MASTER

Evaluating the Impact of AI-based Human Machine Interfaces in comparison with conventional User Interfaces in Autonomous Vehicles

Suresh, A.

Award date:
2021

Department of Mechanical Engineering
Masters in Automotive Technology
User Experience Design - Automotive Human Factors Research Group

Master Thesis

Evaluating the Impact of AI-based Human Machine Interfaces in comparison with conventional User Interfaces in Autonomous Vehicles

Abhishek Suresh

Supervisors

dr. rer. nat. Bastian Pfleging (*Supervisor and Committee Member*)
dr. Frederik Diederichs, Fraunhofer IAO (*Advisor*)

Committee Members

prof. dr. ir. J.H. Eggen (*Committee Chair*)
dr. Ion Barosan (*Committee Member*)

Thursday 21st January, 2021

Declaration concerning the TU/e Code of Scientific Conduct for the Master's thesis

I have read the TU/e Code of Scientific Conduct[i].

I hereby declare that my Master's thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Conduct

Date

11/01/2021

Name

Abhishek Suresh

ID-number

1327755

Signature

Submit the signed declaration to the student administration of your department.

[i] See: https://www.tue.nl/en/our-university/about-the-university/organization/integrity/scientific-integrity/
The Netherlands Code of Conduct for Scientific Integrity, endorsed by 6 umbrella organizations, including the VSNU, