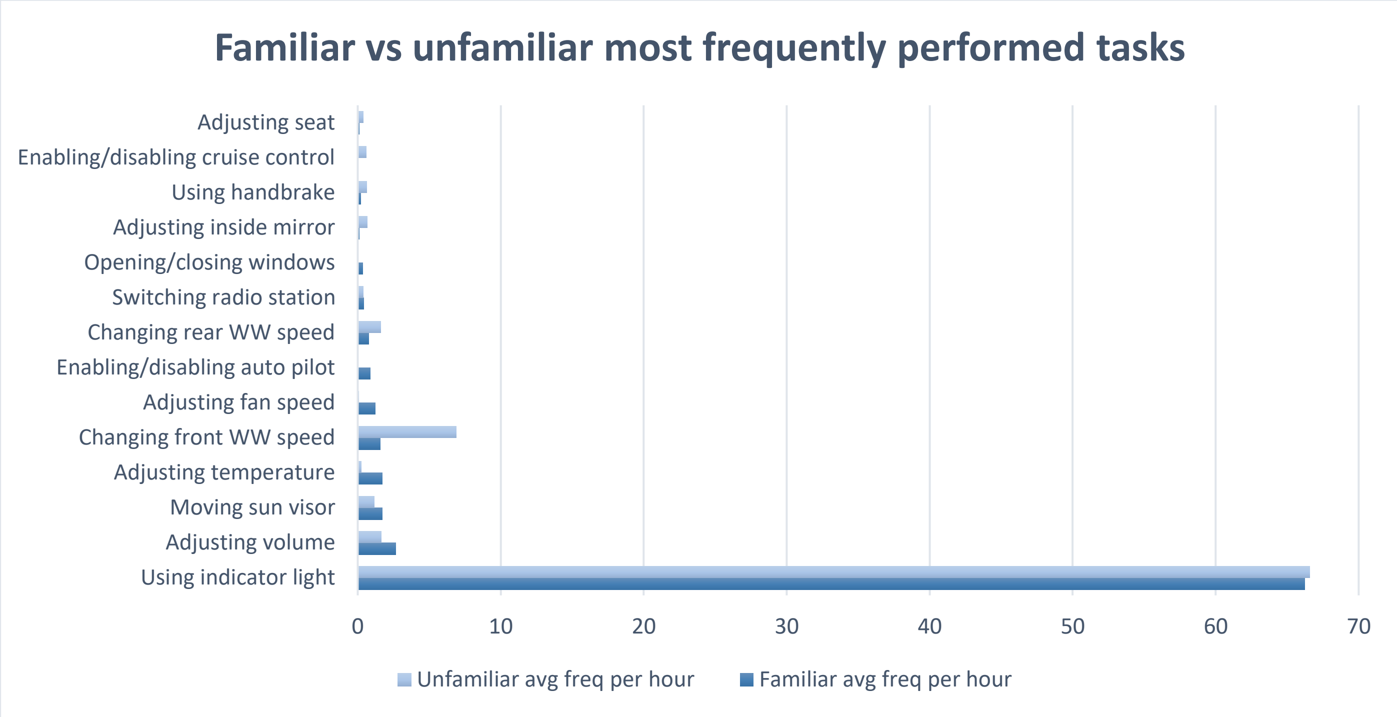


Figure 7: Average frequency per hour for the familiar and unfamiliar car trip

In Figure 8, the average frequencies per trip for the different categories are presented for both the familiar and unfamiliar car. For the familiar car, the average total number of tasks that were performed during a trip was found to be 28.97 with an average frequency of 79.89 per hour. Excluding the indicator light from the task set reduces these numbers to 4.8 and 13.67 respectively. Excluding also the front windshield wiper task reduces these numbers further to 4.27 and 12.11 respectively. When looking at the unfamiliar car, the average amount of tasks performed is 30.9 with an average frequency of 82.82 per hour. Excluding the indicator light reduces these averages to 5.97 and 16.27 and excluding also the front windshield wiper task reduces these averages down further to 3.43 and 9.36. For both the trips with the familiar car and the unfamiliar car, the three categories with the highest average task frequencies are the lights, windshield and radio and media categories. For the familiar car, the next category with the highest task frequencies is found to be the climate control category whilst for the unfamiliar car this is found to be the setup category.

Again, repeated measure ANOVA’s have been performed for both trips to test if the differences in average frequencies between the task categories are statistically significantly different. For the familiar car, the results indicate that only the lights category has a statistically significant difference in average task frequency from the other tasks. The same is true for the unfamiliar car.

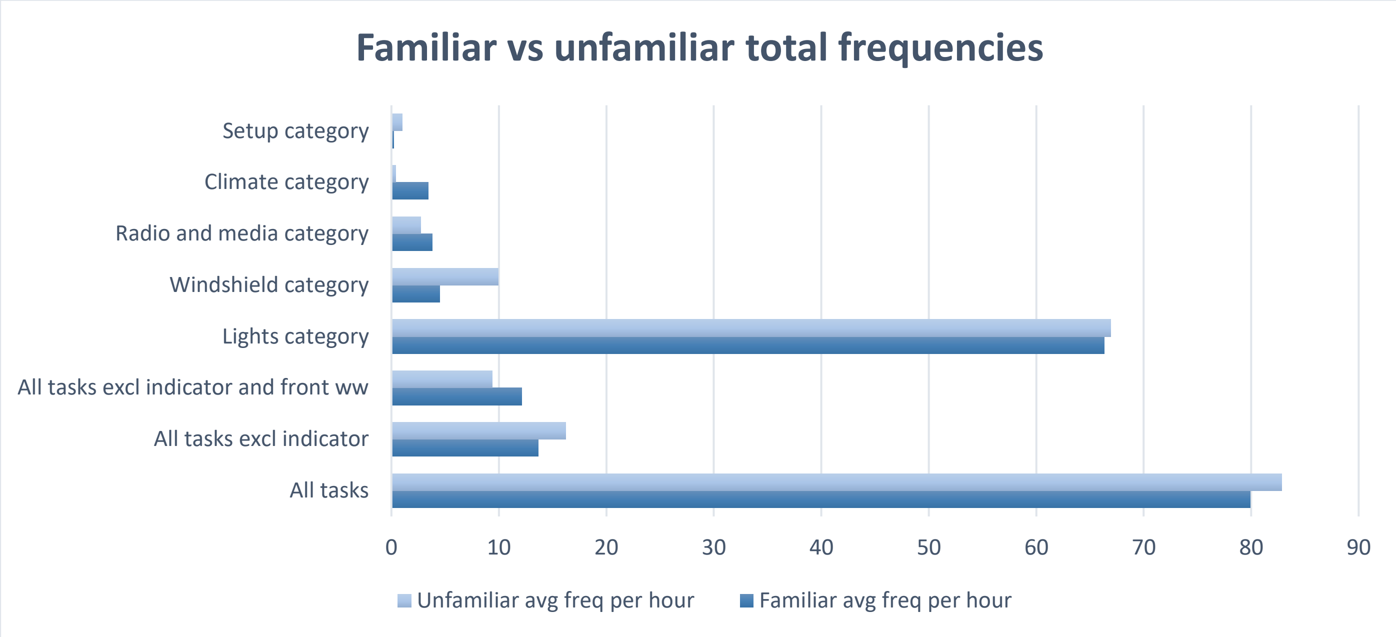


Figure 8: Aggregated average frequencies per trip compared for the two trips

When using an alpha of 0.05, two categories display average frequencies that are statistically significantly different between the two trips. These categories are the climate control category and the setup category. The climate control category has an average frequency of 3.45 tasks per hour when people were driving their cars versus 0.4 tasks per hour when driving an unfamiliar car. The setup category has an average frequency of 0.21 tasks per hour when people were driving their cars versus 1.03 tasks per hour when driving an unfamiliar car. A variable for which the difference in average frequencies is nearly statistically significant is the variable that looks at the total tasks performed excluding the indicator lights and front windshield wiper tasks. This test resulted in a p-value of 0.07. Since the front windshield wiper was such an outlier due to the unintuitive design of the Seat Toledo, this would indicate that people in general perform more tasks in familiar cars.

Table 16: Paired sample t-tests for differences in average aggregated frequencies of the two trips

Variable	Own car	unfamiliar car	t-value	P-value
Total excl indicator light and wwspf	12.11	9.36	1.86	0.07
Climate category	3.45	0.40	4.32	<0.001
Setup category	0.21	1.03	-2.30	0.03

In the previous analyses, it became clear that the front windshield wiper task was used significantly more in the unfamiliar car compared to the cars of the participants. Since the windshield category also includes the sun visor task, this difference was not tested properly in the paired t-tests. Therefore, another paired t-test has been performed for the front windshield wiper task separately. The observations for the two other cars that were used as unfamiliar were deleted, such that only the data from the Seat Toledo has been used. Again, the difference in the average frequency was tested for statistical significance. Interestingly, even though the difference in the average frequency is quite big, the result of the t-test indicates that the difference is not statistically significantly different. This seems to be due to the big standard deviation in the average frequency for the Seat Toledo. However, the data still clearly shows that some participants struggled to set the windshield wipers to interval mode. This can be seen by looking at the high maximum for the windshield wiper speed task in the Seat as well as the high standard deviation. When comparing this to the data from the trips where the participants were driving their own car, it becomes clear that the differences are rather large. The results of this t-test can be found in Table 17.

Table 17: Paired sample t-test for difference in average frequency in the Seat versus the cars from the participants for the windshield wiper task

Variable	Own car	Seat Toledo	St dev own car	ST dev Seat Toledo	t-value	P-value
Front windshield wiper	1.67	7.12	3.03	19.21	-1.469	0.153

4.4 Influence of the factors (socio-demographic, contextual, car familiarity, car characteristics) on the task frequencies

In this section, the impact of all factors on the average task frequencies is studied. First, the percentage of tasks that have been performed on the different road types and situations is presented. This is shown per category to see if there are any differences between the different categories. Since most of the trips had the longest driving time on city roads, the percentages in this column are the highest for most categories. Still, there are some differences between the categories. 30% of the time, the climate control tasks were performed while standing still versus only 9% for windshield tasks and 21% for radio and media tasks

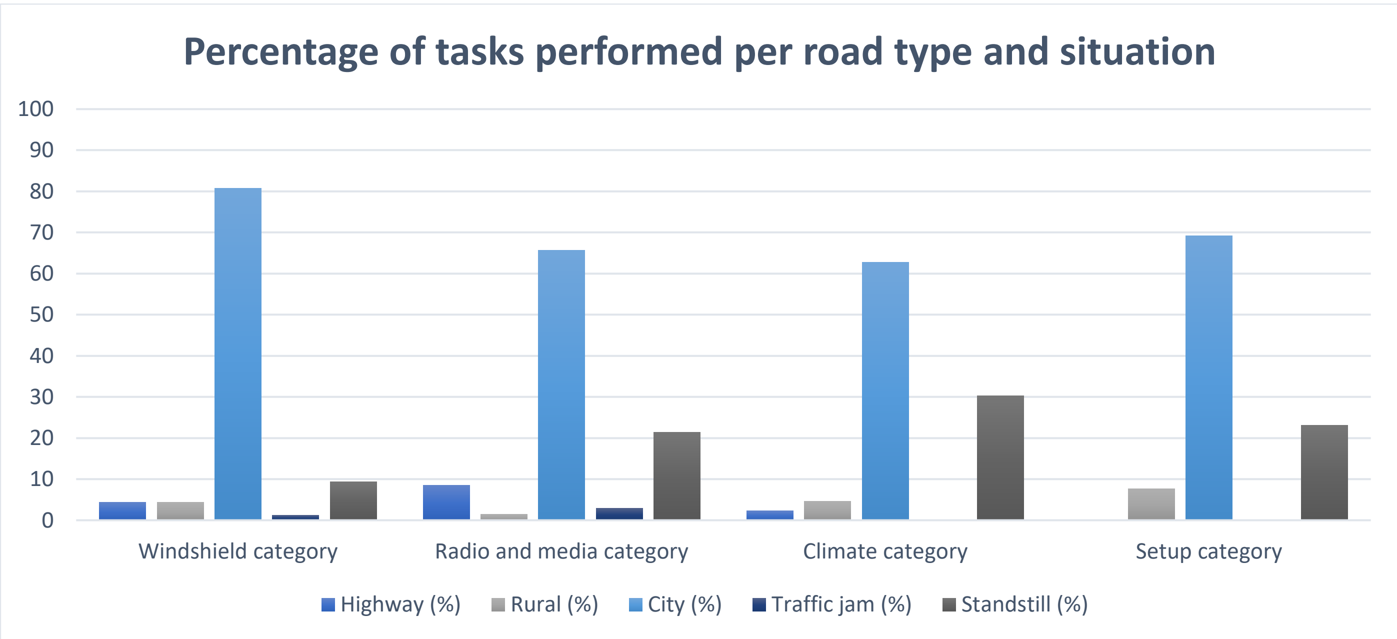


Figure 9: Percentage of tasks performed per road type and situation

The context in which a task is performed could have an impact on the safety with which that task is performed. A task could be seen as less dangerous when the car is standing still. Therefore, a chi-square test has been performed to test if the distribution of tasks that are performed while driving versus standing still is the same for different categories. For the chi-squared test, only categories have been used that have at least one observation for standing still. This was done to make sure that the requirements for the chi-squared test were met, which is explained in Chapter 3. The “Other” category was left out of the test since this includes the handbrake. In theory, the handbrake should only ever be used during a standstill. However, for this test, the goal is to find out if the distribution is different for tasks that can be performed during driving as well as during a standstill. Therefore, only the radio and media, climate control, windshield and setup categories have been taken into account. This resulted in a statistically significant p-value. The distribution of tasks performed while standing still versus while driving is thus not the same for these four categories. The results of this analysis are shown in Table 18.

Table 18: Percentage of tasks performed while driving versus standing still

Chi-square test for distribution p-value = 0.00 Chi-square = 13.83 DF = 3		
Category	**Driving (HW+RU+CI+JA) (%)**	**Standing still (%)**
Radio and media category	78,57	21,43
Climate category	69,77	30,23
Windshield category	90,68	9,32
Setup category	76,92	23,08

To study the impact of the other variables in the study, Poisson regression models have been constructed using the familiar car data. Figure 4 in Chapter 3 shows which variables have been entered initially. To construct the Poisson models, the method explained in chapter 3.7.4 has been used. In Appendix E, the individual Poisson models for the different factors can be found and only models with a statistically significant coefficient are presented. Also, some near statistically significant models have been presented based on their interpretability. Using these models and the correlations between the factors that can be found in Appendix E, variables have been added together in the Poisson regression to end up with a final model that is interpretable. The statistically significant correlations can be found in Table 19.

The task categories with the highest frequencies are used in the models as the dependent variables in the Poisson regression. The Poisson regression was run to predict these frequencies based on the factors in the study. The task categories with the highest frequencies are the radio and media category, the climate control category, the windshield category and the lights category. Since the lights category mainly consists of the indicator light, it has not been used as a dependent variable. No relations are expected between the factors and the frequencies of the indicator light task. Also, one Poisson regression was run to predict the number of total tasks that are performed per hour excluding the indicator light. The results can be found in Table 20. Due to the small sample size, these models should be seen as explorative models.

So far, the familiar car has been defined as the car that the participants brought while the unfamiliar car has been defined as the car that the researcher brought. The idea is that they are more familiar with their own car than the car that was provided by the researcher. Table 15 shows that this is indeed the case. However, it is also interesting to use the scores on the perceived HMI car familiarity variable to see how these scores influence the frequencies. This statement measures how familiar the participants perceived the HMI of the car on a scale from 1 to 10. Therefore, this factor is entered as an independent variable for the Poisson regressions.

Table 19 shows that correlations between the independent factors exist. Age has a statistically significant correlation with the number of years with a driving license. This correlation is the strongest and the coefficient is nearly one. These variables therefore measure almost the same thing. The driving frequency of the participants is statistically significantly correlated with the perceived HMI car familiarity. This makes sense, as people who drive a lot are expected to be more familiar with their car. Also, the car characteristics have correlations. It seems that when a car is equipped with one of the more modern features, it is usually also equipped with the others. The last correlation can be found between the peak hour variable and the dark variable. This makes sense as the peak hours are from 4 pm to 7 pm. In the winter in the Netherlands, it is already dark around these times.

Table 19: significant correlations between the factors

Factor	**Correlation 1**	**Correlation 2**	**Correlation 3**	**Correlation 4**
Age	Years with a driving license (+)			
Driving frequency	Perceived HMI car familiarity (+)			
Touch screen	Adaptive climate control (+)	Adaptive cruise control (+)	Automatic lights (+)	
Transmission type	Adaptive climate control (+)	Adaptive cruise control (+)		
Adaptive climate control	Adaptive cruise control (+)	Automatic lights (+)	Transmission type (+)	Touch screen (+)
Adaptive cruise control	Automatic lights (+)	Touch screen (+)	Transmission type (+)	Adaptive climate control (+)
Peak hour	Dark (+)			

The results of the Poisson model for the total tasks indicate that the model with the factor perceived HMI familiarity fits significantly better than the intercept-only model. The data is overdispersed, as can be seen by the Deviance and Pearson chi-square test. The coefficients divided by the degrees of freedom are higher than 1. The coefficient of one variable was found to be statistically significant. People seem to overall perform 28% more tasks for every point higher on the perceived HMI car familiarity scale.

The results of the Poisson model for the radio and media category indicate that the model with the factors included fits significantly better than the intercept-only model. The data is slightly overdispersed, as can be seen by the Deviance and Pearson chi-square test. The coefficients divided by the degrees of freedom are higher than 1. Three variables were entered in the final model. Male drivers seem to perform 224% more radio and media tasks than female drivers. Furthermore, for every year that a person is older 2% fewer radio and media tasks are performed. Also, the coefficient for perceived HMI car familiarity is significant and indicates that people perform 44% more radio and media tasks when they scored familiarity one point higher.

The results of the Poisson model for the windshield category indicate that the model with the factors included again fits significantly better than the intercept-only model. The data is slightly overdispersed, as seen by the Deviance and Pearson chi-square test. The coefficients divided by the degrees of freedom are higher than 1. The coefficients of four variables were found to be statistically significant. People seem to perform 773% more windshield tasks in rainy conditions and 817% more windshield tasks in sunny conditions. In rainy conditions, this is due to the windshield wipers and in sunny conditions due to the sun visor being used more often. Furthermore, when people drive in cars with an automatic transmission they perform 63% less windshield tasks. This could be because cars with an automatic transmission more often have automatic windshield wipers. This variable was not included in this study so this correlation cannot be tested. Also, the regression indicates that men perform 47% less windshield tasks.

The results of the Poisson model for the climate control category indicate that the model with the factors included fits significantly better than the intercept-only model. The data is overdispersed, however, as can be seen by the Deviance and Pearson chi-square test. The coefficients divided by the degrees of freedom are higher than 1. The final model consists of two variables. The coefficient for gender is significant only at the 10% level. It is included since it is interpretable. The data suggests that men use 42% less climate control tasks than women. This could be because women are more perceptible to cold (Kaikaew et al., 2018). The data also suggests that people use 65% fewer tasks in cars with adaptive cruise control. This could be because modern cars perform more tasks automatically. In appendix E, it can be seen that both automatic lights and adaptive cruise control have a statistically significant coefficient in the individual models. However, since these variables have a strong positive correlation, only one is entered in the final model. Entering both of these variables results in a non-significant model.

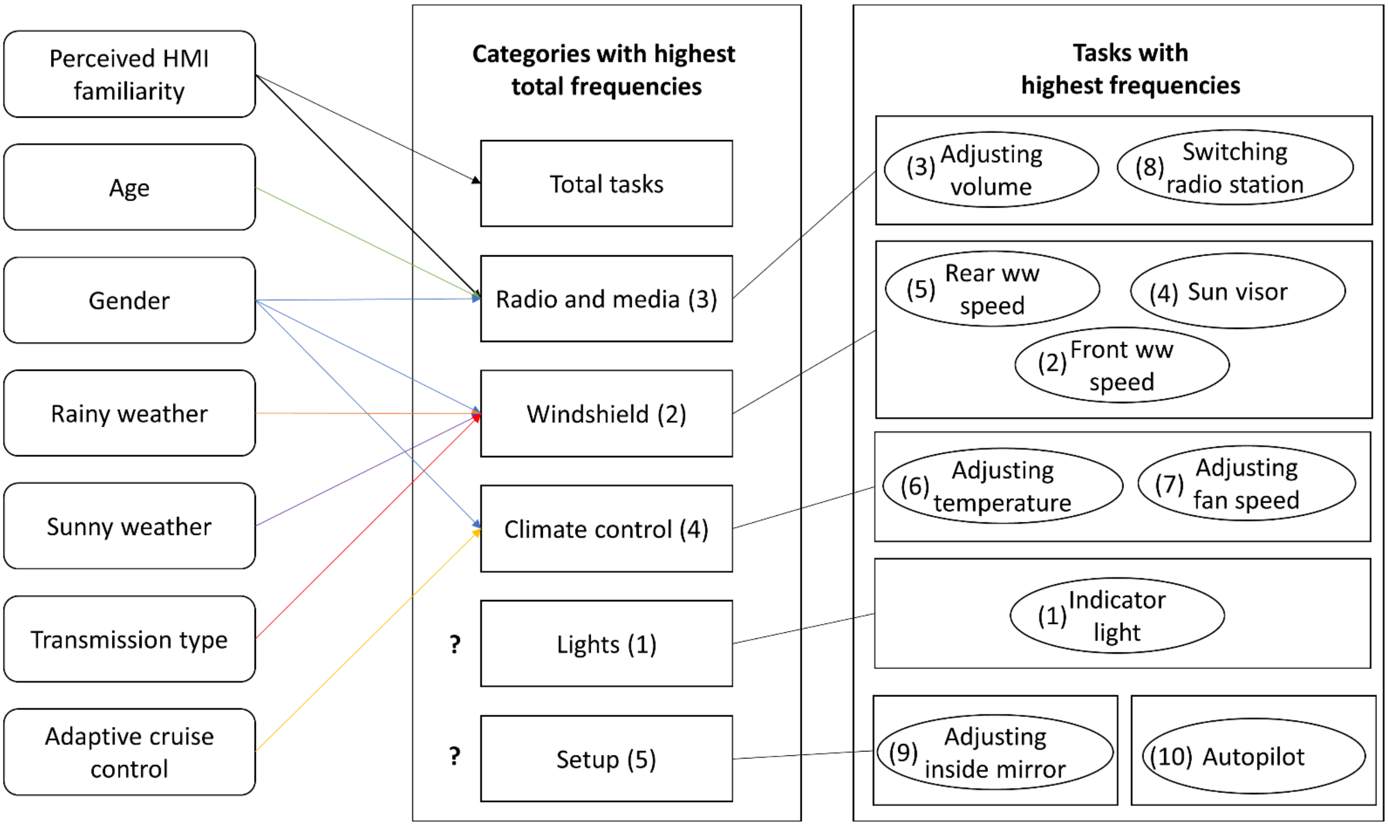
Table 20: Poisson regression models

	Total tasks Model 1 (N=30)		**Radio and media Model 2 (N=30)**		**Windshield Model 3 (N=30)**		**Climate control Model 4 (N=30)**	
	Value	P-value	Value	P-value	Value	P-value	Value	P-value
Likelihood ratio chi-squared	7.545	0.006	18.220	<0.001	54.118	<0.001	10.463	0.005
	Value	Value/df	Value	Value/df	Value	Value/df	Value	Value/df
Deviance	62.158	2.220	32.478	1.249	33.179	1.327	43.747	1.620
Pearson Chi-square	69.145	2.469	35.116	1.351	34.366	1.375	46.818	1.734
	Exp(B)	P-value	Exp(B)	P-value	Exp(B)	P-value	Exp(B)	P-value
Perceived HMI familiarity	1.276	0.009	1.436	0.048				
Age			0.980	0.060				
Gender			3.239	0.016	0.534	0.034	0.576	0.091
Rainy					8.725	<0.001		
Sunny					9.167	<0.001		
Transmission type					0.372	0.004		
Adaptive cruise control							0.355	0.020

*trip duration is used as the offset variable

Summary results

- The figure below shows a summary of the results.
- The factors that statistically significantly influence the total task frequencies of different categories are shown.
- The categories with the highest total task frequencies are presented and ranked, as well as the 10 tasks with the highest frequencies.
- It is indicated which tasks are part of which categories



5. Discussion

5.1 Results

The results of this study revealed the most frequently performed tasks using car HMIs and found that these frequencies differ when people drive their own car compared to an unfamiliar car. When looking at all tasks combined excluding the front windshield wiper and indicator light tasks, this study indicates that people perform more tasks in their own car. This is likely because participants are more comfortable in familiar cars and therefore more confident to perform these tasks safely while driving. This can be seen as a form of self-regulation where people only perform tasks when they think they can safely perform them (Wandtner et al., 2016).

For example, the results indicate that people use more climate control tasks in their own car than in an unfamiliar car. This could have multiple reasons. Even though the participants were told to act as if the car was theirs, they still might have been scared to change the settings. However, this is unlikely, since a lot of other tasks did not show a difference in average frequency between the two trips. Another reason could be that the “unfamiliar car” included climate control while not all familiar cars “included” this feature. Adaptive climate control allows people to set a temperature and the car will keep the air temperature at the set temperature. Therefore, there is no need to constantly adjust the warmth slider or fan speed slider that are usually found in cheaper or older cars. However, when testing if adaptive climate control influences the frequency with which climate control tasks are performed, the results did not indicate a significant relationship. This could be due to the small sample size. Another reason could be the fact that the researcher had to drive the car to the participants, which resulted in the unfamiliar car being heated up already to a comfortable temperature, such that the participants did not feel the need to change any climate control settings. Furthermore, during the trips, participants more often adjusted their mirrors and seat when driving the unfamiliar car. This is in line with expectations since the mirror and seat position in people’s own cars are usually already set. An interesting finding is that most people changed their mirror and seat position before the trip started but still made changes during the trip as well.

The results from the Poisson regressions are in line with these results, as the perceived HMI car familiarity was found to have a relation with the task frequencies. Higher perceived familiarity with a car HMI seems to increase the total frequency with which tasks are performed. This factor increases the frequency of radio and media tasks the most.

Unfamiliar cars also seem to result in increased distraction when the design of the car is unintuitive. Multiple participants struggled to set the front windshield wipers to an interval due to the unfamiliar design of the unfamiliar car. This shows that a lack of standardization can result in extra distractions. The importance of standardization has been mentioned in other research as well (Ruzic et al., 2022; Uhlving et al., 2023). The findings are also in line with other research on the impact of unfamiliar cars, which shows that unfamiliar cars reduce driving performance (Liu & Hansen, 2019; Lee et. al., 2005; Chisholm et al., 2008). However, the difference in average frequency for the front windshield wiper task between the two trips was not statistically significant. This is a surprising result given that the average frequency is much higher but it is the result of a small sample and a high standard deviation. This study looked at the extreme side of unfamiliarity as the participants did not get any practice time and all but one of the participants drove the car for the first time. However, no conclusions can be drawn about the impact on safety since this study did not include performance measures such as total eyes-off-the-road time that are found to be predictors of crash risk (Green., 1998).

The results also indicate that age influences the frequency of radio and media tasks. Older people in general performed fewer tasks in this category. This is in line with the expectations, as in general, older people struggle more to perform HMI tasks (Cooper et al., 2019). As radio and media tasks are optional and mostly for entertainment purposes, it seems that older people choose to perform these tasks less. Unsurprisingly, sunny and rainy conditions increase the frequency of the windshield tasks that are performed, which is also in line with expectations. People more often use windshield wipers in rainy conditions and more often use the sun visor in sunny conditions. Furthermore, people performed fewer windshield and climate control tasks in more modern cars. This could be because these cars can perform more tasks automatically but more research on this is needed. Gender also influenced the task frequencies, as men seem to use significantly more radio and media tasks while women seem to perform more windshield and climate control tasks. This last relationship could be related to the fact that women are more perceptible to cold (Kaikaew et al., 2018). The expectation is that the outside temperature also influences the number of climate control tasks that are performed. However, this could not be measured in this study since all trips were driven during the winter. Therefore, there was not a large enough variation in temperatures between participants.

Since research on the prevalence of different types of distraction is still lacking, it is hard to compare the results of this study to other prevalence studies (Stelling & Hagenzieker., 2015). Still, the results of this study seem to fit well within the overall literature on HMI distraction. Metz et al. (2014) found that people more often performed demanding visual-manual tasks during standstill. The results of this study also indicate this. People performed more climate control tasks during standstill than radio and media tasks. The most performed radio and media tasks in this study consisted of changing the volume, which could be seen as a less demanding task. The climate control tasks were a bit more demanding since participants needed to look away from the road while performing these tasks. A difference between this study and Metz et al. is that they found smaller differences in average frequencies between the road types. Furthermore, the results of this study are in line with the findings by Dingus et al. (2016), who also found that climate control and radio and media tasks are among the most prevalent tasks. The average frequencies as mentioned in this research should be interpreted with the external circumstances such as weather conditions during the trips and the sample statistics in mind, as these can influence the results.

5.2 Strengths and limitations

A strong point of this research is that manual observations are performed during real-world driving in a naturalistic setting. This resulted in detailed observations for many different tasks. Since real-world driving is observed, this also resulted in ecologically valid data. Participants were also asked if they felt influenced during the trip and most participants stated that this was not the case and that they displayed their usual behaviour. Furthermore, the method used in this study to perform the observations is well substantiated by literature and tested with a small pilot before the real observations took place.

There are also strong points in the data itself. A good mix of different types of cars were driven during the observations and a lot of data was collected even with a smaller sample size. In addition to the good mix of different types of cars, 28 out of the 30 participants also drove the same car with the same starting settings. This makes it possible to accurately compare the results of the participants. Furthermore, this study includes many analyses that were performed to explore the data in detail. This exploration could be the foundation for further research.

This study has looked at the frequency with which different internal HMI tasks are performed during real-world driving and how car familiarity impacts this. Since participants answered to statements to measure car familiarity, perceived car familiarity was measured and not objective car familiarity. To measure the perceived car familiarity, this study used a Likert scale. To be able to use this scale as an interval scale in the regressions, the assumption was made that the steps between all the answers are of equal size. However, it is always debatable if Likert scales can be treated as interval scales (Wu & Leung, 2017). That said, using Likert scales is very common in psychological research and is mostly accepted in the field of research.

In this study, the unfamiliar car that was used for 28 out of the 30 participants was a Seat Toledo from 2014. A limitation of using this car is that not all modern functions can be studied. Since the Seat did not include a touch screen or car play, some tasks that can be performed in newer cars were absent. The unfamiliar car data is therefore not fully representative for the new generation of cars.

Furthermore, this study made use of a relatively small convenience sample. This should be kept in mind when looking at the statistical analyses that were performed. A small sample could result in violating some of the assumptions of parametric tests and therefore skewed results. The results of the Poisson regressions do show some overdispersion and therefore do not meet the assumptions perfectly. Also, the fact that a convenience sample is used means that the results are not directly transferrable to the population, but they should be a good first indication. Since the counting has been done manually, mistakes could have been made. No cameras were present during the trips so it was not possible to fix mistakes later on. Also, some people found it difficult to follow the instructions of no conversations. Further research could remedy some of these problems by using cameras and tracking people for longer periods.

5.3 Further research

Future research could focus on extending this study by using a bigger sample and by tracking people for longer periods, as this can result in more accurate data for the population and for the individual participants. Also, cameras might be used to eliminate counting errors. It could create a more comfortable situation for the participants as well. Some participants found it scary that someone was observing them and found it uncomfortable to ignore the researcher fully. This study can also be extended by performing observations under different external circumstances and in different countries to see if there are any differences. Further research could also focus on the impact of factors that were not included in this study, such as psychological factors. These could be factors such as people’s perceived aggressiveness during driving. Also, questions about driving styles could be asked.

More research is also needed on the car familiarity scale, particularly about what is measured exactly by different types of questions and if these measurements are accurate. This study already gave the first insights into this question by performing a factor analysis and finding out that three of the four statements used in this study indeed seem to measure an underlying perceived HMI factor whilst the statement about how often people drive a car seems to measure a different factor. By using latent factors, it is more clear for the respondents what is being asked and accidental errors can cancel each other out.

Further research could also focus on the knowledge gaps in chapter 2.5 that are not addressed in this study. This includes conducting studies on the effect of performing HMI tasks on driving performance using newer simulators, conducting a study on the effect of performing HMI tasks on driving performance in the Netherlands and conducting more studies on the effect of repetition with an HMI on driving performance using different performance measures.

6. Conclusion

6.1 Answers to the research questions

Euro NCAP and RDW have set the target to encourage regulations in the field of passenger car HMIs (Euro NCAP, 2022). This could help to reduce the traffic fatalities that are caused by distraction (Stelling & Hagenzieker, 2015; United States Department of Transportation, 2023). However, knowledge was missing on the frequencies with which different HMI tasks are performed. This posed a problem for the development of the assessment methodology that RDW and Euro NCAP are working on. This study aimed to contribute to solving this problem by manually observing people's HMI behaviour during real-world driving and in a naturalistic setting. In particular, the study focused on the frequencies with which people perform different tasks and what factors influence this. This section answers the sub-questions and main research question.

What are relevant internal HMI tasks that should be considered?

Only tasks at the tactical level as described by Michon (1985) should be considered for this study. For this study, tactical tasks are seen as tasks that still require some thought, are performed during a trip, do not happen fully automatically and take in the order of seconds to perform. These types of tasks can be distracting (Strayer et al., 2017) and are therefore the focus of this research. Examples of these tasks are adjusting the volume or switching to another radio station.

Which factors could be relevant in influencing the frequency with which internal HMI tasks are performed?

The factors that could influence the task frequencies were found to be gender, age, driving frequency, amount of years owning a driver's license, route familiarity, car familiarity, weather and time and day of driving. Some car characteristics such as the climate control type, cruise control type, transmission type, automatic lights and a touch screen could also influence the task frequencies.

Which internal HMI tasks are performed most frequently?

Based on the data that was gathered, the 10 most frequently used tasks included using the indicator light, using the front windshield wipers, adjusting the volume, moving the sun visor, using the rear windshield wiper, adjusting the climate control temperature, adjusting the fan speed, switching radio station, adjusting the inside mirror and enabling or disabling the autopilot. Besides the indicator light, this shows that people mostly perform tasks related to the windshield, radio and media system and climate control system.

What influence does car familiarity have on the frequency with which people perform internal HMI tasks?

Car familiarity seems to have an impact on the frequencies with which certain types of tasks are performed. The observations show that intuitive design and familiarity are important. When this is lacking, people can struggle to perform certain tasks which results in them being distracted more often. This was the case with the front windshield wiper task during this study.

People seem to perform more tasks in familiar cars when excluding the front windshield wiper and indicator tasks. Furthermore, people seem to statistically significantly use more climate control tasks in a familiar car compared to an unfamiliar car. In the unfamiliar car, people used statistically significantly more tasks related to setting up a car. Tasks in this category include adjusting the mirrors and adjusting the seat position. Most people already tried setting this up before the trip, but often, it still required some tweaking during the trip.

Which factors influence the frequency with which people perform internal HMI tasks?

Perceived familiarity with a car HMI seems to increase the total frequency with which tasks are performed. This factor seems to increase the frequency of radio and media tasks the most. People who are more familiar with a car HMI seem to perform more tasks. Male drivers also seem to perform more radio and media tasks, whilst female drivers seem to perform more windshield and climate control related tasks. Furthermore, age seems to influence the frequency of radio and media tasks. Older people in general performed less tasks in this category. Sunny and rainy conditions increase the frequency of the windshield tasks that are performed. For sunny conditions, this is due to the increase of the use of the sun visor. For rainy conditions, this is due to the increase in the use of the windshield wipers. Lastly, two car characteristics were found to influence the frequencies. People who drive in cars with an automatic transmission seem to use less windshield tasks and people who drive in cars with adaptive cruise control seem to perform less climate control tasks. A reason for this could be that these more modern cars perform more tasks automatically

What user interface tasks are performed relatively frequently during real-world driving in passenger cars and what factors influence the frequencies?

The prevalence of different forms of distraction is not well-studied in the current literature (Stelling & Hagenzieker., 2015). This is especially true for internal HMI task prevalence. Apart from studies by Metz et al. (2014) and Dingus et al. (2016), most studies in the current literature are focused on the impact of distraction on different performance measures and innovative technologies to reduce distraction. By using the method that was proposed in this study, it was possible to measure the prevalence of different internal HMI tasks. This thesis contributes to the literature by identifying the frequencies with which different tasks are performed and identifying which factors could influence this. It thereby increases the understanding of internal HMI use in cars. This study also provides a methodology that can be used by future studies on the topic of distraction prevalence. The findings of this study indicate that the most used tasks are the indicator light, tasks in the windshield category and tasks in the radio and media and climate control categories. The differences in frequencies between different tasks are relatively large and even with such a small sample size, statistically significant differences were found between the top 10 most used tasks. In general, people seem to perform around 12 tasks per hour in familiar cars versus 9 tasks per hour in unfamiliar cars when excluding the front windshield wiper and indicator light tasks. This shows that there are a significant amount of distraction moments per hour due to internal HMI use. Different factors seem to impact the frequencies of different types of tasks. The factors that had an influence included car familiarity, gender, age and weather conditions. The type of car also seems to impact the task frequencies, but more research on this is needed. An incidental finding of this study is that unfamiliar cars can result in an increased number of tasks performed when the design of the car is unintuitive and can therefore lead to increased distraction. Future research on the safety impact of performing internal HMI tasks can use the results of this study to better understand the relative risks of the different tasks in real-world driving.

6.2 Practical relevance

The results of this study can be used by Euro NCAP and RDW as input for new regulations and their internal HMI safety assessment methodology. To reduce driver distraction, this new regulation could encourage manufacturers to have these tasks easily available during driving and to standardize these tasks across all cars. By publishing the ratings produced by the assessment methodology, manufacturers can be influenced to improve the safety of their car HMIs. These ratings can also help consumers to make better buying decisions. Ultimately, this should lead to safer car HMIs that reduce driver distraction in both familiar and unfamiliar cars. Car-sharing companies can also use these results to better understand the safety risks involved with the internal HMI of different cars. This could help their car purchase decisions for their car fleet. Peer-to-peer car-sharing companies could use the results to create regulations for their platform. They could disallow certain cars on their platform if they have a low rating. The same can be applied to car rental companies. The results of this study can also directly be used by car manufacturers. They could use the results of this study when designing their internal HMI layout to improve the user experience by making sure that the most used tasks are easy to perform. Also, based on the Poisson regressions, rough predictions can be made on the task frequencies of different populations with different characteristics.

6.3 Recommendations

The recommendation is to encourage manufacturers to have the most frequently used tasks, which are the indicator light task, tasks in the windshield category and tasks in the radio and media and climate control categories easily available. This means that they should not be several levels deep in menus and should be able to be performed within a short time frame. Also, the recommendation is to standardize how these tasks are performed, as this study shows that unintuitive designs can create extra distractions. This need for standardization has been mentioned in multiple other studies as well (Uhlving et al., 2023; Ruzic, 2022). The results of this study can also be used as input for the assessment methodology of RDW and Euro NCAP that rates the safety level of car HMIs. The recommendation is that the most frequently used tasks should be weighed as more important for the safety level and should be taken into account when assessing the rating. It is also recommended that the results of this study and the safety ratings are shared with the consumers and car industry, such that car manufacturers have a greater stimulus to improve the safety of their HMIs. Another recommendation would be for car-sharing and car-rental companies to use these results to improve the safety of their car fleet. They should only allow cars with intuitive internal HMI designs with a high safety rating. This is especially important for these companies since most consumers will drive their cars for the first time. This study has shown that unintuitive designs in unfamiliar cars can increase distraction.

6.4 Reflection

The practical side of this study was quite challenging. It took a large amount of time to set up the experiment. It was especially difficult and time-consuming to arrange the Seat Toledo that was used in this study. Also, gift cars had to be arranged and the procedure for getting ethical permission took longer than expected. This was initially due to the fact that it was unclear which unfamiliar car was going to be used and how it would be insured. This study also resulted in more data than anticipated. Since the observations were best performed with pen and paper, it took a lot of time to enter all the data in the computer files. It was also time-consuming to perform all the observations manually, as many hours were spent driving along with participants. After the data was entered, a lot of calculations still needed to be done to calculate the aggregated data such as the total tasks performed per category. This all took more time than anticipated. Moreover, due to the small sample size, it was hard to conduct some of the statistical analyses. It was especially hard to come up with a good method for the Poisson regression, as the study contained too many variables in comparison to the sample size. More participants could have resulted in more statistically significant relations but due to the large amount of time that was needed for each participant, it was unfeasible for this study. In the end, it was worth the work, since the results of the study can be very useful for RDW, Euro NCAP, car manufacturers, car-sharing companies and rental companies to name a few. This study also fills the knowledge gap on the prevalence of different forms of internal HMI distraction.

Bibliography

AAA Center for Driving Safety and Technology – AAA Exchange. (z.d.). https://exchange.aaa.com/safety/distracted-driving/aaa-center-for-driving-safety-and-technology/

Aiken, L. S., West, S. G., Pitts, S. C., Baraldi, A. N., & Wurpts, I. C. (2012b). Multiple Linear Regression. Handbook of Psychology, Second Edition, 2, 1. c

Allianz Direct. Dit zijn de meest verkochte elektrische auto's in 2022 | Allianz Direct. (2023). https://www.allianzdirect.nl/blog/meest-verkochte-elektrische-auto-2022/#:~:text=Op%20%231%20in%20de%20lijst,is%20goed%20voor%20ruim%2036%25.

Arca, A. A., Chouljian, E., & Mouloua, M. (2022). The role of Gender Differences in Distracted Driving Behavior: A Psychophysiological approach. Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting, 66(1), 1280–1284. https://doi.org/10.1177/1071181322661453

Autodelen. (2022, 16 maart). Cijfers en onderzoek. Geraadpleegd op 9 oktober 2022, van https://autodelen.info/cijfers-autodelen

Auto Express. Tesla Model y review - Pictures | Auto Express. (2023a). https://www.autoexpress.co.uk/tesla/model-y/359882/tesla-model-y-review-pictures

Auto Express. Kia Picanto review – Pictures | Auto Express. (2023b). https://www.autoexpress.co.uk/108590/kia-picanto-review-images

Blanco, M., Biever, W., Gallagher, J., & Dingus, T. A. (2006). The impact of secondary task cognitive processing demand on driving performance. Accident Analysis & Prevention, 38(5), 895–906. https://doi.org/10.1016/j.aap.2006.02.015

Bovag. (2022). Ruim 1,2 miljoen automaten erbij sinds 2015. https://mijn.bovag.nl/actueel/nieuws/2022/februari/ruim-1,2-miljoen-automaten-erbij-sinds-2015

Bovag. (2023). Lichte daling autoverkoop in 2022. https://www.bovag.nl/nieuws/lichte-daling-autoverkoop-in-2022

Breitschaft, S. J., Pastukhov, A., & Carbon, C. (2021). Where's My Button? Evaluating the User Experience of Surface Haptics in Featureless Automotive User Interfaces. IEEE Transactions on Haptics, 15(2), 292–303. https://doi.org/10.1109/toh.2021.3131058

Broström, R., Aust, M. L., Wahlberg, L., & Källgren, L. (2013). What drives off-road glance durations during multitasking: capacity, practice or strategy? Proc. of the Third International Conference on Driver Distraction and Inattention. 04-06.09.2013 Göteborg, paper 08-P. https://trid.trb.org/view/1405238

Burdett, B. R., Charlton, S. G., & Starkey, N. J. (2019). Mind wandering during everyday driving: an on-road study. Accident Analysis & Prevention, 122, 76–84. https://doi.org/10.1016/j.aap.2018.10.001

CBS. (2022). Mobiliteitstrend; per rit, vervoerwijzen, reismotief, leeftijd en geslacht. https://opendata.cbs.nl/#/CBS/nl/dataset/84755NED/table?ts=1697116530423

Chisholm, S., Caird, J. K., & J, L. (2008). The effects of practice with MP3 players on driving performance. Accident Analysis & Prevention, 40(2), 704–713. https://doi.org/10.1016/j.aap.2007.09.014

Cooper, J. D., Wheatley, C. L., McCarty, M. M., Motzkus, C. J., Lopes, C. L., Erickson, G. G., Baucom, B. R., Horrey, W. J., & Strayer, D. L. (2019). Age-Related Differences in the Cognitive, Visual, and Temporal Demands of In-Vehicle Information Systems. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01154

Coxe, S., West, S. G., & Aiken, L. S. (2009). The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. Journal Of Personality Assessment, 91(2), 121–136. https://doi.org/10.1080/00223890802634175

Cuentas-Hernandez, S., Li, Y., King, M., & Oviedo-Trespalacios, Ó. (2023). The impact of road traffic context on secondary task engagement while driving. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1139373

Davidse, R.J., Doumen, M.J.A., Duijvenvoorde, K. van & Louwerse, W.J.R. (2011). Bermongevallen in Zeeland: karakteristieken en oplossingsrichtingen. R-2011-20. SWOV, Leidschendam.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Pérez, M. Á., Buchanan-King, M., & Hankey, J. M. (2016). Driver Crash Risk Factors and Prevalence Evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113

Dingus, T. A., Owens, J. M., Guo, F., Fang, Y., Pérez, M. Á., McClafferty, J., Buchanan-King, M., & Fitch, G. M. (2019). The prevalence of and crash risk associated with primarily cognitive secondary tasks. Safety Science, 119, 98–105. https://doi.org/10.1016/j.ssci.2019.01.005

Euro NCAP. Euro NCAP Vision 2030: a Safer Future for Mobility. (2022). https://www.euroncap.com/en/press-media/press-releases/euro-ncap-vision-2030-a-safer-future-for-mobility/

Euro NCAP. How To Read The Stars. (2023). https://www.euroncap.com/en/about-euro-ncap/how-to-read-the-stars/

Gillin, A., & Gillin, A. (2023). Here are the top causes of distracted driving. GJEL Accident Attorneys. https://www.gjel.com/blog/causes-distracted-driving.html

Green, P. (1998). Visual and Task Demands of Driver Information Systems (Technical ReportUMTRI-98-16), Ann Arbor, MI: The University of Michigan Transportation Research Institute(in preparation).

Green, P. (1999). The 15-second rule for driver information systems. Proceedings of the ITS America Ninth Annual Meeting. Washington, D.C.: Intelligent Transportation Society of America

de Haas, M. (2020). Autogebruik over de dag – naar sector en onderwijs. Kennisinstituut voor Mobiliteitsbeleid. https://www.kimnet.nl/binaries/kimnet/documenten/notities/2020/10/15/spreiding-autogebruik-over-de-dag-en-verdeling-naar-bedrijfssector-en-onderwijs/Autogebruik_over_de_dag_naar_sector_en_onderwijs.pdf

Hamari, J., Sjöklint, M., & Ukkonen, A. (2015). The sharing economy: Why people participate in collaborative consumption. Journal of the Association for Information Science and Technology, 67(9), 2047–2059. https://doi.org/10.1002/asi.23552

Harms, I. M., Burdett, B. R., & Charlton, S. G. (2021). The Role of Route Familiarity in Traffic Participants' Behaviour and Transport Psychology Research: A Systematic review. Transportation research interdisciplinary perspectives, 9, 100331. https://doi.org/10.1016/j.trip.2021.100331

Harms, I. M., van Dijken, J. H., Brookhuis, K. A., & de Waard, D. (2019). Walking without awareness. Frontiers in Psychology, 10:1846, 1-12. doi:10.3389/ fpsyg.2019.01846

van Heukelingen, C., Schakel, L. & Molnár-in't Veld, H (2023). 5. Gemiddelde afgelegde afstand per verplaatsing. Centraal Bureau voor de Statistiek. https://www.cbs.nl/nl-nl/longread/rapportages/2023/onderweg-in-nederland--odin---2022-plausibiliteitsrapportage/5-gemiddelde-afgelegde-afstand-per-verplaatsing

Horberry, T., Anderson, J., Regan, M., Triggs, T. J., & Brown, J. M. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accident Analysis & Prevention, 38(1), 185–191. https://doi.org/10.1016/j.aap.2005.09.007

Iqbal, S. T., E. Horvitz, Y.-C. Ju, and E. Mathews (2011). "Hang on a sec!: effects of proactive mediation of phone conversations while driving". In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM. 463–472.

Kaikaew, K., Van Den Beukel, J. C., Neggers, S. J. C. M. M., Themmen, A. P. N., Visser, J. A., & Grefhorst, A. (2018). Sex difference in cold perception and shivering onset upon gradual cold exposure. Journal Of Thermal Biology, 77, 137–144. https://doi.org/10.1016/j.jtherbio.2018.08.016

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. The New England Journal of Medicine, 370(1), 54–59. https://doi.org/10.1056/nejmsa1204142

Kun, A. L. (2018). Human-Machine Interaction for Vehicles: Review and Outlook. Foundations and Trends in Human-computer Interaction, 11(4), 201–293. https://doi.org/10.1561/1100000069

Lee, S. E., Dingus, T. A., Klauer, S. G., Neale, V. L., & Sudweeks, J. (2005). Naturalistic data collection of driver performance in familiar and unfamiliar vehicles. Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting, 49(22), 1994–1998. https://doi.org/10.1177/154193120504902223

Liu, Y., & Hansen, J. H. L. (2019). Analysis of Driving Performance Based on Driver Experience and Vehicle Familiarity: A UTDrive/Mobile-UTDrive App Study. SAE International Journal Of Transportation Safety, 7(2). https://doi.org/10.4271/09-07-02-0010

Lo, V. S. Y., & Green, P. E. (2013). Development and Evaluation of Automotive Speech Interfaces: Useful Information from the Human Factors and the Related Literature. International Journal of Vehicular Technology, 2013, 1–13. https://doi.org/10.1155/2013/924170

Ma, J., Li, J., & Gong, Z. (2022). Evaluation of driver distraction from in-vehicle information Systems: A simulator study of interaction modes and secondary tasks classes on eight production cars. International Journal of Industrial Ergonomics, 92, 103380. https://doi.org/10.1016/j.ergon.2022.103380

Metz, B., Landau, A., & Just, M. (2014). Frequency of secondary tasks in driving – results from naturalistic driving data. Safety Science, 68, 195–203. https://doi.org/10.1016/j.ssci.2014.04.002

Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In Springer eBooks (pp. 485–524). https://doi.org/10.1007/978-1-4613-2173-6_19

NHTSA. The 100-Car Naturalistic Driving Study. (2006). (DOT HS 810 593). https://www.nhtsa.gov/sites/nhtsa.gov/files/100carmain.pdf

Ou, Y., Liu, Y. C., & Shih, F. (2013). Risk prediction model for drivers' in-vehicle activities – Application of task analysis and back-propagation neural network. Transportation Research Part F-traffic Psychology and Behaviour, 18, 83–93. https://doi.org/10.1016/j.trf.2012.12.013

Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. Transportation research procedia, 25, 1796–1805. https://doi.org/10.1016/j.trpro.2017.05.148

Perez, M., Owens, J., Viita, D., Angell, A., Ranney, T. A., Baldwin, G. H. S., Parmer, E., Martin, J., Garrott, W. R., & Mazzae, E. N. (2013, May). Radio tuning effects on visual and driving performance – Simulator and test track studies. (Report No. DOT HS 811 781). Washington, DC: National Highway Traffic Safety Administration.

Prabhakar, G., & Biswas, P. (2021). A Brief Survey on Interactive Automotive UI. Transportation engineering, 6, 100089. https://doi.org/10.1016/j.treng.2021.100089

Rasmussen J (1983) Skills, rules, and knowledge;signals, signs and symbols, and other distinctions inhuman performance models. IEEE Transactions onSystems, Man, and Cybernetics 13(3): 257-66
(PDF) Rasmussen's model of human behavior in laparoscopy training. Available from: https://www.researchgate.net/publication/10714076_Rasmussen%27s_model_of_human_behavior_in_laparoscopy_training [accessed Oct 03 2023].

Redelmeier, D. A., & Tibshirani, R. (1997). Association between Cellular-Telephone calls and motor vehicle collisions. The New England Journal of Medicine, 336(7), 453–458. https://doi.org/10.1056/nejm199702133360701

Ružić, D. (2022). The ergonomic aspects of a passenger car climate control panel. IOP Conference Series: Materials Science and Engineering, 1271(1), 012019. https://doi.org/10.1088/1757-899x/1271/1/012019

SAE. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. (2021). *SAE international*. https://doi.org/10.4271/j3016_202104

Schmidt, A., Dey, A. K., Kun, A. L., & Spiessl, W. (2010). Automotive user interfaces. https://doi.org/10.1145/1753846.1753949

Stelling, A., & Hagenzieker, M. P. (2015). Schatting aantal verkeersdoden door afleiding. In SWOV (R-2015-13). SWOV. https://swov.nl/system/files/publication-downloads/r-2015-13.pdf

Strayer, D. L., Cooper, J. D., Goethe, R. M., McCarty, M. M., Getty, D., & Biondi, F. (2017). Visual and Cognitive Demands of Using In-Vehicle Infotainment Systems. Foundation for Traffic Safety, 202-638–5944. https://aaafoundation.org/wp-content/uploads/2017/11/VisualandCognitive.pdf

Universiteit Utrecht. LibGuides: Google Scholar: Verdieping. (2022). https://libguides.library.uu.nl/google-scholar/verdieping#s-lg-box-wrapper-16993799

Uhlving, V. M., Weyde, K. V., Hagenzieker, M., Aasvik, O., & Rostoft, M. T. (2023). Advanced driver support systems and screen interfaces in car - How do they affect driver behaviour? TØI, 978-82-480-1519–2. https://www.toi.no/getfile.php?mmfileid=77085

U.S. Department of Transportation. (2023). Distracted driving in 2021 (DOT HS 813 443). NHTSA. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813443

ViBilägare. Physical buttons outperform touchscreens in new cars, test finds. (2022). https://www.vibilagare.se/english/physical-buttons-outperform-touchscreens-new-cars-test-finds

Wandtner, B., Schumacher, M., & Schmidt, E. (2016). The role of self-regulation in the context of driver distraction: A simulator study. Traffic Injury Prevention, 17(5), 472–479. https://doi.org/10.1080/15389588.2015.1102231

Wang, Y., Reimer, B., Mehler, B., Zhang, J., Mehler, A., & Coughlin, J. F. (2010). The impact of repeated cognitive tasks on driving performance and visual attention. In Advances in human factors and ergonomics series (pp. 668–678). https://doi.org/10.1201/ebk1439834992-70

Wu, Y., Boyle, L. N., McGehee, D. V., Angell, L., & Foley, J. D. (2012). Task analysis of vehicle entry and backing. https://doi.org/10.1145/2390256.2390290

Wu, H., & Leung, S. O. (2017). Can Likert Scales be Treated as Interval Scales?—A Simulation Study. *Journal of Social Service Research*, *43*(4), 527–532. https://doi.org/10.1080/01488376.2017.1329775

Wynne, R. A., Beanland, V., & Salmon, P. M. (2019). Systematic Review of driving Simulator Validation Studies. *Safety Science*, *117*, 138–151. https://doi.org/10.1016/j.ssci.2019.04.004

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. Tutorials in Quantitative Methods for Psychology, 9(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p07

Appendix A: Scientific paper

Towards safe use of general controls in cars

A real-world driving study assessing internal HMI task frequencies and influencing factors

D.A.M. Auerbach

Abstract - The human-machine interfaces (HMIs) in passenger cars have become more complex over the years, with touch screens replacing physical buttons and more options becoming available. This can result in multiple layers of menus that have to be navigated to perform simple tasks and therefore increase distraction. In addition to this, there is no standardisation for passenger car HMIs. To face these problems, new standards will have to be created. However, knowledge regarding the prevalence of different internal HMI tasks is missing. This complicates efforts to create new standards since it is unclear which internal HMI tasks are most important to assess. It also complicates efforts to create an assessment methodology to rate the safety of different internal HMI interfaces, as tasks that are used more often could be seen as more important. This thesis fills the knowledge gap by identifying the most used internal HMI tasks and the factors that influence their frequencies based on real-world driving observations in a naturalistic setting. In particular, this study especially looks at the influence of car familiarity. Multiple data analysis methods, including Poisson regressions, ANOVA's, paired t-tests and factor analysis are used to obtain the required knowledge. In general, people seem to perform around 12 tasks per hour in familiar cars versus 9 tasks per hour in unfamiliar cars, but significant differences between tasks exist. The results indicate that the most performed is using the indicator light. Other frequently performed tasks fall into the windshield, radio and media and climate control categories. The factors that influenced the task frequencies included car familiarity, gender, age and weather conditions. The type of car also seems to impact the task frequencies, but more research on this is needed.

Keywords – Real-world observations, Naturalistic driving, Internal HMI Distraction, Task Frequency, Passenger cars, Poisson regression

I. Introduction

The human-machine interfaces (HMIs) in passenger cars have become more complex over the years, with touch screens replacing physical buttons and more options becoming available. These recent HMI designs might look clean and provide more options, but they also take longer to operate and can result in navigating through multiple layers of menus to perform simple tasks [1]. The distraction that occurs while operating these more complex interfaces could impose problems regarding road traffic safety [2]. In the Netherlands, it is estimated that 24-31% of car crashes occur because of some form of distracted driving [3], which includes operating car HMIs. The fact that these car HMIs are not standardized is often mentioned as a problem in literature [4] [5], as it can result in reduced safety particularly when driving unfamiliar cars. This is backed up by current research, which finds that crash risk increases when driving unfamiliar cars [6] [7]. This lack of standardization could become an even bigger problem due to a growing amount of people who are using services such as car sharing, where they are driving unfamiliar cars [8].

RDW and Euro NCAP want to address these issues and set new standards [9]. Euro NCAP wants to do this by using its star rating system [10]. This system rates the level of safety of different aspects of cars and the goal is to also rate the safety of car HMIs. These safety ratings can then be used to influence car manufacturers.

However, as can be concluded from Chapter 2, knowledge regarding the prevalence of different internal HMI tasks is missing. This complicates efforts to create new standards since it is unclear which internal HMI tasks are most important to assess.

This thesis aims to contribute to this knowledge gap by investigating how often different tasks are performed using the internal HMI of a car and what factors might influence this. In particular, this study especially looks at the influence of car familiarity. To achieve this, manual observations are performed during real-world naturalistic driving and the influence of different factors are studied with Poisson regressions. The results of this study can be used by Euro NCAP and RDW for their star rating system and can be used as input for new regulations. The star ratings can help consumers make better buying decisions and influence car manufacturers. To reduce driver distraction, new regulations could encourage manufacturers to create designs in which the most performed tasks are easily available and standardized across cars. Ultimately, this should lead to safer car HMIs that reduce driver distraction in both familiar and unfamiliar cars. To reach the objective of the study, the following research question was proposed: "*What are the most performed internal HMI tasks in passenger cars and what factors (Socio-demographic, contextual, car familiarity, car characteristics) influence the frequency with which tasks are performed?*"

II. Literature Review

Due to the increasing complexity and importance of car HMIs, an increasing amount of interest has emerged in this topic [11]. In this Chapter, a literature review is performed on the current state-of-the-art research regarding passenger car HMIs. This literature review is performed to find knowledge gaps. To find the literature, Google Scholar was used and studies were selected based on relevance sorting and how well they fit the topic. This resulted in the studies that are discussed in the next sections.

a. HMI technologies

Several innovative technologies for HMIs in cars have been researched. [12] and [13] performed literature research and found that speech recognition and non-visual interfaces are likely to have a positive influence on reducing driver distraction. However, the results regarding task completion time were mixed. [14] also looked at voice control using a driving simulator and found it to be a suitable technology for basic tasks and tasks of medium complexity.

Many cars today also include screens in which all or most tasks are integrated. [4] studied the impact of these screens on driving behaviour to assess their potential effect on safety. To measure this, they performed drives on a test track and made use of an eye-tracking system. Results indicate that these integrated touch screens can be potentially dangerous. [14] found that touch screens can be more suitable for complex tasks than standard buttons. However, the latter seems to be more suitable for basic tasks.

[15] looked at the influence of alert notifications and mediation on calls while driving. This technology alerts the driver on upcoming critical road conditions and can mediate a call by putting it on hold during these conditions. The results indicate that this alerting technology can improve driving performance and was preferred by most of the participants.

Furthermore, [12] looked at some advantages and disadvantages of virtual touch interfaces and wearable devices. The findings indicate that wearable devices are not yet researched well as a way to interact with the user interface of a car and that virtual touch systems are only researched using infrared at the moment. Haptic feedback is another technology that could be implemented in the human interfaces of cars. [16] looked at the influence of haptic feedback on driving performance. They found that haptic feedback did not have a significant influence on the driving task, but did improve the overall user experience.

b. Performance measures

In the current literature, different performance measures are used to evaluate the impact of driver distraction on safety. Most studies that look into performance measures include either real-world driving or driving in a simulator. Out of the eight studies that included driving, five used driving simulators, whilst the other three were based on real-world driving. A study by [5] differs from this since a heuristic approach was used to assess the performance. This means that the researcher analysed and discussed the internal HMI of a typical passenger car based on 5 criteria that are important from an ergonomic and safety point of view. These criteria include findability, reachability, identification, interpretability, operability and understandability. The focus of this research was also only on the climate control panel. The findings of this study indicate that the cars that were investigated lacked standardization regarding the layout and that there was too much cluttering of the controls. Also, many controls required visual guidance for correct operation.

One of the most frequently mentioned performance measures in the literature is the time to complete a task [17] [18] [19] [20] [21]. In the past, studies have suggested a 15-second rule, meaning that the time to complete a task should not exceed 15 seconds [18]. The time to complete a task is an important indicator of crash risk. A study by [17], showed that total eyes-off-the-road time is a predictor of crash risk and that this measure had a high correlation with task completion time. [5] also mentions that HMIs can be tested by analysing the time required for certain tasks using driving simulators. Research also suggests that older drivers have increased task completion times [21].

Two other important performance measures used to evaluate the safety of human-machine interfaces in cars are visual and cognitive demand [15] [21] [22]. To measure this, participants usually have to perform visually and cognitively demanding tasks while driving. First, a baseline score on these tasks is measured. Then, the participants have to engage in HMI tasks while performing the tasks that are used to measure the demand. By comparing the two scores, it becomes clear how demanding the HMI task was. In general, ISO standard tasks are used to determine cognitive and visual demand. The findings of [22] indicate that there are differences in visual and cognitive demands between different tasks. The most demanding task was destination entry for navigation. Voice-based commands demanded lower levels of visual demand but increased total interaction time significantly. There was also a significant difference in demand across the different vehicles. In addition to this, research suggests that older drivers have more difficulty operating infotainment systems in cars resulting in a higher cognitive and visual workload [21]. [15] took a different approach to determine the cognitive load of drivers. They indirectly assessed the cognitive load by looking at the number of missed red lights and stop signs. In addition to this, they also looked at the number of lane departures.

Another way to measure driving safety is to take into account the number of collisions. However, this is not a very suitable indicator for testing due to the low rate of collisions in both real-life trials and most driving simulator trials [23]. Therefore, the common approach is to use performance measures that are known to increase crash risk. A study by [19] is one of the few that took into account the number of collisions. This study made use of a driving simulator and only one study using real-world driving that included the number of collisions was found [6].

Off-road glance duration is yet another frequently used performance measure [19] [20] [24]. The results of [24] indicate that off-road glance durations while performing different tasks vary significantly between different people, that practice reduces the off-road glance time and that individual glance strategies play a big role in glance duration regardless of in-vehicle task complexity. Furthermore, [20] found the number of glances to be the most important predictor of crash risk due to performing in-vehicle information system tasks.

[20] also found that longitudinal velocity was found to be the most important predictor for crash risk when performing traditional in-vehicle tasks. Velocity is also included in a study by [25]. This study indicated that physiological measures such as heart rate and skin conductance are more sensitive to changes in workload than driving performance measures such as the standard deviation of velocity and lane position. [6] also used velocity as well as lane position as a performance measure. Some other performance measures that can be found in the literature include turning errors, response time and number of hand movements [15] [20].

As can be seen from the literature, the performance measures used to assess driving performance widely vary between studies. A literature review by [26] also mentions this but also states that there is not a single driving performance measure that can accurately describe all distraction aspects. Most studies do show that distracted driving negatively influences a variety of these performance measures.

c. Safety impact of Internal HMI tasks

Looking at the literature that studies the effect of different tasks on driving performance and cognitive demand, it becomes clear that certain tasks are more often mentioned than others. HMI tasks that are often mentioned are making phone calls, radio tuning and setting up navigation. Of the seven studies that were identified, five mentioned the impact of phone calls on driving performance or driver distraction [14] [22] [27] [28] [29]. Three studies mentioned the effects of navigation tasks [2] [14] [22] and five studies mentioned the effects of radio or media tasks [14] [22] [28] [29] [30]. The consensus found in the literature is that these tasks are found to decrease driving performance and increase cognitive and visual demand. Research regarding phone calls includes both hands-free calling as well as hand-held calling. [28] found no significant safety advantage of hands-free over hand-held calling.

The effects of more traditional in-vehicle tasks such as the operation of the air-conditioner and opening or closing windows are less common in literature. [14] performed the only study that was found for this review in which the effects of operating the air conditioning were studied. This study found that different types of tasks are more or less suited to certain types of modalities, which were classified as knobs and buttons on the centre stack, touch screen, steering wheel buttons, and voice control. Knobs and buttons were found to be most suitable for basic tasks whilst the other modalities could be better for more advanced tasks.

Furthermore, [31] performed a study investigating the frequencies of secondary tasks in driving. They found that passenger-related distraction is a very big part of the total amount of distraction during driving. The driver was distracted in some form during 25% to 40% of the driving time. The most frequently used secondary tasks are telephoning, internal human-machine interface inputs and using the mobile phone.

d. Knowledge gaps

Most studies in the current literature are focused on the impact of distraction on different performance measures. The impact of distraction on driving performance is therefore well studied. There is also a lot of research about innovative HMI technologies and how they might be used to reduce driver distraction. To the author’s knowledge, no studies have looked at how often different HMI tasks are performed during real-world driving using the internal HMI of a car, which is in line with the findings by [32], who also found that there is a lack of research on the prevalence of different types of distraction. Only two studies were found that looked into the frequency with which different tasks occur, but these studies did not make a distinction between different internal HMI tasks [7] [31]. This poses problems for the goals set by RDW and Euro NCAP [9]. It is not clear which tasks are most important to assess, how often different tasks are performed during driving and what factors could influence this. In particular, it could be interesting to look at the influence of car familiarity as car-sharing services are becoming more common.

When looking at the literature, it can also be observed that some studies used older driving simulators. Over time, driving simulators have become increasingly realistic and can maybe produce more ecologically valid results since it is closer to real life. Therefore, it is important to keep studying the effects of interacting with HMIs in these newer driving simulators, particularly for scenarios that are difficult to find in real life or scenarios that are viewed as too unsafe for real-world testing.

The literature furthermore indicates that repeated use of an HMI enhances driving performance during HMI tasks. However, this research field is still lacking, as only two studies were found to investigate this with a limited set of performance measures [19] [24]. Since the literature makes clear that no single driving performance measure can capture all the effects of distraction [26], it is not sufficient to use a single indicator for evaluating the safety of HMIs.

Geographical areas could also influence the results of the different studies mentioned. People in different areas of the world might have different levels of familiarity with technological systems. These kinds of geographical differences could change the results of the individual studies mentioned significantly. No studies conducted in the Netherlands were found that examined the impact of performing various HMI tasks on driving performance.

This thesis aims to contribute to the knowledge gap regarding the frequencies with which different internal HMI tasks are performed during driving in passenger cars. This gap has been chosen due to its immediate practical relevance. The frequencies can be used to create regulations and also to create an assessment methodology. The results of this kind of research could therefore aid Euro NCAP and RDW in reaching their goals [9]. This thesis also looks into the factors that might influence the frequency with which tasks are performed. As car sharing is becoming more common [8], it is particularly interesting to look at the impact of car familiarity. Studying the impact of different factors improves the interpretation of the results and could be input for creating measures to reduce the frequency of HMI use

III. Methodology

a. Experimental design

In a 2x1 within-subjects design, the number of times that drivers performed specific tasks during a naturalistic driving experiment was counted by driving along. To be able to quickly observe and count, an observation checklist was created that structured the different types of tasks intuitively into categories. The participants had to drive two trips: 1) In their own car and 2) in an unfamiliar car. About half of the participants started by driving their own car while the other half started with the unfamiliar car. The initial settings of the unfamiliar car were the same for all participants. A convenience sample was used for this study, which should suffice, due to the explorative nature of the study. Ethical approval was granted on November 16th 2023 by the Human Research Ethics Committee (HREC) of the TU Delft with application number 3679.

Only familiar routes were driven during the trips since these are the routes that are mostly driven by people and therefore most representable for the average trip. Being familiar with a route means that participants perform their user interface tasks as they usually do. Using unfamiliar routes may introduce different results. A study by [33] found that the familiarity of a route has a big impact on driving behaviour. One of the findings in this study is that increased familiarity resulted in lower task difficulty and increased engagement in secondary tasks. Since unfamiliar routes require more attention, fewer user interface tasks are likely performed compared to familiar routes. The route familiarity was self-reported by the participants on a scale from 1 to 10. A route is considered familiar when the participants score the familiarity higher than or equal to 8, which is a similar approach to [34] and [35]. The route length for each trip was less than 30 minutes to reach a maximum of 60 minutes per participant.

b. Procedure

For the observations. the unfamiliar car was driven to the starting location by the researcher, which was usually the home address of the participants. For most participants, the two trips were driven during the same appointment. This was done to minimize weather and context differences. However, due to time constraints, 7 out of the 30 participants had to drive the two trips on separate days.

Before the trips, the participants were asked about their age, gender and driving experience. They were also asked to come up with a route that was familiar to them. They were told that the total time per trip would have to be less than 30 minutes but more than 15 minutes. The participants would then drive from the origin to the destination and back again in both cars. During the trips, the researcher was sitting in the passenger seat noting down every time that a task was performed. After driving both trips, the participants were asked a few more questions. Also, their familiarity with both cars was measured. After the questions, the characteristics of the participant's car were noted. It was noted whether or not the HMI of the car was able to perform the tasks included in the study and whether or not the car was equipped with adaptive climate control, adaptive cruise control, a touch screen, automatic lights and an automatic transmission.

The choice was also made to not engage in any conversations during the trip, as most trips with a car are made without passengers [36]. Engaging in conversations can distract the driver and can result in the driver performing fewer secondary tasks than they would normally perform [31]. The observations were performed from the 29th of November 2023 until the 8th of January 2024 and a small pilot was performed to test the procedures.

c. Materials

The unfamiliar car for the study was provided by the researcher. 28 out of the 30 participants drove a Seat Toledo from 2014 as their unfamiliar car. The climate control temperature of this car was set to 19 degrees and the radio was set to radio 538 on volume 8. Also, all lights and the windshield wipers were turned off such that all participants started with the same settings. The seat toledo had a manual transmission and did not include adaptive cruise control, automatic lights or a touchscreen but did include adaptive climate control. Two participants drove a different car as their unfamiliar car. For one of the participants, this was due to time constraints. The other participant who did not drive the Seat was only able to drive in cars with an automatic transmission, which the Seat did not have. The other two cars that were used as unfamiliar cars are a Renault Megane from 2012 and a Renault Modus from 2005. For the familiar car trips, people used their personal vehicles. When looking at the car characteristics of this fleet, a good variety of different types of cars can be found in the sample. Information about the car fleet can be found in Table 1.

Table 1: Car fleet

Car fleet characteristics of participants				
Variable	**Category**	**Amount**	**%**	**N**
Adapt. climate control	Yes	19	63.33	30
	No	11	36.67	
Adapt. cruise control	Yes	11	36.67	30
	No	19	63.33	
Auto lights	Yes	21	70	30
	No	9	30	
Transmission	Manual	17	56.67	30
	Automatic	13	43.33	
Touch screen	Yes	20	66.67	30
	No	10	33.33	

Furthermore, an observation checklist was used during the observations, which included all tasks intuitively structured into different categories. Only tactical tasks as mentioned by [37] have been considered for this study. In this study, the definition of tactical tasks is the following: Tasks that are not performed fully automatically by the driver and take some seconds to perform. This resulted in tasks such as adjusting radio volume or changing the climate control temperature. In the end, the following 11 categories were distinguished: Radio and media, Climate control, Calling, Lights, Adjusting car setup, Windshield, Danger signalling, Cruise control, Adjusting settings, Using extra features and Other. The observation checklist also included all other factors that were documented. By setting lines in the correct boxes for each task, it was possible to count the number of occurrences per road type and situation effectively for both trips.

d. Measures

The distinguished road types for this study include city roads, rural roads and highways, which is a similar approach to [31]. The two situations that were included are standing still and driving in a traffic jam. Based on the counts for the separate tasks and the trip durations, the individual task frequencies and total frequencies per category have been calculated. To be able to calculate the frequencies per road type, the driving times for each of these have been documented. The total frequency including all tasks has also been calculated.

The counting was performed based on a cooldown timer. An educated guess was made on the number of seconds this should be set to, which was decided to be 5 seconds. In this research, the assumption was made that when a driver performs the same task again within five seconds or continues with a task within five seconds, it is still part of the first task occurrence. The task is seen as being continued and not as a new task occurrence. Turning something on is seen as a separate task from turning something off. The counting began when the driver moved the vehicle for the first time. Only tasks that are part of the original internal HMI have been looked at during the observations.

The factors that were documented for this research based on their possible relation with the task frequencies are the following: Age, Gender, Time with a driving license, Frequency of driving, Outside temperature, Weather type, Peak hour, Day of the week and car familiarity.

For the socio-demographic variables, the decision was made to include age, gender and driving experience. These variables are often included in literature when performing data analysis. [35] and [31] also included the participants' age and driving experience as variables in their study regarding HMI tasks. The expectation was that age has a significant effect on the frequency with which people perform tasks. Older people are often less tech-savvy and therefore struggle more to perform secondary tasks (21). The expectation was therefore that older people show lower frequencies of performing tasks. Gender was not expected to have a significant correlation with the frequency. [38] did find that women tend to display worse driving performance when distracted in comparison to men, which could result in men having higher frequencies, but this is questionable. It was expected that the number of years with a driving license would have a positive correlation with the frequencies since [39] found that novice drivers engaged more frequently in distracting tasks as they gained more driving experience. Also, [33] found that lower task difficulty has a relation with increased engagement in secondary tasks. Since more experienced drivers have less difficulty with the driving task, it was expected that more experienced drivers have higher frequencies with which they perform tasks.

Four contextual factors have been identified as relevant variables for this study. These are outside temperature, weather type, peak hour and day of the week. The weather type and outside temperature are included since they could have a big influence on the use of climate control options and windshield wipers. Different weather types were expected to have different correlations with the frequencies. The hypothesis was that rainy weather increases the amount of windshield tasks that are being used. Lastly, it was noted whether the trip was driven during peak hours or not and on what day of the week. The inspiration for this variable was found in a study by [40], who found that traffic density had an impact on secondary task engagement.

To measure car familiarity, the participants had to state whether they agreed or disagreed with different statements on a scale of 1-10. One statement attempted to measure overall car familiarity while 4 other statements tried to measure car familiarity from different angles. This variable was included since the literature suggests that familiarity can have a big impact on driving behaviour [33]. Also, similar scales for measuring familiarity have been used in the past [34].

The participants were also asked if they could describe what the researcher was looking at. This was done to find out whether the participants already had an idea of what the study was about. Then, they were asked if they felt that the researcher had influenced their behaviour during the trip. Most people could not correctly describe what the researcher was observing while they drove their routes and did not feel like their behaviour was influenced.

e. Participants

Most participants who took part in the study were friends and family of the researcher. In the end, 30 participants were recruited. To be eligible for the study, participants were required to be above the age of 18 and had to own a driver's license. Also, as far as this was possible, participants were selected in a way that a variety of different ages and genders participated. All 30 participants successfully completed the experiment. The information about the participants can be found in Table 2.

Table 2: Socio-demographic variables

Numerical sociodemographic variables				
Variable	**Avg**	**Min**	**Max**	**N**
Age (years)	41.1	20	85	30
Driv. Lic. (years)	22.0	3	65	30
Driv. freq (per week)	7.9	1	28	30
Route familiarity	9.5	8	10	30
Categorical sociodemographic variables				
Variable	**Categ.**	**#**	**%**	**N**
Gender	Male	20	66.7	30
	Female	10	33.3	
Q1: Influenced?	Yes	4	13.3	30
	No	26	86.7	
Q2: purpose?	Yes	1	3.3	30
	No	29	96.7	

f. Data analysis methods

To test if the differences in the average frequencies of the different tasks that were performed are statistically significant, repeated measure ANOVA's were performed. Paired t-tests are performed to test for differences in average frequencies between the two trips that are driven by the participants.

To test if the distribution of tasks that are performed while driving versus standing still is the same for different categories a chi-square test was performed. Tasks that are more typically performed while standing still could be seen as less dangerous than tasks that are typically performed while driving. To meet the requirements of a chi-square test, only tasks or task categories that have enough observations are tested for differences in distribution.

To find if overall car familiarity is an underlying factor of different indicators, a factor analysis was performed. If the statements had a high load on the same factor, they were used to create an overall familiarity scale. A factor analysis is very useful to find and measure underlying factors, which are also called latent variables [41]. A standard approach for factor analysis has been used. Statements that load lower than 0.3 on the factor have been deleted for the overall familiarity scale. The overall scores on the familiarity scale have been calculated with a sum of scores.

The dependent variable in this study is the number of times that a task has been performed adjusted by the driving time for each participant. This type of data can be seen as count data, which is why Poisson regressions have been performed to investigate which factors influence the frequencies. The natural log of the trip durations has been used as the offset variable to take into account the fact that the trip durations varied across participants. The formula for the Poisson regressions in which x_i stands for independent variable i is the following: $\ln(\text{number of times}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \ldots. \beta_i x_i + \ln(\text{trip duration})$

Multiple models have been constructed with two different types of dependent variables. These are the total number of tasks that have been performed per hour and the total number of tasks per category that have been performed per hour. The Poisson models have been constructed using the data from the trips where the participants drove their own car. The reason for this is that the unfamiliar car data has almost no variety in car types, as 28 out of the 30 participants drove the same unfamiliar car. A specific method has been followed to get to the final Poisson models. First, separate models have been made for all factors where only one factor at a time was entered per dependent variable. The most interpretable factors with the highest significance values were then combined in one model. The experimenting stopped when a final model was constructed that was interpretable whilst also having (near) significant variables.

IV. Results

a. Most frequently used tasks

Figure 1 shows the average frequency of the 10 most often performed tasks during the 60 trips, which includes both the trips with the familiar and unfamiliar cars.

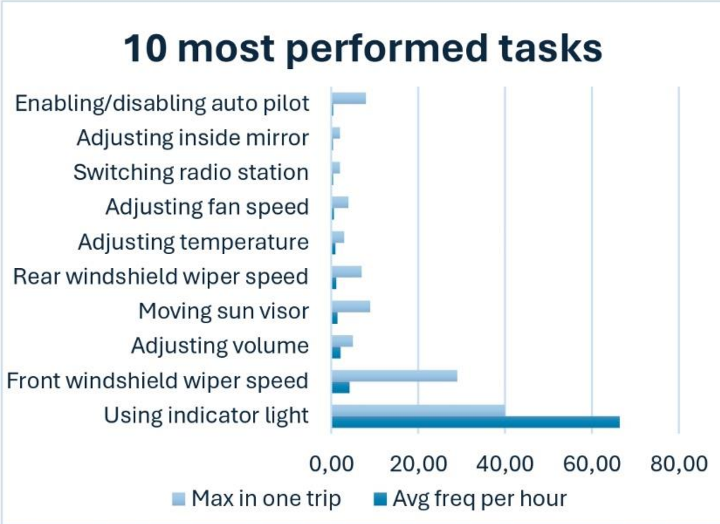


Figure 1: Most performed tasks

The most used task during the 60 trips was found to be the indicator light, followed by adjusting the front windshield wipers' speed and adjusting the volume. On average, the frequency of using the indicator light is found to be 66.4 per hour. The front windshield wipers' speed was adjusted with a frequency of 4.22 times per hour whilst the volume was adjusted with a frequency of 2.15 times per hour. The windshield wiper task is an interesting one, as it displays a relatively high maximum compared to some of the other tasks. The reason for this is that some participants were unable to figure out how to set the windshield wipers to an interval. Therefore, some participants had to constantly trigger the windshield wipers themselves, resulting in a maximum of 29 occurrences on a single trip for this task.

To test if the differences in average frequency for the different tasks are statistically significant, a repeated measures ANOVA was performed, which indicates that the average frequency of the indicator light is statistically different from the average frequency of all other tasks. The average frequency of the volume task is statistically different from the average of switching radio stations and adjusting the inside mirror. All the other tasks only have statistically significantly different average frequencies from using the indicator light and/or adjusting the volume.

To get a better overview of the types of tasks that are used relatively often, the tasks have been divided into several categories. In Figure 2, the average total frequency for tasks in different categories is shown. Only the categories that have more than 10 total tasks performed are presented. Also, the total average frequency including all tasks is shown. Since the indicator light is used so much more than the other tasks, also the total number of tasks performed over all trips excluding the indicator light is presented. Lastly, a variable was created that excludes the front windshield wiper, since this variable was also an outlier.

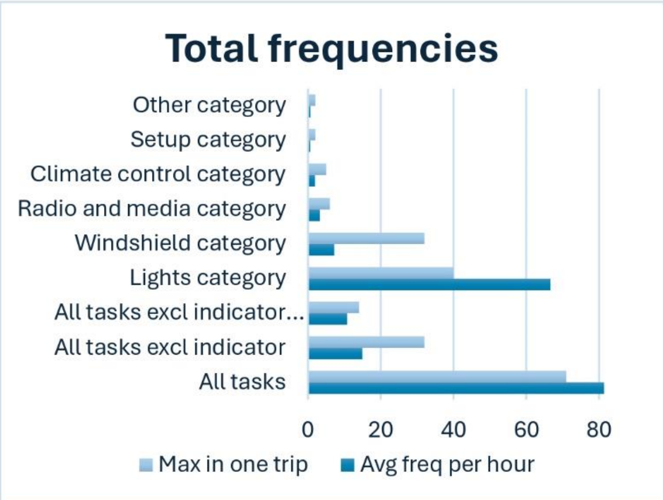


Figure 2: Most performed tasks in categories

On average, participants performed 81.35 tasks per hour of driving. This number is reduced significantly when the indicator light is excluded, as it then drops to 14.95 tasks per hour of driving. Since the indicator light is part of the lights category, it is not surprising that this category has the highest total average frequency of 66.63 tasks per hour. The categories with the second and third highest frequencies are the windshield and radio and media categories with an average of 7.22 and 3.26 tasks per hour respectively. The categories with the lowest total average frequency per trip are calling, changing settings and danger signalling.

The average frequency of the lights category is statistically different from all other categories. The average frequency of the windshield task category is only statistically different from the lights category, the setup category and the “other” category. The average frequency of the radio and media category is also different from these categories. The average frequencies of the other categories in Figure 2 are only statistically significantly different from one or more of these three categories.

b. Factor analysis

To test if the four car familiarity statements in the study measure the same underlying factor, a factor analysis has been performed. The hypothesis is that the four statements together measure the underlying factor of overall car familiarity.

Table 3: Factor analysis

Perceived HMI car familiarity (α = 0.846)	**Factor loading**
I know the features/options that the car has available	0.945
I know where the buttons are located	0.891
I understand the dashboard and the things it displays	0.649

From the factor analysis, it becomes clear that the three statements in Table 3 measure the same underlying factor. This factor has been called the perceived HMI car familiarity, as all the statements are about how familiar the participants perceive the HMI of the car. This new perceived HMI car familiarity variable has been created with a sum of scores. To see if this was possible, the scale's reliability was checked. The scale was found to be reliable (>0.70) with a Cronbach’s alpha of 0.846. The sum of scores has the advantage that it can be interpreted in terms of the original scale. The other statement in the study: “I drive this car a lot” was found to measure a different factor and was deleted in the factor analysis. The perceived HMI car familiarity factor is later used in the Poisson regressions.

c. Familiar vs Unfamiliar

Before looking at the differences between the familiar and unfamiliar car trips, it is important to test if the average familiarity is indeed statistically significantly different across the two trips. Therefore paired t-tests

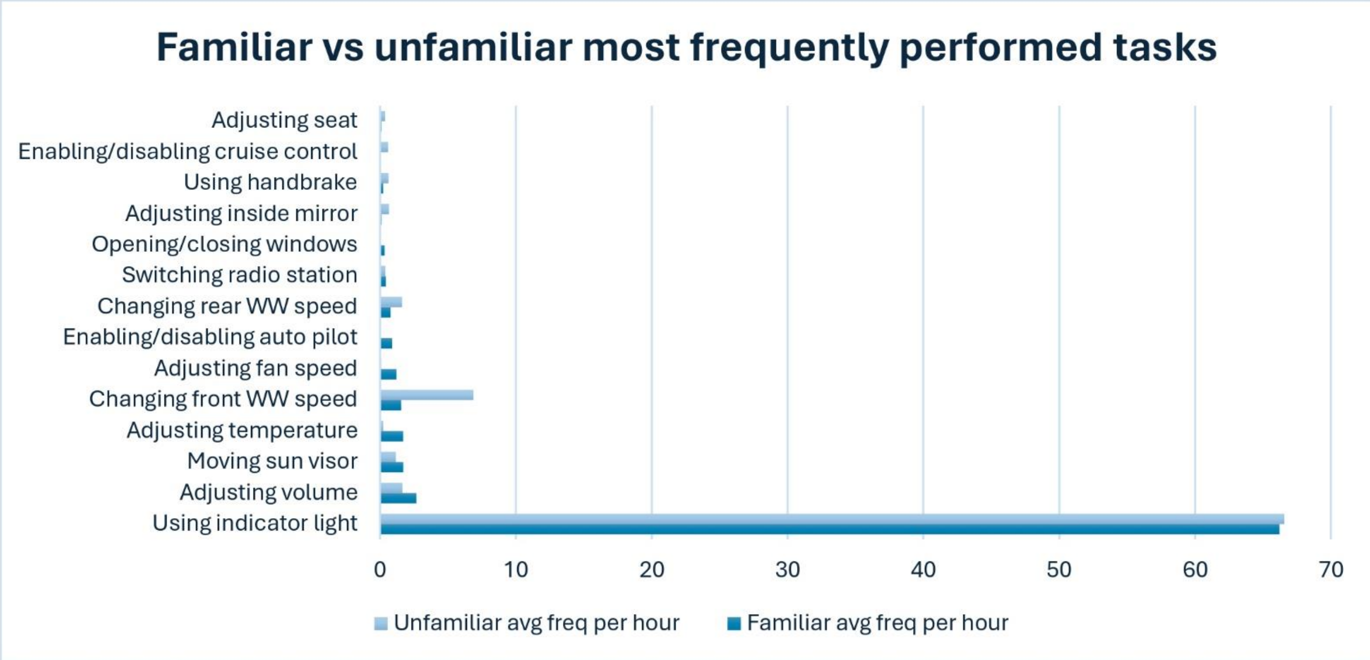


Figure 3: Most performed tasks familiar vs unfamiliar

have been performed for the average scores on each statement and the perceived HMI car familiarity. The paired t-tests show for each statement that the personal vehicles were indeed more familiar on average than the car that was brought by the researcher ($p<0.001$).

The most performed task in both trips was found to be changing the indicator light direction. This task was performed with an average frequency of 66.22 and 66.58 respectively. For the familiar car, the next most used tasks were found to be changing the volume, moving the sun visor, changing the temperature of the climate control system and changing the front windshield wiper speed. For the unfamiliar car, the most performed tasks after changing the indicator light direction are found to be changing the front windshield wiper speed, adjusting the volume, changing the speed of the rear windshield wiper and moving the sun visor. For the unfamiliar car, higher frequencies can be found for tasks that are related to setting up the car. Examples of these tasks are adjusting the inside mirror and setting up the seat. In Figure 3, the average frequencies of the 10 most performed tasks for both trips are compared and presented whilst Figure 4 presents this for the different categories.

For the familiar car, a total average frequency of 79.89 tasks per hour was found. Excluding the indicator light from the task set reduces this number to 13.67 tasks per hour. Excluding also the front windshield wiper task reduces it even further to 12.11 tasks per hour. Looking at the unfamiliar car, the total average frequency is found to be 82.82 tasks per hour. Excluding the indicator light reduces this average to 16.27 tasks per hour and excluding also the front windshield wiper task reduces this further to 9.36 tasks per hour. For both trips, the three categories with the highest average task frequencies are the lights, windshield and radio and media categories. For the familiar car, the next category with the highest task frequencies is found to be the climate control category whilst for the unfamiliar car this is found to be the setup category.

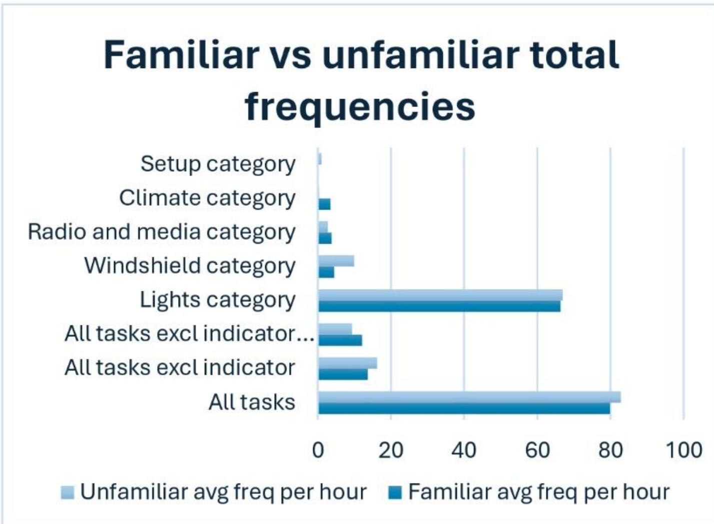


Figure 4: total average frequencies familiar vs unfamiliar

Table 4: Poisson regressions

	Total tasks Model 1 (N=30)		Radio and media Model 2 (N=30)		Windshield Model 3 (N=30)		Climate control Model 4 (N=30)	
	Exp(B)	P-value	Exp(B)	P-value	Exp(B)	P-value	Exp(B)	P-value
Perceived HMI familiarity	1.276	0.009	1.436	0.048				
Age			0.980	0.060				
Gender			3.239	0.016	0.534	0.034	0.576	0.091
Rainy					8.725	<0.001		
Sunny					9.167	<0.001		
Transmission type					0.372	0.004		
Adaptive cruise control							0.355	0.020

When using an alpha of 0.05, two categories display average frequencies that are statistically significantly different between the familiar and unfamiliar car trips. These categories are the climate control category ($p<0.001$) and the setup category ($p=0.03$). The climate control category has an average frequency of 3.45 tasks per hour when people were driving their personal cars versus 0.4 tasks per hour when driving the unfamiliar car. The setup category has an average frequency of 0.21 tasks per hour when people were driving their personal cars versus 1.03 tasks per hour when driving an unfamiliar car. A variable for which the difference in average frequencies is nearly statistically significant is the variable that looks at the total tasks performed excluding the indicator lights and front windshield wiper tasks. This test resulted in a p-value of 0.07. Since the front windshield wiper was such an outlier due to the unintuitive design of the Seat Toledo, this would indicate that people in general perform more tasks in familiar cars.

d. Influence of the factors on the task frequencies

To study the influence of the factors in the study, Poisson regression models, as presented in Table 4, have been constructed using the familiar car data. It is important to note that all models showed (some) overdispersion (value/df>1) but were statistically significantly better than the intercept-only model.

The results of the Poisson model for the total tasks indicate that people seem to overall perform 28% more tasks for every point higher on the perceived HMI car familiarity scale.

The results of the Poisson model for the radio and media category indicate that Male drivers seem to perform 224% more radio and media tasks than female drivers. Furthermore, for every year that a person is older 2% fewer radio and media tasks are performed. Also, the coefficient for perceived HMI car familiarity is significant and indicates that people perform 44% more radio and media tasks when they scored familiarity one point higher.

The results of the Poisson model for the windshield category indicate that people seem to perform 773% more windshield tasks in rainy conditions and 817% more windshield tasks in sunny conditions. In rainy conditions, this is due to the windshield wipers and in sunny conditions due to the sun visor being used more often. Furthermore, when people drive in cars with an automatic transmission they perform 63% less windshield tasks. This could be because cars with an automatic transmission more often have automatic windshield wipers. This variable was not included in this study so this correlation cannot be tested. Also, the regression indicates that men perform 47% less windshield tasks.

The results of the Poisson model for the climate control category indicate that the coefficient for gender is significant only at the 10% level. It is included since it is interpretable. The data suggests that men use 42% less climate control tasks than women. This could be because women are more perceptible to cold [42]. The data also suggests that people use 65% fewer tasks in cars with adaptive cruise control. This could be because modern cars perform more tasks automatically.

Discussion

The results of this study revealed the most frequently performed tasks using car HMIs and found that these frequencies differ when people drive their own car compared to an unfamiliar car. When looking at all tasks combined excluding the front windshield wiper and indicator light tasks, this study indicates that people perform more tasks in their own car. This is likely because participants are more comfortable in familiar cars and therefore more confident to perform these tasks safely while driving. This can be seen as a form of self-regulation where people only perform tasks when they think they can safely perform them [43].

Unfamiliar cars also seem to result in increased distraction when the design of the car is unintuitive. Multiple participants struggled to set the front windshield wipers to an interval due to the unfamiliar design of the unfamiliar car. This shows that a lack of standardization can result in extra distractions. The importance of standardization has been mentioned in other research as well [4] [5]. The findings are also in line with other research on the impact of unfamiliar cars, which shows that unfamiliar cars reduce driving performance [6] [19] [44]. The results of this study also seem to fit well within the overall literature on HMI distraction, as they are in line with the findings by [7], who also found that climate control and radio and media tasks are among the most prevalent tasks.

Limitations of this study include the fact that car familiarity was measured based on statements, that the Seat Toledo that was used lacks some modern features and that a relatively small convenience sample was used. Also, the counting was performed manually, such that it is more prone to mistakes. Lastly, some people found it difficult to follow the instructions of no conversations. Further research could remedy some of these problems by using cameras, tracking people for longer periods and using a bigger sample. This study can also be extended by performing observations under different external circumstances and in different countries to see if there are any differences. Further research could also focus on the impact of factors that were not included in this study, such as psychological factors. These could be factors such as people's perceived aggressiveness during driving. Also, questions about driving styles could be asked. Further research could also focus on the knowledge gaps in chapter 2.5 that are not addressed in this study. This includes conducting studies on the effect of performing HMI tasks on driving performance using newer simulators, conducting a study on the effect of performing HMI tasks on driving performance in the Netherlands and conducting more studies on the effect of repetition with an HMI on driving performance using different performance measures.

Conclusion

Euro NCAP and RDW have set the target to encourage regulations in the field of passenger car HMIs [9]. This could help to reduce the many traffic fatalities that are caused by distraction [32] [45]. However, knowledge was missing on the frequencies with which different HMI tasks are performed. Apart from studies by [7] and [31], most studies in the current literature are focused on the impact of distraction on different performance measures and innovative technologies to reduce distraction. This posed a problem for the development of the assessment methodology that RDW and Euro NCAP are working on.

By using the method that was proposed in this study, it was possible to measure the prevalence of different internal HMI tasks. This thesis contributes to the literature by identifying the frequencies with which different tasks are performed and identifying which factors could influence this. It thereby increases the understanding of internal HMI use in cars. This study also provides a methodology that can be used by future studies on the topic of distraction prevalence. The findings of this study indicate that the most used tasks are the indicator light, tasks in the windshield category and tasks in the radio and media and climate control categories. The differences in frequencies between different tasks are relatively large and even with such a small sample size, statistically significant differences were found between the top 10 most used tasks. In general, people seem to perform around 12 tasks per hour in familiar cars versus 9 tasks per hour in unfamiliar cars when excluding the front windshield wiper and indicator light tasks. This shows that there are a significant amount of distraction moments per hour due to internal HMI use. Different factors seem to impact the frequencies of different types of tasks. The factors that had an influence included car familiarity, gender, age and weather conditions. The type of car also seems to impact the task frequencies, but more research on this is needed. An incidental finding of this study is that unfamiliar cars can result in an increased number of tasks performed when the design of the car is unintuitive and can therefore lead to increased distraction. Future research on the safety impact of performing internal HMI tasks can use the results of this study to better understand the relative risks of the different tasks in real-world driving.

References

[1] ViBilägare. Physical buttons outperform touchscreens in new cars, test finds. (2022). https://www.vibilagare.se/english/physical-buttons-outperform-touchscreens-new-cars-test-finds

[2] Blanco, M., Biever, W., Gallagher, J., & Dingus, T. A. (2006). The impact of secondary task cognitive processing demand on driving performance. Accident Analysis & Prevention, 38(5), 895–906. https://doi.org/10.1016/j.aap.2006.02.015

[3] Davidse, R.J., Doumen, M.J.A., Duijvenvoorde, K. van & Louwerse, W.J.R. (2011). Bermongevallen in Zeeland: karakteristieken en oplossingsrichtingen. R-2011-20. SWOV, Leidschendam.

[4] Uhlving, V. M., Weyde, K. V., Hagenzieker, M., Aasvik, O., & Rostoft, M. T. (2023). Advanced driver support systems and screen interfaces in car - How do they affect driver behaviour? TØI, 978-82-480-1519–2. https://www.toi.no/getfile.php?mmfileid=77085

[5] Ružić, D. (2022). The ergonomic aspects of a passenger car climate control panel. IOP Conference Series: Materials Science and Engineering, 1271(1), 012019. https://doi.org/10.1088/1757-899x/1271/1/012019

[6] Lee, S. E., Dingus, T. A., Klauer, S. G., Neale, V. L., & Sudweeks, J. (2005). Naturalistic data collection of driver performance in familiar and unfamiliar vehicles. Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting, 49(22), 1994–1998. https://doi.org/10.1177/154193120504902223

[7] Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Pérez, M. Á., Buchanan-King, M., & Hankey, J. M. (2016). Driver Crash Risk Factors and Prevalence Evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(10), 2636–2641. https://doi.org/10.1073/pnas.1513271113

[8] Autodelen. (2022, 16 maart). Cijfers en onderzoek. Geraadpleegd op 9 oktober 2022, van https://autodelen.info/cijfers-autodelen

[9] Euro NCAP. Euro NCAP Vision 2030: a Safer Future for Mobility. (2022). https://www.euroncap.com/en/press-media/press-releases/euro-ncap-vision-2030-a-safer-future-for-mobility/

[10] Euro NCAP. How To Read The Stars. (2023). https://www.euroncap.com/en/about-euro-ncap/how-to-read-the-stars/

[11] Schmidt, A., Dey, A. K., Kun, A. L., & Spiessl, W. (2010). Automotive user interfaces. https://doi.org/10.1145/1753846.1753949

[12] Prabhakar, G., & Biswas, P. (2021). A Brief Survey on Interactive Automotive UI. Transportation engineering, 6, 100089. https://doi.org/10.1016/j.treng.2021.100089

[13] Lo, V. S. Y., & Green, P. E. (2013). Development and Evaluation of Automotive Speech Interfaces: Useful Information from the Human Factors and the Related Literature. International Journal of Vehicular Technology, 2013, 1–13. https://doi.org/10.1155/2013/924170

[14] Ma, J., Li, J., & Gong, Z. (2022). Evaluation of driver distraction from in-vehicle information Systems: A simulator study of interaction modes and secondary tasks classes on eight production cars. International Journal of Industrial Ergonomics, 92, 103380. https://doi.org/10.1016/j.ergon.2022.103380

[15] Iqbal, S. T., E. Horvitz, Y.-C. Ju, and E. Mathews (2011). “Hang on a sec!: effects of proactive mediation of phone conversations while driving”. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM. 463–472.

[16] Breitschaft, S. J., Pastukhov, A., & Carbon, C. (2021). Where’s My Button? Evaluating the User Experience of Surface Haptics in Featureless Automotive User Interfaces. IEEE Transactions on Haptics, 15(2), 292–303. https://doi.org/10.1109/toh.2021.3131058

[17] Green, P. (1998). Visual and Task Demands of Driver Information Systems (Technical ReportUMTRI-98-16), Ann Arbor, MI: The University of Michigan Transportation Research Institute(in preparation).

[18] Green, P. (1999). The 15-second rule for driver information systems. Proceedings of the ITS America Ninth Annual Meeting. Washington, D.C.: Intelligent Transportation Society of America

[19] Chisholm, S., Caird, J. K., & J, L. (2008). The effects of practice with MP3 players on driving performance. Accident Analysis & Prevention, 40(2), 704–713. https://doi.org/10.1016/j.aap.2007.09.014

[20] Ou, Y., Liu, Y. C., & Shih, F. (2013). Risk prediction model for drivers’ in-vehicle activities – Application of task analysis and back-propagation neural network. Transportation Research Part F-traffic Psychology and Behaviour, 18, 83–93. https://doi.org/10.1016/j.trf.2012.12.013

[21] Cooper, J. D., Wheatley, C. L., McCarty, M. M., Motzkus, C. J., Lopes, C. L., Erickson, G. G., Baucom, B. R., Horrey, W. J., & Strayer, D. L. (2019). Age-Related Differences in the Cognitive, Visual, and Temporal Demands of In-Vehicle Information Systems. Frontiers in Psychology, 11. https://doi.org/10.3389/fpsyg.2020.01154

[22] Strayer, D. L., Cooper, J. D., Goethe, R. M., McCarty, M. M., Getty, D., & Biondi, F. (2017). Visual and Cognitive Demands of Using In-Vehicle Infotainment Systems. Foundation for Traffic Safety, 202-638–5944. https://aaafoundation.org/wp-content/uploads/2017/11/VisualandCognitive.pdf

[23] Kun, A. L. (2018). Human-Machine Interaction for Vehicles: Review and Outlook. Foundations and Trends in Human-computer Interaction, 11(4), 201–293. https://doi.org/10.1561/1100000069

[24] Broström, R., Aust, M. L., Wahlberg, L., & Källgren, L. (2013). What drives off-road glance durations during multitasking: capacity, practice or strategy? Proc. of the Third International Conference on Driver Distraction and Inattention. 04-06.09.2013 Göteborg, paper 08-P. https://trid.trb.org/view/1405238

[25] Wang, Y., Reimer, B., Mehler, B., Zhang, J., Mehler, A., & Coughlin, J. F. (2010). The impact of repeated cognitive tasks on driving performance and visual attention. In Advances in human factors and ergonomics series (pp. 668–678). https://doi.org/10.1201/ebk1439834992-70

[26] Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. Transportation research procedia, 25, 1796–1805. https://doi.org/10.1016/j.trpro.2017.05.148

[27] Redelmeier, D. A., & Tibshirani, R. (1997). Association between Cellular-Telephone calls and motor vehicle collisions. The New England Journal of Medicine, 336(7), 453–458. https://doi.org/10.1056/nejm199702133360701

[28] Horberry, T., Anderson, J., Regan, M., Triggs, T. J., & Brown, J. M. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. Accident Analysis & Prevention, 38(1), 185–191. https://doi.org/10.1016/j.aap.2005.09.007

[29] Dingus, T. A., Owens, J. M., Guo, F., Fang, Y., Pérez, M. Á., McClafferty, J., Buchanan-King, M., & Fitch, G. M. (2019). The prevalence of and crash risk associated with primarily cognitive secondary tasks. Safety Science, 119, 98–105. https://doi.org/10.1016/j.ssci.2019.01.005

[30] Perez, M., Owens, J., Viita, D., Angell, A., Ranney, T. A., Baldwin, G. H. S., Parmer, E., Martin, J., Garrott, W. R., & Mazzae, E. N. (2013, May). Radio tuning effects on visual and driving performance – Simulator and test track studies. (Report No. DOT HS 811 781). Washington, DC: National Highway Traffic Safety Administration.

[31] Metz, B., Landau, A., & Just, M. (2014). Frequency of secondary tasks in driving – results from naturalistic driving data. Safety Science, 68, 195–203. https://doi.org/10.1016/j.ssci.2014.04.002

[32] Stelling, A., & Hagenzieker, M. P. (2015). Schatting aantal verkeersdoden door afleiding. In SWOV (R-2015-13). SWOV. https://swov.nl/system/files/publication-downloads/r-2015-13.pdf

[33] Harms, I. M., Burdett, B. R., & Charlton, S. G. (2021). The Role of Route Familiarity in Traffic Participants' Behaviour and Transport Psychology Research: A Systematic review. Transportation research interdisciplinary perspectives, 9, 100331. https://doi.org/10.1016/j.trip.2021.100331

[34] Harms, I. M., van Dijken, J. H., Brookhuis, K. A., & de Waard, D. (2019). Walking without awareness. Frontiers in Psychology, 10:1846, 1-12. doi:10.3389/ fpsyg.2019.01846

[35] Burdett, B. R., Charlton, S. G., & Starkey, N. J. (2019). Mind wandering during everyday driving: an on-road study. Accident Analysis & Prevention, 122, 76–84. https://doi.org/10.1016/j.aap.2018.10.001

[36] CBS. (2022). Mobiliteitstrend; per rit, vervoerwijzen, reismotief, leeftijd en geslacht. https://opendata.cbs.nl/#/CBS/nl/dataset/84755NED/table?ts=1697116530423

[37] Michon, J. A. (1985). A critical view of driver behavior models: What do we know, what should we do? In Springer eBooks (pp. 485–524). https://doi.org/10.1007/978-1-4613-2173-6_19

[38] Arca, A. A., Chouljian, E., & Mouloua, M. (2022). The role of Gender Differences in Distracted Driving Behavior: A Psychophysiological approach. Proceedings of the Human Factors and Ergonomics Society . . . Annual Meeting, 66(1), 1280–1284. https://doi.org/10.1177/1071181322661453

[39] Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. The New England Journal of Medicine, 370(1), 54–59. https://doi.org/10.1056/nejmsa1204142

[40] Cuentas-Hernandez, S., Li, Y., King, M., & Oviedo-Trespalacios, Ó. (2023). The impact of road traffic context on secondary task engagement while driving. Frontiers in Psychology, 14. https://doi.org/10.3389/fpsyg.2023.1139373

[41] Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. Tutorials in Quantitative Methods for Psychology, 9(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p07

[42] Kaikaew, K., Van Den Beukel, J. C., Neggers, S. J. C. M. M., Themmen, A. P. N., Visser, J. A., & Grefhorst, A. (2018). Sex difference in cold perception and shivering onset upon gradual cold exposure. Journal Of Thermal Biology, 77, 137–144. https://doi.org/10.1016/j.jtherbio.2018.08.016

[43] Wandtner, B., Schumacher, M., & Schmidt, E. (2016). The role of self-regulation in the context of driver distraction: A simulator study. Traffic Injury Prevention, 17(5), 472–479. https://doi.org/10.1080/15389588.2015.1102231

[44] Liu, Y., & Hansen, J. H. L. (2019). Analysis of Driving Performance Based on Driver Experience and Vehicle Familiarity: A UTDrive/Mobile-UTDrive App Study. SAE International Journal Of Transportation Safety, 7(2). https://doi.org/10.4271/09-07-02-0010

[45] U.S. Department of Transportation. (2023). Distracted driving in 2021 (DOT HS 813 443). NHTSA. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813443

Appendix B: Operationalization of variables

Table 21: variable names in the database for the sociodemographic variables

Age (years)**:** Age	**Gender** (Male/female/other)**:** Gender
Driving license (years)**:** Driv_lic	**Driving frequency** (Times per week last year)**:** Driv_fre
Route familiarity (1-10)**:** R_famil	**I am familiar with this car** (1-10): C_famil

I drive this car a lot (1-10): C_drive
I know the features/options that the car has available (1-10): C_feature
I know where the buttons are located (1-10): C_button
I understand the dashboard and the things it displays (1-10): C_dashb

Table 22: variable names in the database for other variables and dummy coding

Route characteristics	
Highway road (time)	HW_time
Rural road (time)	RU_time
City road (time)	CI_time
Traffic jam (time)	JA_time
Board computer system type	
No touch screen (0)	BCS_type
Touch screen (1)	BCS_type
Transmission type	
Manual (0)	Trans_type
Automatic (1)	Trans_type
Adaptive climate control	
No (0)	Ada_Clim
Yes (1)	Ada_clim
Adaptive cruise control	
No (0)	Ada_cc
Yes (1)	Ada_cc
Automatic lights	
No (0)	Aut_lights
Yes (1)	Aut_lights
Weather*	

	Yes	No
Sunny	1	0
Cloudy	1	0
Rainy	1	0
Dark	1	0
Outside temperature	Temp	
Trip duration and trip distance		
Trip duration (minutes)	Tr_dur	
Trip distance (kilometres)	Tr_dist	
Questions **Did you feel like I influenced your behaviour in any way by sitting next to you?** Q1 **Did my introduction already give you the impression that I was going to look at HMI controls?** Q2		

*When it was dark and raining, it was counted as raining

Please note that there are separate databases for this study. One database contains the number of times that certain tasks have been performed while the other database contains the frequencies with which certain tasks have been performed. The variable names in the table below have been used for both databases. This means that in the first database, these names stand for the number of times that a task has been performed and in the second database these names stand for the frequency with which a task has been performed. Also, please note that in the second database, the names for the standing still variable do not appear, as the frequency for tasks while standing still could not be calculated due to a lack of time data for standing still. There are also separate databases for the own car data and unfamiliar car data.

Table 23: Variable names in the databases

	Is able to	(While driving)				(Standing still)	**TOTAL**
		(Highway)	(Rural)	(City)	(Jam)		
	Out of 30	**Radio and Media**					
Adjusting volume	Adj_vol_abl	Adj_vol_HW	Adj_vol_RU	Adj_vol_CI	Adj_vol_JA	Adj_vol_ST	Adj_vol_tot
Switching radio station	Swi_rad_abl	Swi_rad_HW	Swi_rad_RU	Swi_rad_CI	Swi_rad_JA	Swi_rad_ST	Swi_rad_tot
Switching app	NA	Swi_app_HW	Swi_app_RU	Swi_app_CI	Swi_app_JA	Swi_app_ST	Swi_app_tot
Switching media input	Swi_med_abl	Swi_med_HW	Swi_med_RU	Swi_med_CI	Swi_med_JA	Swi_med_ST	Swi_med_tot

Switching songs	Swi_son_abl	Swi_son_HW	Swi_son_RU	Swi_son_CI	Swi_son_JA	Swi_son_ST	Swi_son_tot
Switching view on main screen	NA	Swi_view_HW	Swi_view_RU	Swi_view_CI	Swi_view_JA	Swi_view_ST	Swi_view_tot
Turning on/off entertainment system	NA	On_off_radm_HW	On_off_radm_RU	On_off_radm_CI	On_off_radm_JA	On_off_radm_ST	On_off_radm_tot
Switching view on dashboard	NA	Swi_dash_HW	Swi_dash_RU	Swi_dash_CI	Swi_dash_JA	Swi_dash_ST	Swi_dash_tot
Total for category		Radmedia_HW	Radmedia_RU	Radmedia_CI	Radmedia_JA	Radmedia_ST	Radmedia_tot
		Climate control					
Turn. on/off AC	On_off_AC_abl	On_off_AC_HW	On_off_AC_RU	On_off_AC_CI	On_off_AC_JA	On_off_AC_ST	On_off_AC_tot
Changing temperature	Cha_tem_abl	Cha_tem_HW	Cha_tem_RU	Cha_tem_CI	Cha_tem_JA	Cha_tem_ST	Cha_tem_tot
Adjusting fan speed	Adj_fansp_abl	Adj_fansp_HW	Adj_fansp_RU	Adj_fansp_CI	Adj_fansp_JA	Adj_fansp_ST	Adj_fansp_tot
Adjusting fan layout	Adj_fanla_abl	Adj_fanla_HW	Adj_fanla_RU	Adj_fanla_CI	Adj_fanla_JA	Adj_fanla_ST	Adj_fanla_tot
Act/deact heated seat	Act_heatedS_abl	Act_heatedS_HW	Act_heatedS_RU	Act_heatedS_CI	Act_heatedS_JA	Act_haetedS_ST	Act_heatedS_tot
Adj. recirculating mode	Adj_rec_abl	Adj_rec_HW	Adj_rec_RU	Adj_rec_CI	Adj_rec_JA	Adj_rec_ST	Adj_rec_tot
Opening/closing windows	Op_cl_win_abl	Op_cl_win_HW	Op_cl_win_RU	Op_cl_win_CI	Op_cl_win_JA	Op_cl_win_ST	Op_cl_win_tot
Open/close roof	NA	Op_cl_roof_HW	Op_cl_roof_RU	Op_cl_roof_CI	Op_cl_roof_JA	Op_cl_roof_ST	Op_cl_roof_tot
Total for category		Climate_HW	Climate_RU	Climate_CI	Climate_JA	Climate_ST	Climate_tot
		Phone calls					
Answering phone call	Ans_cal_abl	Ans_cal_HW	Ans_cal_RU	Ans_cal_CI	Ans_cal_JA	Ans_cal_ST	Ans_cal_tot
Calling someone	Cal_abl	Cal_HW	Cal_RU	Cal_CI	Cal_JA	Cal_ST	Cal_tot

Total for category		Phone_HW	Phone_RU	Phone_CI	Phone_JA	Phone_ST	Phone_tot
		Lights					
Turn. on/off headlights		On_off_head_HW	On_off_head_RU	On_off_head_CI	On_off_head_JA	On_off_head_ST	On_off_head_tot
Turn. on/off high beam		On_off_bea_HW	On_off_bea_RU	On_off_bea_CI	On_off_bea_JA	On_off_bea_ST	On_off_bea_tot
Turn. on/off mist light		On_off_mist_HW	On_off_mist_RU	On_off_mist_CI	On_off_mist_JA	On_off_mist_ST	On_off_mist_tot
Turn. on/off interior light	On_off_int_abl	On_off_int_HW	On_off_int_RU	On_off_int_CI	On_off_int_JA	On_off_int_ST	On_off_int_tot
Turn. on/off indicator light		On_off_ind_HW	On_off_ind_RU	On_off_ind_CI	On_off_ind_JA	On_off_ind_ST	On_off_ind_tot
Total for category		Lights_HW	Lights_RU	Lights_CI	Lights_JA	Lights_ST	Lights_tot
		Cruise control					
Turn on/off cruise control	On_off_cru_abl	On_off_cru_HW	On_off_cru_RU	On_off_cru_CI	On_off_cru_JA	On_off_cru_ST	On_off_cru_tot
Adj. cruise control speed	Adj_crsp_abl	Adj_crsp_HW	Adj_crsp_RU	Adj_crsp_CI	Adj_crsp_JA	Adj_crsp_ST	Adj_crsp_tot
Cancel/resume cruising	Can_res_cru_abl	Can_res_cru_HW	Can_res_cru_RU	Can_res_cru_CI	Can_res_cru_JA	Can_res_cru_ST	Can_res_cru_tot
Total for category		Cruise_HW	Cruise_RU	Cruise_CI	Cruise_JA	Cruise_ST	Cruise_tot
		Danger signalling					
Using the horn		Usi_hor_HW	Usi_hor_RU	Usi_hor_CI	Usi_hor_JA	Usi_hor_ST	Usi_hor_tot
Turn. on/off hazard lights		On_off_haz_HW	On_off_haz_RU	On_off_haz_CI	On_off_haz_JA	On_off_haz_ST	On_off_haz_tot
Total for category		Signal_HW	Signal_RU	Signal_CI	Signal_JA	Signal_ST	Signal_tot
		Windshield tasks					
Chang. WW speed front*		Cha_WWspf_HW	Cha_WWspf_RU	Cha_WWspf_CI	Cha_WWspf_JA	Cha_WWspf_ST	Cha_WWspf_tot
Chang. WW speed back*		Cha_WWspb_HW	Cha_WWspb_RU	Cha_WWspb_CI	Cha_WWspb_JA	Cha_WWspb_ST	Cha_WWspb_tot
Using window fluid front		Usi_fluf_HW	Usi_fluf_RU	Usi_fluf_CI	Usi_fluf_JA	Usi_fluf_ST	Usi_fluf_tot
Using window fluid back	Usi_flub_abl	Usi_flub_HW	Usi_flub_RU	Usi_flub_CI	Usi_flub_JA	Usi_flub_ST	Usi_flub_tot
Chang. sun visor position	Cha_vis_abl	Cha_vis_HW	Cha_vis_RU	Cha_vis_CI	Cha_vis_JA	Cha_vis_ST	Cha_vis_tot

Act/deact W-heater front	Act_Wheatf_abl	Act_Wheatf_HW	Act_Wheatf_RU	Act_Wheatf_CI	Act_Wheatf_JA	Act_Wheatf_ST	Act_Wheatf_tot
Act/deact W-heater back	Act_Wheatb_abl	Act_Wheatb_HW	Act_Wheatb_RU	Act_Wheatb_CI	Act_Wheatb_JA	Act_Wheatb_ST	Act_Wheatb_tot
Enabling automatic windshield wipers	NA	Ena_autww_HW	Ena_autww_RU	Ena_autww_CI	Ena_autww_JA	Ena_autww_ST	Ena_autww_tot
Total for category		Windshield_HW	Windshield_RU	Windshield_CI	Windshield_JA	Windshield_ST	Windshield_tot
		Adjusting car setup					
Adj. inside mirror		Adj_insmir_HW	Adj_insmir_RU	Adj_insmir_CI	Adj_insmir_JA	Adj_insmir_ST	Adj_insmir_tot
Adj. outside mirror(s)		Adj_outsmir_HW	Adj_outsmir_RU	Adj_outsmir_CI	Adj_outsmir_JA	Adj_outsmir_ST	Adj_outsmir_tot
Adjusting seat		Adj_seat_HW	Adj_seat_RU	Adj_seat_CI	Adj_seat_JA	Adj_seat_ST	Adj_seat_tot
Total for category		Mirror_HW	Mirror_RU	Mirror_CI	Mirror_JA	Mirror_ST	Mirror_tot
		Adjusting settings					
Changing auto distance	Cha_autdis_abl	Cha_autdis_HW	Cha_autdis_RU	Cha_autdis_CI	Cha_autdis_JA	Cha_autdis_ST	Cha_autdis_tot
Changing sound mix	Cha_soumix_abl	Cha_soumix_HW	Cha_soumix_RU	Cha_soumix_CI	Cha_soumix_JA	Cha_soumix_ST	Cha_soumix_tot
Changing driving mode	NA	Cha_mode_HW	Cha_mode_RU	Cha_mode_CI	Cha_mode_JA	Cha_mode_ST	Cha_mode_tot
Total for category		Settings_HW	Settings_RU	Settings_CI	Settings_JA	Settings_ST	Settings_tot
		Using extra features					
Enabl. auto steering	Ena_autste_abl	Ena_autste_HW	Ena_autste_RU	Ena_autste_CI	Ena_autste_JA	Ena_autste_ST	Ena_autste_tot
Enabl. auto parking	Ena_autpar_abl	Ena_autpar_HW	Ena_autpar_RU	Ena_autpar_CI	Ena_autpar_JA	Ena_autpar_ST	Ena_autpar_tot
Enabl. auto distance	Ena_autdis_abl	Ena_autdis_HW	Ena_autdis_RU	Ena_autdis_CI	Ena_autdis_JA	Ena_autdis_ST	Ena_autdis_tot
Enabl. auto lane-keeping	Ena_autlank_abl	Ena_autlank_HW	Ena_autlank_RU	Ena_autlank_CI	Ena_autlank_JA	Ena_autlank_ST	Ena_autlank_tot
Total for category		Features_HW	Features_RU	Features_CI	Features_JA	Features_ST	Features_tot
Other							

Using handbrake	NA	Usi_brake_HW	Usi_brake_RU	Usi_brake_CI	Usi_brake_JA	Usi_brake_ST	Usi_brake_tot
Opening/closing compartment	NA	Op_cl_comp_HW	Op_cl_comp_RU	Op_cl_comp_CI	Op_cl_comp_JA	Op_cl_comp_ST	Op_cl_comp_tot
Moving arm rest	NA	Mov_arm_HW	Mov_arm_RU	Mov_arm_CI	Mov_arm_JA	Mov_arm_ST	Mov_arm_tot
Plugging in/removing cable from a socket	NA	Plug_cable_HW	Plug_cable_RU	Plug_cable_CI	Plug_cable_JA	Plug_cable_ST	Plug_cable_tot
Total for category		Other_HW	Other_RU	Other_CI	Other_JA	Other_ST	Other_tot
Total		HW_tot	RU_tot	CI_tot	JA_tot	ST_tot	TOTAL

Appendix C: Initial observation checklist

In Table 24, the initial observation checklist is presented. The frequencies will be noted for different road types, whether the driver is experiencing a traffic jam and whether the driver is standing still. For each time that a driver performs a task as defined in the study setup, a line will be drawn in the correct box. When a particular part of the familiar route is closed and thus the driver has to drive a detour, the stripes will be adjusted to make a distinction between the normal route and the detour part. Also, it will be noted whether a car “is able to” perform the tasks listed at all. This is noted to be able to draw correct conclusions later on which tasks are used relatively often. When a certain task cannot be performed in most of the cars that are used, the “is able to” category will help to explain why the frequencies for this task are relatively low.

Table 24: Observation checklist and Dutch version of statements

	Is able to	(While driving)				(Standing still)
		(Highway)	(Rural)	(City)	(Jam)	
		Radio and Media				
Adjusting volume						
Switching radio station						
Switching app						
Switching media input						
Switching songs						
		Climate control				
Turn. on/off AC						
Changing temperature						
Adjusting fan speed						
Adjusting fan layout						
Act/deact heated seat						
Adj. recirculating mode						
Opening/closing windows						
		Phone calls				
Answering phone call						
Calling someone						
		Lights				
Turn. on/off headlights						
Turn. on/off high beam						
Turn. on/off mist light						
Turn. on/off interior light						
Turn. on/off indicator light						
		Cruise control				
Turn on/off cruise control						
Adj. cruise control speed						
Cancel/resume cruising						
		Danger signalling				
Using the horn						
Turn. on/off hazard lights						
		Windshield tasks				
Chang. WW speed front*						
Chang. WW speed back*						
Using window fluid front						
Using window fluid back						

Chang. sun visor position						
Act/deact W-heater front						
Act/deact W-heater back						
		Mirrors				
Adj. inside mirror						
Adj. outside mirror(s)						
		Adjusting settings				
Changing auto distance						
Changing sound mix						
		Using extra features				
Enabl. auto steering						
Enabl. auto parking						
Enabl. auto distance						
Enabl. auto lane-keeping						
		Room for extra tasks				

Route characteristics	
Highway road (time)	
Rural road (time)	
City road (time)	
Traffic jam (time)	
Board computer system type	
matrix screen / No screen	
Screen with menu's	
Touch screen with menu's	
Transmission type	
Manual	
Automatic	
Other car characteristics	
Adaptive climate control (yes/no)	
Adapt. CC (yes/no)	
Automatic lights (yes/no)	
Weather	
Sunny	
Rainy	
Cloudy	
Dark	
Outside temperature	
Trip duration and trip distance	
Trip duration (minutes)	
Trip distance (kilometres)	
Room for comments	

Questions
Did you feel like I influenced your behaviour in any way by sitting next to you?   **Did my introduction already give you the impression that I was going to look at HMI controls?**
I = One task performed on a familiar route T = One task performed on a detour
Thank you for participating in the research. As previously mentioned in the invitation, this research will be performed in collaboration with the TU Delft and the RDW to better understand driving behaviour. To achieve this goal, we will drive along with as many people as possible and make observations. The idea is to do as you always do. To make this possible, we would like to ask you to pretend that you are alone in the car. This means that we will not have any conversations during the trip. Don't worry, we won't pay attention to how well you drive and it is certainly not a driving test. The data we collect is completely anonymous. And once again, you are doing it well when you just do everything like you normally do when we are not present.
Respondent number: **Age** (years)**:** **Gender** (Male/female/other)**:**   **Driving license** (years)**:** **Driving frequency** (Times per week last year)**:**   **Route familiarity** (1-10)**:** **I am familiar with this car** (1-10): **I drive this car a lot** (1-10): **I know the features/options that the car has available** (1-10): **I know where the buttons are located** (1-10): **I understand the dashboard and the things it displays** (1-10):

*WW = windshield wipers
**W-heater = window heater
***AC = air conditioner
****CC = Cruise control

Bedankt voor het deelnemen aan het onderzoek. Zoals eerder vermeld in de uitnodiging zal dit onderzoek uitgevoerd worden in samenwerking met de TU Delft en het RDW om rijgedrag beter te leren begrijpen. Om dit doel te bereiken zullen we met zo veel mogelijk mensen meerijden en observaties doen. Het is de bedoeling om te rijden zoals je altijd doet. Om dit mogelijk te maken willen we je dan ook vragen om te doen alsof je alleen in de auto bent. Dit betekent dus ik tijdens de trip ook geen gesprekken zal voeren. Geen zorgen, we zullen niet letten op hoe goed je rijdt en het is ook zeker geen rijexamen. De data die we verzamelen is volledig anoniem. En nogmaals, je doet het goed wanneer je gewoon doet wat je normaal ook doet als wij er niet bij zijn.
Ik ben bekend met deze auto (1-10): **Ik rij deze auto veel** (1-10): **Ik weet welke functies/opties de auto heeft** (1-10): **Ik weet waar de knoppen zitten** (1-10): **Ik begrijp het dashboard en de dingen die het weergeeft** (1-10):

Questions
Heb je het gevoel dat ik jouw gedrag op welke manier dan ook beïnvloed heb door naast je te zitten? **Wist je door mijn introductie al dat ik zou gaan kijken naar HMI taken of had je geen idee?**

Appendix D: Raw results

Table 25: Average number of times a task has been performed per trip over all participants (Own car data)

Is able to		(While driving)				(Standing still)	TOTAL
		(Highway)	(Rural)	(City)	(Jam)		
Out of 30		Radio and Media					
Adjusting volume	30	0,10	0,00	0,63	0,00	0,23	0,97
Switching radio station	30	0,00	0,00	0,07	0,00	0,07	0,13
Switching app	N/A	0,00	0,00	0,07	0,00	0,00	0,07
Switching media input	30	0,00	0,00	0,03	0,00	0,00	0,03
Switching songs	30	0,00	0,00	0,07	0,00	0,00	0,07
Switching screen view	N/A	0,00	0,00	0,03	0,00	0,00	0.03
Turn. on/off media system	N/A	0,00	0,03	0,00	0,00	0,00	0,03
Switching dashboard view	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	30	0,10	0,03	0,90	0,00	0,30	1,33
		Climate control					
Turn. on/off AC	29	0,00	0,00	0,00	0,00	0,00	0,00
Changing temperature	30	0,00	0,03	0,50	0,00	0,07	0,60
Adjusting fan speed	30	0,00	0,03	0,27	0,00	0,17	0,47
Adjusting fan layout	30	0,00	0,00	0,00	0,00	0,03	0,03
Act/deact heated seat	12	0,00	0,00	0,00	0,00	0,00	0,00
Adj. recirculating mode	30	0,00	0,00	0,00	0,00	0,00	0,00
Opening/closing windows	29	0,00	0,00	0,07	0,00	0,07	0,13
Open/close roof	N/A	0,00	0,00	0,00	0,00	0.03	0.03
Total for category	12	0,00	0,07	0,83	0,00	0,36	1,26
		Phone calls					
Answering phone call	23	0,00	0,00	0,00	0,00	0,00	0,00
Calling someone	23	0,00	0,00	0,03	0,00	0,00	0,03
Total for category	23	0,00	0,00	0,03	0,00	0,00	0,03
		Lights					
Turn. on/off headlights		0,00	0,00	0,00	0,00	0,00	0,00
Turn. on/off high beam		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off mist light		0,00	0,00	0,00	0,00	0,00	0,00
Turn. on/off interior light		0,00	0,00	0,00	0,00	0,00	0,00
Turn. on/off indicator light		0,90	2,23	20,97	0,07	0,00	24,17
Total for category		0,90	2,23	21,00	0,07	0,00	24,20
		Cruise control					
Turn on/off cruise control	23	0,00	0,00	0,00	0,00	0,00	0,00
Adj. cruise control speed	23	0,00	0,00	0,00	0,00	0,00	0,00
Cancel/resume cruising	23	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	23	0,00	0,00	0,00	0,00	0,00	0,00
		Danger signalling					
Using the horn		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off hazard lights		0,00	0,00	0,00	0,00	0,00	0,00
Total for category		0,00	0,00	0,03	0,00	0,00	0,03
		Windshield tasks					
Chang. WW speed front*		0,00	0,03	0,50	0,00	0,00	0,53
Chang. WW speed back*		0,00	0,00	0,13	0,00	0,10	0,23
Using window fluid front		0,03	0,03	0,03	0,00	0,00	0,10
Using window fluid back	29	0,00	0,00	0,00	0,00	0,00	0,00

Chang. sun visor position	30	0,00	0,07	0,40	0,07	0,13	0,67
Act/deact W-heater front	1	0,00	0,00	0,00	0,00	0,00	0,00
Act/deact W-heater back	29	0,00	0,00	0,03	0,00	0,00	0,03
Enabl. Auto WW	N/A	0,00	0,00	0,03	0,00	0,00	0,03
Total for category	1	0,03	0,13	1,13	0,07	0,23	1,60
		Adjusting car setup					
Adj. inside mirror		0,00	0,00	0,00	0,00	0,03	0,03
Adj. outside mirror(s)		0,00	0,00	0,00	0,00	0,00	0,00
Adj. Seat		0,00	0,00	0,00	0,00	0,03	0,03
Total for category		0,00	0,00	0,00	0,00	0,07	0,07
		Adjusting settings					
Changing auto distance	11	0,00	0,00	0,00	0,00	0,00	0,00
Changing sound mix	28	0,00	0,00	0,00	0,00	0,00	0,00
Changing driving mode	N/A	0,00	0,00	0,03	0,00	0,00	0,03
Total for category	11	0,00	0,00	0,03	0,00	0,00	0,03
		Using extra features					
Enabl. auto steering	2	0,00	0,13	0,00	0,13	0,00	0,27
Enabl. auto parking	1	0,00	0,00	0,00	0,00	0,00	0,00
Enabl. auto distance	11	0,00	0,00	0,00	0,00	0,00	0,00
Enabl. auto lane-keeping	13	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	1	0,00	0,13	0,00	0,13	0,00	0,27
	Other						
Move cable from socket	N/A	0,00	0,00	0,07	0,00	0,00	0,07
Using handbrake	N/A	0,00	0,00	0,00	0,00	0,07	0,07
Open/close compartment	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Move arm rest	N/A	0,00	0,00	0,00	0,00	0,00	0,00
		0,00	0,00	0,07	0,00	0,07	0,13
Total		**1,03**	**2,59**	**24,02**	**0,27**	**1,03**	**28,95**

Table 26: Average task frequency per trip (own car data)

	Is able to	**(**While driving)				(Standing still)	**TOTAL AVERAGE**
		(Highway)	(Rural)	(City)	(Jam)		
	Out of 30	**Radio and Media**					
Adjusting volume	30	7,50	0,00	2,28	0,00	NA	2,67
Switching radio station	30	0,00	0,00	0,28	0,00	NA	0,42
Switching app	N/A	0,00	0,00	0,20	0,00	NA	0,20
Switching media input	30	0,00	0,00	0,11	0,00	NA	0,11
Switching songs	30	0,00	0,00	0,22	0,00	NA	0,22
Switching screen view	N/A	0,00	0,00	0,11	0,00	NA	0,11
Turn. on/off media system	N/A	0,00	0,88	0,00	0,00	NA	0,07
Switching dashboard view	N/A	0,00	0,00	0,00	0,00	NA	0,00
Total for category	30	7,50	0,88	3,19	0,00	NA	3,79
		Climate control					
Turn. on/off AC	29	0,00	0,00	0,00	0,00	NA	0,00
Changing temperature	30	0,00	0,88	1,64	0,00	NA	1,70
Adjusting fan speed	30	0,00	0,59	0,90	0,00	NA	1,21
Adjusting fan layout	30	0,00	0,00	0,00	0,00	NA	0,10
Act/deact heated seat	12	0,00	0,00	0,00	0,00	NA	0,00
Adj. recirculating mode	30	0,00	0,00	0,00	0,00	NA	0,00
Opening/closing windows	29	0,00	0,00	0,18	0,00	NA	0,34
Open/close roof	N/A	0,00	0,00	0,00	0,00	NA	0,10

Total for category	12	0,00	1,47	2,73	0,00	NA	3,45
		Phone calls					
Answering phone call	23	0,00	0,00	0,00	0,00	NA	0,00
Calling someone	23	0,00	0,00	0,18	0,00	NA	0,11
Total for category	23	0,00	0,00	0,18	0,00	NA	0,11
		Lights					
Turn. on/off headlights		0,00	0,00	0,00	0,00	NA	0,00
Turn. on/off high beam		0,00	0,00	0,11	0,00	NA	0,11
Turn. on/off mist light		0,00	0,00	0,00	0,00	NA	0,00
Turn. on/off interior light		0,00	0,00	0,00	0,00	NA	0,00
Turn. on/off indicator light		65,50	39,75	69,53	6,00	NA	66,22
Total for category		65,50	39,75	69,64	6,00	NA	66,33
		Cruise control					
Turn on/off cruise control	23	0,00	0,00	0,00	0,00	NA	0,00
Adj. cruise control speed	23	0,00	0,00	0,00	0,00	NA	0,00
Cancel/resume cruising	23	0,00	0,00	0,00	0,00	NA	0,00
Total for category	23	0,00	0,00	0,00	0,00	NA	0,00
		Danger signalling					
Using the horn		0,00	0,00	0,10	0,00	NA	0,10
Turn. on/off hazard lights		0,00	0,00	0,00	0,00	NA	0,00
Total for category		0,00	0,00	0,10	0,00	NA	0,10
		Windshield tasks					
Chang. WW speed front*		0,00	0,35	1,54	0,00	NA	1,56
Chang. WW speed back*		0,00	0,00	0,43	0,00	NA	0,77
Using window fluid front		2,50	0,35	0,10	0,00	NA	0,25
Using window fluid back	29	0,00	0,00	0,00	0,00	NA	0,00
Chang. sun visor position	30	0,00	1,39	1,57	6,00	NA	1,71
Act/deact W-heater front	1	0,00	0,00	0,00	0,00	NA	0,00
Act/deact W-heater back	29	0,00	0,00	0,11	0,00	NA	0,11
Enabl. Auto WW	N/A	0,00	0,00	0,11	0,00	NA	0,11
Total for category	1	2,50	2,09	3,85	6,00	NA	4,50
		Adjusting car setup					
Adj. inside mirror		0,00	0,00	0,00	0,00	NA	0,11
Adj. outside mirror(s)		0,00	0,00	0,00	0,00	NA	0,00
Adj. Seat		0,00	0,00	0,00	0,00	NA	0,10
Total for category		0,00	0,00	0,00	0,00	NA	0,21
		Adjusting settings					
Changing auto distance	11	0,00	0,00	0,00	0,00	NA	0,00
Changing sound mix	28	0,00	0,00	0,00	0,00	NA	0,00
Changing driving mode	N/A	0,00	0,00	0,08	0,00	NA	0,08
Total for category	11	0,00	0,00	0,08	0,00	NA	0,08
		Using extra features					
Enabl. auto steering	2	0,00	3,53	0,00	20,00	NA	0,89
Enabl. auto parking	1	0,00	0,00	0,00	0,00	NA	0,00
Enabl. auto distance	11	0,00	0,00	0,00	0,00	NA	0,00
Enabl. auto lane-keeping	13	0,00	0,00	0,00	0,00	NA	0,00
Total for category	1	0,00	3,53	0,00	20,00	NA	0,89
	Other						
Move cable from socket	N/A	0,00	0,00	0,21	0,00	NA	0,21
Using handbrake	N/A	0,00	0,00	0,00	0,00	NA	0,22
Open/close compartment	N/A	0,00	0,00	0,00	0,00	NA	0,00
Move arm rest	N/A	0,00	0,00	0,00	0,00	NA	0,00
		0,00	0,00	0,21	0,00	NA	0,43

Total		**75,50**	**47,73**	**79,99**	**32,00**	**NA**	**79,89**

Table 27: Average number of times a task has been performed per trip over all participants (unfamiliar car data)

	Is able to	**(**While driving)				(Standing still)	**TOTAL AVERAGE**
		(Highway)	(Rural)	(City)	(Jam)		
	Out of 30	**Radio and Media**					
Adjusting volume	Yes	0,07	0,00	0,43	0,00	0,13	0,63
Switching radio station	Yes	0,03	0,00	0,07	0,03	0,03	0,17
Switching app	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Switching media input	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Switching songs	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Switching screen view	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Turn. on/off media system	N/A	0,00	0,00	0,13	0,00	0,00	0,13
Switching dashboard view	N/A	0,00	0,00	0,00	0,03	0,03	0,07
Total for category	Yes	0,1	0,00	0,63	0,07	0,19	1
		Climate control					
Turn. on/off AC	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Changing temperature	Yes	0,03	0,00	0,07	0,00	0,00	0,10
Adjusting fan speed	Yes	0,00	0,00	0,00	0,00	0,03	0,03
Adjusting fan layout	Yes	0,00	0,00	0,00	0,00	0,03	0,03
Act/deact heated seat	No	0,00	0,00	0,00	0,00	0,00	0,00
Adj. recirculating mode	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Opening/closing windows	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Open/close roof	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	No	0,03	0,00	0,07	0,00	0,07	0,17
		Phone calls					
Answering phone call	No	0,00	0,00	0,00	0,00	0,00	0,00
Calling someone	No	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	No	0,00	0,00	0,00	0,00	0,00	0,00
		Lights					
Turn. on/off headlights		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off high beam		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off mist light		0,00	0,00	0,00	0,00	0,00	0,00
Turn. on/off interior light		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off indicator light		1,03	2,37	21,43	0,10	0,00	24,93
Total for category		1,03	2,37	21,53	0,10	0,00	25,03
		Cruise control					
Turn on/off cruise control	Yes	0,00	0,10	0,07	0,00	0,00	0,17
Adj. cruise control speed	Yes	0,00	0,00	0,03	0,00	0,00	0,03
Cancel/resume cruising	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	Yes	0,00	0,10	0,10	0,00	0,00	0,20
		Danger signalling					
Using the horn		0,00	0,00	0,03	0,00	0,00	0,03
Turn. on/off hazard lights		0,00	0,00	0,00	0,00	0,00	0,00
Total for category		0,00	0,00	0,03	0,00	0,00	0,03
		Windshield tasks					
Chang. WW speed front*		0,17	0,00	2,23	0,00	0,13	2,53
Chang. WW speed back*		0,03	0,00	0,50	0,00	0,10	0,63
Using window fluid front		0,00	0,00	0,10	0,00	0,00	0,10
Using window fluid back	Yes	0,00	0,00	0,00	0,00	0,00	0,00
Chang. sun visor position	Yes	0,00	0,10	0,37	0,00	0,03	0,50

Act/deact W-heater front	No	0,00	0,00	0,00	0,00	0,00	0,00
Act/deact W-heater back	Yes	0,17	0,00	2,23	0,00	0,13	2,53
Enabl. Auto WW	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	No	0,20	0,10	3,20	0,00	0,27	3,77
		Adjusting car setup					
Adj. inside mirror		0,00	0,03	0,20	0,00	0,00	0,23
Adj. outside mirror(s)		0,00	0,00	0,00	0,00	0,00	0,00
Adj. Seat		0,00	0,00	0,10	0,00	0,03	0,13
Total for category		0,00	0,03	0,30	0,00	0,03	0,37
		Adjusting settings					
Changing auto distance	No	0,00	0,00	0,00	0,00	0,00	0,00
Changing sound mix	Yes	0,00	0,00	0,03	0,00	0,00	0,03
Changing driving mode	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	No	0,00	0,00	0,03	0,00	0,00	0,03
		Using extra features					
Enabl. auto steering	No	0,00	0,00	0,00	0,00	0,00	0,00
Enabl. auto parking	No	0,00	0,00	0,00	0,00	0,00	0,00
Enabl. auto distance	No	0,00	0,00	0,00	0,00	0,00	0,00
Enabl. auto lane-keeping	No	0,00	0,00	0,00	0,00	0,00	0,00
Total for category	No	0,00	0,00	0,00	0,00	0,00	0,00
	Other						
Move cable from socket	N/A	0,00	0,00	0,00	0,00	0,00	0,00
Using handbrake	N/A	0,00	0,00	0,00	0,00	0,20	0,20
Open/close compartment	N/A	0,00	0,00	0,00	0,00	0,07	0,07
Move arm rest	N/A	0,00	0,00	0,03	0,00	0,00	0,03
	NA	0,00	0,00	0,03	0,00	0,27	0,3
Total		**1,36**	**2,6**	**25,92**	**0,17**	**0,83**	**30,9**

Table 28: Average task frequency per trip (Unfamiliar car data)

Is able to		(While driving)				(Standing still)	**TOTAL**
		(Highway)	(Rural)	(City)	(Jam)		
Out of 30		**Radio and Media**					
Adjusting volume	Yes	5,00	0,00	1,40	0,00	NA	1,64
Switching radio station	Yes	2,50	0,00	0,19	15,00	NA	0,39
Switching app	N/A	0,00	0,00	0,00	0,00	NA	0,00
Switching media input	Yes	0,00	0,00	0,00	0,00	NA	0,00
Switching songs	Yes	0,00	0,00	0,00	0,00	NA	0,00
Switching screen view	N/A	0,00	0,00	0,00	0,00	NA	0,00
Turn. on/off media system	N/A	0,00	0,00	0,80	0,00	NA	0,47
Switching dashboard view	N/A	0,00	0,00	0,00	5,00	NA	0,24
Total for category	Yes	7,50	0,00	2,39	20,00	NA	2,73
		Climate control					
Turn. on/off AC	Yes	0,00	0,00	0,00	0,00	NA	0,00
Changing temperature	Yes	2,50	0,00	0,17	0,00	NA	0,23
Adjusting fan speed	Yes	0,00	0,00	0,00	0,00	NA	0,07
Adjusting fan layout	Yes	0,00	0,00	0,00	0,00	NA	0,10
Act/deact heated seat	No	0,00	0,00	0,00	0,00	NA	0,00
Adj. recirculating mode	Yes	0,00	0,00	0,00	0,00	NA	0,00
Opening/closing windows	Yes	0,00	0,00	0,00	0,00	NA	0,00
Open/close roof	N/A	0,00	0,00	0,00	0,00	NA	0,00
Total for category	No	2,50	0,00	0,17	0,00	NA	0,40

		Phone calls					
Answering phone call	No	0,00	0,00	0,00	0,00	NA	0,00
Calling someone	No	0,00	0,00	0,00	0,00	NA	0,00
Total for category	No	0,00	0,00	0,00	0,00	NA	0,00
		Lights					
Turn. on/off headlights		0,00	0,00	0,20	0,00	NA	0,11
Turn. on/off high beam		0,00	0,00	0,11	0,00	NA	0,11
Turn. on/off mist light		0,00	0,00	0,00	0,00	NA	0,00
Turn. on/off interior light		0,00	0,00	0,33	0,00	NA	0,13
Turn. on/off indicator light		73,00	40,44	67,99	25,00	NA	66,58
Total for category		73,00	40,44	68,62	25,00	NA	66,92
		Cruise control					
Turn on/off cruise control	Yes	0,00	2,65	0,40	0,00	NA	0,59
Adj. cruise control speed	Yes	0,00	0,00	0,20	0,00	NA	0,12
Cancel/resume cruising	Yes	0,00	0,00	0,00	0,00	NA	0,00
Total for category	Yes	0,00	2,65	0,60	0,00	NA	0,71
		Danger signalling					
Using the horn		0,00	0,00	0,10	0,00	NA	0,10
Turn. on/off hazard lights		0,00	0,00	0,00	0,00	NA	0,00
Total for category		0,00	0,00	0,10	0,00	NA	0,10
		Windshield tasks					
Chang. WW speed front*		12,50	0,00	6,70	0,00	NA	6,88
Chang. WW speed back*		2,50	0,00	1,50	0,00	NA	1,62
Using window fluid front		0,00	0,00	0,37	0,00	NA	0,28
Using window fluid back	Yes	0,00	0,00	0,00	0,00	NA	0,00
Chang. sun visor position	Yes	0,00	2,06	1,07	0,00	NA	1,16
Act/deact W-heater front	No	0,00	0,00	0,00	0,00	NA	0,00
Act/deact W-heater back	Yes	0,00	0,00	0,00	0,00	NA	0,00
Enabl. Auto WW	N/A	0,00	0,00	0,00	0,00	NA	0,00
Total for category	No	15,00	2,06	9,65	0,00	NA	9,94
		Adjusting car setup					
Adj. inside mirror		0,00	0,39	0,69	0,00	NA	0,66
Adj. outside mirror(s)		0,00	0,00	0,00	0,00	NA	0,00
Adj. Seat		0,00	0,00	0,28	0,00	NA	0,37
Total for category		0,00	0,39	0,97	0,00	NA	1,03
		Adjusting settings					
Changing auto distance	No	0,00	0,00	0,00	0,00	NA	0,00
Changing sound mix	Yes	0,00	0,00	0,09	0,00	NA	0,08
Changing driving mode	N/A	0,00	0,00	0,00	0,00	NA	0,00
Total for category	No	0,00	0,00	0,09	0,00	NA	0,08
		Using extra features					
Enabl. auto steering	No	0,00	0,00	0,00	0,00	NA	0,00
Enabl. auto parking	No	0,00	0,00	0,00	0,00	NA	0,00
Enabl. auto distance	No	0,00	0,00	0,00	0,00	NA	0,00
Enabl. auto lane-keeping	No	0,00	0,00	0,00	0,00	NA	0,00
Total for category	No	0,00	0,00	0,00	0,00	NA	0,00
	Other						
Move cable from socket	N/A	0,00	0,00	0,00	0,00	NA	0,00
Using handbrake	N/A	0,00	0,00	0,00	0,00	NA	0,62
Open/close compartment	N/A	0,00	0,00	0,00	0,00	NA	0,20
Move arm rest	N/A	0,00	0,00	0,09	0,00	NA	0,08
	NA	0,00	0,00	0,09	0,00	NA	0,90
Total		**98,00**	**45,54**	**82,68**	**45,00**	**NA**	**82,82**

Appendix E: Data analysis

Table 29: Trip context

Trip context				
Variable	**Category**	**Amount**	**Percentage**	**N**
Weather	Sunny	8	13.33	60
	Rainy	14	23.33	
	Cloudy	28	46.67	
	Dark	10	16.67	
Peak traffic hours	Yes	15	25	60
	No	45	75	

The most common weather type was found to be cloudy weather and most trips were performed outside of peak traffic hours. For the weather and peak traffic hours variables, the data from both trips has been used.

Table 30: Car familiarity statements average scores for the whole study

Statements measuring car familiarity (avg. = 7.16)					
Variable	**Average**	**Min**	**Max**	**Std dev**	**N**
Frequency driving the car	5.27	1.00	10.00	4.18	60
Feature and option knowledge of the car	6.92	1.00	10.00	2.61	60
Button location knowledge of the car	7.60	4.00	10.00	2.02	60
Dashboard understanding of the car	8.85	5.00	10.00	1.31	60

The average car familiarity across both cars was found to be 7.12. To calculate this average, 60 data points were used, as the participants scored this for both their own car and the unfamiliar car. The four statements that try to measure car familiarity in more depth had an average score of 7.16 across the different statements. The individual statement with the highest average was found to be the statement about the dashboard familiarity whilst the lowest average was found for the statement about the frequency of driving a particular car.

Table 31: Car characteristics of the two other unfamiliar cars

Variable	Category	Amount	Percentage	N
Ad. Climate	Yes	1	50	2
	No	1	50	
Ad. Cruise	Yes	0	0	2
	No	2	100	
Auto Lights	Yes	1	50	2
	No	1	50	
Transmission	Manual	1	100	2
	Automatic	1	50	
Touch screen	Yes	0	0	2
	No	2	100	

Table 32: Number of trips per day of the week

Variable	Category	Amount	Percentage	N
Day	Monday	5	16.67	60
	Tuesday	11	36.67	
	Wednesday	14	46.67	
	Thursday	13	43.33	
	Friday	7	23.33	
	Saturday	8	26.67	
	Sunday	2	6.67	

Most trips were performed during the week but still 10 out of the total 60 trips were performed in the weekend.

The table below shows the 10 most used tasks during the 60 trips. The table is ordered from the most used tasks to the least used tasks. Also, the average number of times a particular task has been performed over the 60 trips is shown, as well as the average frequency per hour. Furthermore, the minimum amount and maximum amount of times that a task has been performed during a single trip are shown. The table also shows how many of the cars of the participants were able to perform the tasks. It is also shown whether or not the Seat Toledo and the other two cars were able to perform the tasks.

Table 33: Top 10 most used tasks using the data from both trips combined

Task **N = 60**	**Average**	**Total**	**Min**	**Max**	**Avg freq per hour**	**Own car**	**Seat / other cars**
Using indicator light	24.55	1473.00	11.00	40.00	66.40	30	Yes / 2
Changing front windshield wiper speed	1.53	92.00	0.00	29.00	4.22	30	Yes / 2
Adjusting volume	0.80	48.00	0.00	5.00	2.15	30	Yes / 2
Moving sun visor	0.58	35.00	0.00	9.00	1.44	30	Yes / 2
Changing rear windshield wiper speed	0.43	26.00	0.00	7.00	1.19	30	Yes / 2
Adjusting temperature	0.35	21.00	0.00	3.00	0.96	30	Yes / 2
Adjusting fan speed	0.25	15.00	0.00	4.00	0.64	30	Yes / 2
Switching radio station	0.15	9.00	0.00	2.00	0.41	30	Yes / 2
Adjusting inside mirror	0.13	8.00	0.00	2.00	0.38	30	Yes / 2
Enabling/disabling auto pilot	0.13	8.00	0.00	8.00	0.44	2	No / 0

To test if the differences in average frequencies between the tasks are statistically significant a repeated measures ANOVA has been performed. Afterwards, pairwise comparisons have been performed to find out which pairs have statistically different frequencies. The results of these tests can be found in the table below.

Table 34: Repeated measures ANOVA for the means of the frequencies of the 10 tasks in Table 33

Task counts	**F**	**P-value**
Mauchly's test of sphericity		<0.001
Greenhouse-Geisser	402.03	<0.001

Task	**Different mean from***
Using indicator light	All tasks
Adjusting volume	Changing indicator light, Switching radio, adjusting inside mirror
Rest	Using indicator light and/or adjusting volume

*Adjustment for multiple comparisons: Bonferroni.

The Greenhouse-Geisser coefficient is used for the repeated measures ANOVA as the data does not satisfy the sphericity characteristic. The repeated ANOVA test shows that at least one of the frequencies is statistically significantly different. Therefore, pairwise comparisons were made and Bonferroni adjustment was used to adjust for the fact that multiple comparisons were made. The results show that the average frequency of the indicator light is statistically different from the average frequency of all other tasks. The average frequency of the volume is also statistically different from a few of the other tasks. These are switching radio stations and adjusting the inside mirror. All the other tasks only have statistically significantly different average frequencies from using the indicator light and/or adjusting the volume.

Table 35 shows an aggregated approach. Here, the total number of tasks performed over all trips is presented. Also, the total number of tasks performed per category is shown for all categories that have more than 10 tasks performed. Since the indicator light is used so much more than the other tasks, also the total number of tasks performed over all trips excluding the indicator light is presented. Lastly, a variable was created that excludes the front windshield wiper, since this variable was also an outlier.

Table 35: Aggregated descriptive analysis using the data from both trips combined

Variable N = 60	**Average**	**Total**	**Min**	**Max**	**Avg freq per hour**
Number of tasks used in total					
Total	29.93	1796.00	13.00	71.00	81.35
Total excluding indicator light	5.38	323.00	0.00	32.00	14.95
Total excluding indicator light and front windshield wiper	3.85	231.00	0.00	14.00	10.74
Number of tasks used per category					
Lights category	24.62	1477.00	11.00	40.00	66.63
Windshield category	2.68	161.00	0.00	32.00	7.22
Radio and media category	1.17	70.00	0.00	6.00	3.26
Climate control category	0.72	43.00	0.00	5.00	1.93
Setup category	0.22	13.00	0.00	2.00	0.62
Other category	0.22	13.00	0.00	2.00	0.67

To test if the differences in average frequencies between the categories are statistically significant a repeated measures ANOVA has been performed. Afterwards, pairwise comparisons have been performed to find out which pairs have statistically different frequencies. The results of these tests can be found in the table below.

Table 36: Repeated measures ANOVA for the means of the frequencies of the categories in Table 35

Task counts	F	Sig.
Greenhouse-Geisser	371.20	<0.001
Category	**Different mean from***	
Lights category	All categories	
Windshield category	Lights category, Setup category, Other category	
Radio and media category	Lights category, Setup category, Other category	
Rest	Lights category and/or Windshield category and/or Radio and media category	

*Adjustment for multiple comparisons: Bonferroni.

The Greenhouse-Geisser coefficient is used for the repeated measures ANOVA as the data does not satisfy the sphericity characteristic. The repeated ANOVA test shows that at least one of the frequencies is statistically significantly different. Therefore, pairwise comparisons were made and Bonferroni adjustment was used to adjust for the fact that multiple comparisons were made. The results show that the average frequency of the lights category is statistically different from all other categories. The windshield task category is only statistically different from the lights category, the setup category and the “other” category. The radio and media category is also different from these categories. All other categories are only statistically significantly different from one or more of the first three in the table.

Table 37: Number of total tasks performed per road type and situation

Number of tasks performed per road type and situation					
Variable	**Average**	**Total**	**% of total**	**Average Frequency (per hour)**	**# of trips using the road type**
City	24.98	1499.00	83.46	81.33	60
Rural	2.60	156.00	8.69	46.63	34
Highway	1.20	72.00	4.01	86.75	8
Standing still	0.93	56.00	3.12	NA	NA
Traffic jam	0.22	13.00	0.72	38.50	8

The highway road type has the highest average frequency, followed by the city road type. The average frequency of the rural road type and in traffic jams is nearly half of the frequency of the other two.

Table 38: Linear regression model with perceived HMI car familiarity as the dependent variable

	Perceived HMI car familiarity	
	Coefficient	P-waarde
I am familiar with this car	0.872	<0,001
	Model 1 (N=30)	
R-square	0.463	

In the tables below, the scores on the car familiarity statements are presented for both the familiar car and the unfamiliar car.

Table 39: Car familiarity statements for own car

Variable	**Average**	**Min**	**Max**	**Std dev**	**N**
Statements measuring car familiarity own car (avg. = 7.16)					
Frequency driving the car	9.17	4	10	1.37	30
Feature and option knowledge of the car	8.6	4	10	1.54	30
Button location knowledge of the car	8.97	6	10	1.19	30
Dashboard understanding of the car	9.47	8	10	0.78	30
I am familiar with this car	9.5	7	10	0.78	30
Perceived HMI car familiarity	9.01	6.33	10	1.03	30

Table 40: Car familiarity statements for unfamiliar car

Variable	Average	Min	Max	Std dev	N
Statements measuring car familiarity "unfamiliar car" (avg. = 7.16)					
Frequency driving the car	1.37	1	9	1.47	30
Feature and option knowledge of the car	5.23	1	10	2.36	30
Button location knowledge of the car	6.23	4	10	1.74	30
Dashboard understanding of the car	8.23	5	10	1.45	30
I am familiar with this car	4.73	1	10	2.90	30
Perceived HMI car familiarity	6.57	4	10	1.59	30

The next step is then to compare the average frequencies of different tasks for both trips. In Table 41, some descriptive data for the 10 most performed tasks is presented again. However, this time, the data is shown separately for both trips. This allows for a comparison between the two. For both trips, the average amount of times that a task has been performed is presented, as well as the total amount of times and the average frequency per hour.

Table 41: Top 10 most used tasks for both trips separately

Own car				unfamiliar car			
Task **N = 30**	**Average**	**Total**	**Average frequency (per hour)**	**Task** **N = 30**	**Average**	**Total**	**Average Frequency (per hour)**
Using indicator light	24.17	725	66.22	**Using indicator light**	24.93	748	66.58
Adjusting volume	1.00	29	2.67	**Changing front WW speed**	2.53	76	6.88
Moving sun visor	0.67	20	1.71	**Adjusting volume**	0.63	19	1.64
Adjusting temperature	0.60	18	1.70	**Changing rear WW speed**	0.63	19	1.62
Changing front WW speed	0.53	16	1.56	**Moving sun visor**	0.50	15	1.16
Adjusting fan speed	0.47	14	1.21	**Adjusting inside mirror**	0.23	7	0.66
Enabling/disabling auto pilot	0.27	8	0.89	**Using handbrake**	0.20	6	0.62
Changing rear WW speed	0.23	7	0.77	**Switching radio station**	0.17	5	0.39
Switching radio station	0.13	4	0.42	**Enabling/disabling cruise control**	0.17	5	0.59
Opening/closing windows	0.13	4	0.34	**Adjusting seat**	0.13	4	0.37

*WW = windshield wiper

To test if the differences in average frequencies between the tasks are statistically significant, a repeated measures ANOVA has been performed again. Afterwards, pairwise comparisons have been performed to find out which pairs have statistically significantly different frequencies. The results of these tests are presented for the familiar car and unfamiliar car in Tables 42 and 43 respectively.

Table 42: Repeated measures ANOVA for the means of the frequencies of the 10 tasks in Table 41 (Own car):

Task counts	**F**	**Sig.**
Mauchly’s test of sphericity		<0.001
Greenhouse-Geisser	384.39	<0.001

Task	**Different mean from***
Using indicator light	All tasks
Adjusting volume	Using indicator light, Switching radio, opening/closing windows
Rest	Using indicator light and/or adjusting volume

*Adjustment for multiple comparisons: Bonferroni.

The Greenhouse-Geisser coefficient is used for the repeated measures ANOVA for the familiar car as the data does not satisfy the sphericity characteristic. The repeated ANOVA test shows that at least one of the frequencies is statistically significantly different. Therefore, pairwise comparisons were made and Bonferroni adjustment was used to adjust for the fact that multiple comparisons were made. The results show that the average frequency of the indicator light is different from all other tasks. The average frequency of the volume task is found to be statistically different from the average frequencies of the indicator light, switching radio and opening/closing windows task. The other tasks only showed statistically significant differences with either the indicator light and/or adjusting volume tasks.

Table 43: Repeated measures ANOVA for the means of the frequencies of the 10 tasks in Table 41 (RDW car):

Task counts	**F**	**Sig.**
Mauchly’s test of sphericity		<0.001
Greenhouse-Geisser	214.93	<0.001

Task	**Different mean from***
Using indicator light	All tasks
Rest	Using indicator light

*Adjustment for multiple comparisons: Bonferroni.

Also for the unfamiliar car trips, the greenhouse-geisser coefficient is used. The repeated ANOVA test again shows that at least one of the frequencies is statistically significantly different. The results of the pairwise comparisons show that the average frequency of the indicator light is different from all other tasks. However, the average frequency of these other tasks is only statistically different from the indicator light task for the unfamiliar car.

Table 44 shows the average amounts, total amounts, and average frequency for different aggregated variables. The first three variables show the average total amount of tasks that participants performed during a trip, the total amount of tasks that were performed across all participants and all trips and the total average frequency per hour of all tasks. The first variable includes all tasks, while the second and third variable exclude the indicator light and indicator light plus adjusting the front windshield wiper speed respectively. The same is shown for all different categories of tasks. On the left of the table, this is shown for the familiar car while on the right of the table, it is shown for the unfamiliar car.

For both cars, a repeated measures ANOVA has been performed to see if the differences in average task frequencies are statistically significantly different between the aggregated variables within each car type as mentioned in Table 44. The results of these ANOVAs can be found in Tables 45 and 46.

Table 44: Aggregated descriptive analysis showing the data from both trips separately

Own car				**unfamiliar car**			
Variable N = 30	**Average**	**Total**	**Avg freq (per hour)**	**Variable N = 30**	**Average**	**Total**	**Avg freq (per hour)**
Total amount of tasks used							
Total	28,97	869,00	79,89	**Total**	30,9	927	82,82
Total excl indicator light	4,80	144,00	13,67	**Total excl indicator light**	5,97	179	16,24
Total excl ind and wwspf	4,27	128,00	12,11	**Total excl ind and wwspf**	3,43	103	9,36
Tasks performed per category							
Lights category	24,20	726,00	66,33	**Lights category**	25,03	751	66,92
Windshield category	1,60	48,00	4,50	**Windshield category**	3,77	113	9,94
Radio and media category	1,33	40,00	3,79	**Radio and media category**	1,00	30	2,73
Climate category	1,27	38,00	3,45	**Setup category**	0,37	11	1,03

Table 45: Repeated measures ANOVA for the means of the frequencies of the 4 categories for people's own car:

<table>
<tr><th>Task counts</th><th>F</th><th>Sig.</th></tr>
<tr><td>Mauchly's test of sphericity</td><td></td><td><0.001</td></tr>
<tr><td>Greenhouse-Geisser</td><td>371.33</td><td><0.001</td></tr>
<tr><th>Category</th><th colspan="2">Different mean from*</th></tr>
<tr><td>Lights category</td><td colspan="2">All categories</td></tr>
<tr><td>Rest</td><td colspan="2">Lights category</td></tr>
</table>

*Adjustment for multiple comparisons: Bonferroni.

The greenhouse-Geisser coefficient is used again for the repeated ANOVA and the results indicate that at least one of the average task frequencies is statistically significantly different from the others. To test this further, pairwise comparisons have been made, using Bonferroni as the adjustment method for multiple comparisons. The results show that the average frequency of the lights category statistically significantly differs from all the other categories.

Table 46: Repeated measures ANOVA for the means of the frequencies of the 4 categories for the unfamiliar car:

<table>
<tr><th>Task counts</th><th>F</th><th>Sig.</th></tr>
<tr><td>Mauchly's test of sphericity</td><td></td><td><0.001</td></tr>
<tr><td>Greenhouse-Geisser</td><td>172.85</td><td><0.001</td></tr>
<tr><th>Category</th><th colspan="2">Different mean from*</th></tr>
<tr><td>Lights category</td><td colspan="2">All categories</td></tr>
<tr><td>Rest</td><td colspan="2">Lights category</td></tr>
</table>

*Adjustment for multiple comparisons: Bonferroni.

Also for the unfamiliar car, the results indicate that at least one of the average task frequencies is statistically significantly different from the others. Performing pairwise comparisons shows that only the lights category has a statistically significantly different average task frequency. It is again different from all other categories.

To test if the differences in average task frequencies are statistically significantly different between the two car types, multiple paired t-tests have been performed. The results of these paired t-tests can be found in Table 47. The average task frequencies for both trips are presented, as well as the t-value and p-value for the paired t-test.

Table 47: Paired t-tests for differences in average frequencies

Variable	Own car	unfamiliar car	t-value	P-value
Total	79.89	82.82	-0.82	0.42
Total excl indicator light	13.67	16.24	-0.71	0.49
Total excl indicator light and wwspf	12.11	9.36	1.86	0.07
Lights_tot	66.33	66.92	-0.35	0.73
Windshield_tot	4.50	9.94	-1.32	0.20
Radmedia_tot	3.79	2.73	1.22	0.23
Climate_tot	3.45	0.40	4.32	<0.001
Features_tot	0.89	0.00	1.00	0.33
Other_tot	0.43	0.90	-1.11	0.28
Setup_tot	0.21	1.03	-2.30	0.03
Phone_tot	0.11	0.00	1.00	0.33
Signal_tot	0.10	0.10	NA	NA
Settings_tot	0.08	0.08	-1.00	0.33
Cruise_tot	0.00	0.71	-1.00	0.33

Table 48: Distribution per road type and situation of the number of tasks performed

Category	Highway (%)	Rural (%)	City (%)	Traffic jam (%)	Standstill (%)	N
Lights_tot	3,93	9,34	86,39	0,34	0,00	1477
Windshield_tot	4,35	4,35	80,75	1,24	9,32	161
Radmedia_tot	8,57	1,43	65,71	2,86	21,43	70
Climate_tot	2,33	4,65	62,79	0,00	30,23	43
Features_tot	0,00	50,00	0,00	50,00	0,00	8
Other_tot	0,00	0,00	23,08	0,00	76,92	13
Setup_tot	0,00	7,69	69,23	0,00	23,08	13
Phone_tot	0,00	0,00	100,00	0,00	0,00	1
Signal_tot	0,00	0,00	100,00	0,00	0,00	2
Settings_tot	0,00	0,00	100,00	0,00	0,00	2
Cruise_tot	0,00	50,00	50,00	0,00	0,00	6

The following tables show the statistically significant individual Poisson models that were run and the correlation table for the factors

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	Hypothesis Test df	Hypothesis Test Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-3,131	,2294	-3,581	-2,681	186,249	1	<,001	,044	,028	,068
[Gender=0]	,637	,3244	,001	1,273	3,858	1	,050	1,891	1,001	3,572
[Gender=1]	0^a	.	.	.	.	.	.	1	.	.
(Scale)	1^b									

Dependent Variable: Climate_tot
Model: (Intercept), Gender, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	Hypothesis Test df	Hypothesis Test Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-3,169	,2294	-3,619	-2,720	190,838	1	<,001	,042	,027	,066
[Aut_light=No]	,752	,3244	,116	1,388	5,378	1	,020	2,122	1,124	4,008
[Aut_light=Yes]	0^a	.	.	.	.	.	.	1	.	.
(Scale)	1^b									

Dependent Variable: Climate_tot
Model: (Intercept), Aut_light, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-3,685	,4082	-4,485	-2,885	81,462	1	<.001	,025	,011	,056
[Ada_CC=No]	1,096	,4449	,224	1,968	6,069	1	,014	2,992	1,251	7,156
[Ada_CC=Yes]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

Dependent Variable: Climate_tot
Model: (Intercept), Ada_CC, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-3,685	,4082	-4,485	-2,885	81,462	1	<.001	,025	,011	,056
[Ada_CC=No]	,888	,4935	-,079	1,856	3,240	1	,072	2,431	,924	6,396
[Ada_CC=Yes]	0[a]	.	.	.	.	.	.	1	.	.
[Aut_light=No]	,379	,3599	-,326	1,085	1,112	1	,292	1,462	,722	2,959
[Aut_light=Yes]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

Dependent Variable: Climate_tot
Model: (Intercept), Ada_CC, Aut_light, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-2,520	,1690	-2,851	-2,189	222,264	1	<.001	,080	,058	,112
[Gender=0]	-1,309	,4781	-2,246	-,372	7,492	1	,006	,270	,106	,690
[Gender=1]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

Dependent Variable: Radmedia_tot
Model: (Intercept), Gender, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-3,577	,5000	-4,557	-2,597	51,167	1	<.001	,028	,010	,075
[Weather=Cloudy]	1,008	,5417	-,054	2,070	3,464	1	,063	2,741	,948	7,925
[Weather=Dark]	1,274	,5916	,114	2,433	4,637	1	,031	3,575	1,121	11,399
[Weather=Rainy]	-,129	,7638	-1,626	1,368	,028	1	,866	,879	,197	3,928
[Weather=Sunny]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

Dependent Variable: Radmedia_tot
Model: (Intercept), Weather, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-7,039	1,8698	-10,704	-3,374	14,170	1	<.001	<.001	2,247E-5	,034
HMI	,461	,1993	,070	,851	5,346	1	,021	1,585	1,073	2,343
(Scale)	1[a]									

Dependent Variable: Radmedia_tot
Model: (Intercept), HMI, offset = Log_trip_dur

a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-2,162	,4064	-2,959	-1,365	28,299	1	<.001	,115	,052	,255
Age	-,017	,0102	-,037	,003	2,683	1	,101	,983	,964	1,003
(Scale)	1[a]									

Dependent Variable: Radmedia_tot
Model: (Intercept), Age, offset = Log_trip_dur

a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-5,663	1,7018	-8,998	-2,327	11,072	1	<.001	,003	,000	,098
[Gender=0]	-1,203	,4838	-2,152	-,255	6,186	1	,013	,300	,116	,775
[Gender=1]	0[a]	.	.	.	.	.	.	1	.	.
HMI	,340	,1801	-,013	,693	3,572	1	,059	1,406	,987	2,001
(Scale)	1[b]									

Dependent Variable: Radmedia_tot
Model: (Intercept), Gender, HMI, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-6,826	1,6676	-10,095	-3,558	16,755	1	<.001	,001	4,130E-5	,029
HMI	,327	,1777	-,021	,676	3,396	1	,065	1,387	,979	1,965
Gender	1,186	,4820	,241	2,130	6,052	1	,014	3,273	1,273	8,418
Dark	,464	,3663	-,254	1,182	1,605	1	,205	1,590	,776	3,260
(Scale)	1[a]									

Dependent Variable: Radmedia_tot
Model: (Intercept), HMI, Gender, Dark, offset = Log_trip_dur

a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-1,967	,2236	-2,405	-1,529	77,391	1	<.001	,140	,090	,217
[Weather=Cloudy]	-1,945	,4655	-2,857	-1,033	17,459	1	<.001	,143	,057	,356
[Weather=Dark]	-1,945	,7416	-3,398	-,491	6,878	1	,009	,143	,033	,612
[Weather=Rainy]	,159	,3162	-,461	,779	,252	1	,615	1,172	,631	2,178
[Weather=Sunny]	0[a]	.	.	.	.	.	.	1	.	.
(Scale)	1[b]									

Dependent Variable: Windshield_tot
Model: (Intercept), Weather, offset = Log_trip_dur

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-2,374	,1690	-2,706	-2,043	197,296	1	<.001	,093	,067	,130
Trans_type_dummy	-,727	,3248	-1,364	-,091	5,013	1	,025	,483	,256	,913
(Scale)	1[a]									

Dependent Variable: Windshield_tot
Model: (Intercept), Trans_type_dummy, offset = Log_trip_dur
a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-2,219	,2000	-2,611	-1,827	123,122	1	<.001	,109	,073	,161
Gender	-,721	,2889	-1,287	-,154	6,221	1	,013	,486	,276	,857
(Scale)	1[a]									

Dependent Variable: Windshield_tot
Model: (Intercept), Gender, offset = Log_trip_dur
a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-3,741	,8565	-5,420	-2,062	19,078	1	<.001	,024	,004	,127
HMI	,243	,0927	,062	,425	6,889	1	,009	1,276	1,064	1,530
(Scale)	1[a]									

Dependent Variable: Total_excl_indicator
Model: (Intercept), HMI, offset = Log_trip_dur
a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-1,736	,1447	-2,019	-1,452	143,912	1	<.001	,176	,133	,234
Driv_fre	,026	,0141	-,002	,053	3,344	1	,067	1,026	,998	1,055
(Scale)	1[a]									

Dependent Variable: Total_excl_indicator
Model: (Intercept), Driv_fre, offset = Log_trip_dur
a. Fixed at the displayed value.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.	Exp(B)	95% Wald Confidence Interval for Exp(B) Lower	95% Wald Confidence Interval for Exp(B) Upper
(Intercept)	-3,583	,8865	-5,321	-1,845	16,335	1	<.001	,028	,005	,158
HMI	,217	,1011	,019	,415	4,596	1	,032	1,242	1,019	1,514
Driv_fre	,011	,0160	-,021	,042	,442	1	,506	1,011	,980	1,043
(Scale)	1[a]									

Dependent Variable: Total_excl_indicator
Model: (Intercept), HMI, Driv_fre, offset = Log_trip_dur
a. Fixed at the displayed value.

Correlations

		Age	Gender	Driv_lic	Driv_fre	HMI	BSC_type_dummy	Trans_type_dummy	Ada_clim_dummy	Ada_cc_dummy	Aut_light_dummy	Rush_dummy	Sunny	Cloudy	Rainy	Dark
Age	Pearson Correlation	1	,074	,992**	,291	,090	-,004	,089	,039	,229	,178	,243	,053	-,187	,089	,096
	Sig. (2-tailed)		,699	<,001	,119	,637	,982	,638	,840	,224	,347	,196	,780	,323	,640	,614
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Gender	Pearson Correlation	,074	1	,073	,225	,125	,250	,190	,196	,098	,000	-,111	,000	-,095	,000	,126
	Sig. (2-tailed)	,699		,700	,233	,512	,183	,314	,300	,607	1,000	,558	1,000	,617	1,000	,505
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Driv_lic	Pearson Correlation	,992**	,073	1	,285	,146	,006	,075	,010	,257	,200	,199	,052	-,147	,073	,061
	Sig. (2-tailed)	<,001	,700		,127	,441	,975	,694	,957	,170	,290	,292	,786	,438	,702	,747
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Driv_fre	Pearson Correlation	,291	,225	,285	1	,408*	,330	,104	,361	,182	,178	-,004	,118	-,109	,056	-,042
	Sig. (2-tailed)	,119	,233	,127		,025	,075	,583	,050	,335	,346	,982	,533	,565	,769	,826
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
HMI	Pearson Correlation	,090	,125	,146	,408*	1	,148	-,010	,031	-,100	,079	-,136	-,171	,013	,105	,054
	Sig. (2-tailed)	,637	,512	,441	,025		,435	,960	,870	,600	,677	,473	,367	,947	,582	,776
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
BSC_type_dummy	Pearson Correlation	-,004	,250	,006	,330	,148	1	,190	,489**	,391*	,617**	,056	,000	-,095	,177	-,063
	Sig. (2-tailed)	,982	,183	,975	,075	,435		,314	,006	,032	<,001	,770	1,000	,617	,350	,740
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Trans_type_dummy	Pearson Correlation	,089	,190	,075	,104	-,010	,190	1	,386*	,451*	,279	-,164	,235	,050	-,101	-,211
	Sig. (2-tailed)	,638	,314	,694	,583	,960	,314		,035	,012	,136	,385	,210	,794	,596	,264
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Ada_clim_dummy	Pearson Correlation	,039	,196	,010	,361	,031	,489**	,386*	1	,435*	,558**	,093	,035	-,312	,208	,155
	Sig. (2-tailed)	,840	,300	,957	,050	,870	,006	,035		,016	,001	,626	,856	,094	,271	,414
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Ada_cc_dummy	Pearson Correlation	,229	,098	,257	,182	-,100	,391*	,451*	,435*	1	,498**	,071	,138	,172	-,208	-,155
	Sig. (2-tailed)	,224	,607	,170	,335	,600	,032	,012	,016		,005	,710	,466	,363	,271	,414
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Aut_light_dummy	Pearson Correlation	,178	,000	,200	,178	,079	,617**	,279	,558**	,498**	1	,017	-,218	-,015	,327	-,098
	Sig. (2-tailed)	,347	1,000	,290	,346	,677	<,001	,136	,001	,005		,928	,247	,939	,077	,608
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Rush_dummy	Pearson Correlation	,243	-,111	,199	-,004	-,136	,056	-,164	,093	,071	,017	1	-,276	-,164	-,079	,599**
	Sig. (2-tailed)	,196	,558	,292	,982	,473	,770	,385	,626	,710	,928		,140	,385	,679	<,001
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Sunny	Pearson Correlation	,053	,000	,052	,118	-,171	,000	,235	,035	,138	-,218	-,276	1	-,437*	-,250	-,224
	Sig. (2-tailed)	,780	1,000	,786	,533	,367	1,000	,210	,856	,466	,247	,140		,016	,183	,235
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Cloudy	Pearson Correlation	-,187	-,095	-,147	-,109	,013	-,095	,050	-,312	,172	-,015	-,164	-,437*	1	-,437*	-,391*
	Sig. (2-tailed)	,323	,617	,438	,565	,947	,617	,794	,094	,363	,939	,385	,016		,016	,033
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Rainy	Pearson Correlation	,089	,000	,073	,056	,105	,177	-,101	,208	-,208	,327	-,079	-,250	-,437*	1	-,224
	Sig. (2-tailed)	,640	1,000	,702	,769	,582	,350	,596	,271	,271	,077	,679	,183	,016		,235
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Dark	Pearson Correlation	,096	,126	,061	-,042	,054	-,063	-,211	,155	-,155	-,098	,599**	-,224	-,391*	-,224	1
	Sig. (2-tailed)	,614	,505	,747	,826	,776	,740	,264	,414	,414	,608	<,001	,235	,033	,235	
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix F: Flyer and informed consent form

Instructions on informed consent form

You are being invited to participate in a Master's thesis research. This study is being done in collaboration with the TU Delft and RDW.

The purpose of this research study is to better understand driving behaviour and will take you approximately 60 minutes to complete. The data will be used for research purposes and can be used by RDW as input for regulations. We will be asking you to answer a few short questions and drive two trips on a route that is familiar to you. One trip will be with your own car and the other will be with an unfamiliar car provided by the researcher. The car that is provided to you is fully insured by RDW and will not be a financial risk to the participants. During the trips, we will ask you to act as if you are alone in the car and drive as you normally would.

To the best of our ability, your answers in this study will remain confidential. We will minimize any risks by making sure that the data is completely anonymous. No privacy-sensitive information or information that can result in a high risk of identification of the participant will be stored. The data will only include non-privacy-sensitive information such as age and time since acquiring a driver's license. Some data regarding characteristics of the route that is driven and the car that is used will be collected but the origin and destination of the route will not be mentioned such that the route cannot be identified. The researcher will collect data only through written notes which will later be entered digitally. The collected data will be used for research purposes and will be shared with RDW.

Your participation in this study is entirely voluntary and you can withdraw at any time. You are free to omit any questions.

If you have any questions, feel free to contact the researchers via the following details:

Daniël Auerbach (Dauerbach@rdw.nl)
Marjan Hagenzieker (m.p.hagenzieker@tudelft.nl)

INFORMED CONSENT FORM (PLEASE TICK THE APPROPRIATE BOXES)	Yes	No
A: GENERAL AGREEMENT – RESEARCH GOALS, PARTICPANT TASKS AND VOLUNTARY PARTICIPATION		
1. I have read and understood the study information dated [28/*11/2023*], or it has been read to me. I have been able to ask questions about the study and my questions have been answered to my satisfaction.	☐	☐
2. I consent voluntarily to be a participant in this study and understand that I can refuse to answer questions and I can withdraw from the study at any time during the observations, without having to give a reason.	☐	☐
3. I understand that taking part in the study involves: • Answering some short questions. • Driving two trips on a familiar route. One with a familiar car and one with an unfamiliar car. • The researcher collecting data during driving by making notes.	☒	☐
4. I understand that in the event of any incidents with my own car, I am required to use my own insurance coverage.	☐	☐
5. I understand that any fines that occur while I am driving are at my own expense.	☐	☐
B: RESEARCH PUBLICATION, DISSEMINATION AND APPLICATION		
6. I understand that after the research study, the de-identified information I provide will be used for research purposes and can be used as input for regulations.	☐	☐
C: (LONGTERM) DATA STORAGE, ACCESS AND REUSE		
7. I give permission for the de-identified data that I provide to be archived on the hard drive of the researcher, a RDW-secured location and the TU Delft thesis repository so it can be used for future research, learning and creating regulations.	☐	☐

Signatures

________________________ ________________________ ________

Name of participant [printed] Signature Date

I, as researcher, have accurately read out the information or presented the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

________________________ ________________________ ________

Researcher name [printed] Signature Date

TU Delft RDW

Participants Needed for Research Study on Driving Behavior.

SCAN MIJ

Do you want to contribute to science by participating in a TU Delft and RDW research project while getting compensation? Please let us know by scanning the QR code and booking a time slot (Personal details will not be shared). We will contact you with more information!

You May Qualify If You

- Are older than 18 years of age
- Own a driving license
- Have a car available that you can drive in

Participant benefit

Participants in the study will be compensated through a 10-euro gift card.

The data we collect will be fully anonymous

Participation includes: (10-euro gift card)

- Driving a familiar route with us in your available car
- Driving a familiar route with us in an RDW car (FULLY INSURED)
- Maximum of 60 minutes of driving total
- Answering a few short questions

Location: We will come to you!

FOR MORE INFORMATION:

Contact us at Dauerbach@RDW.nl

Time slot link:
https://calendar.app.google/5jzyCyrGaPoW7ENP7

RDW, VRT TI | Europaweg 205, 2711 ER Zoetermeer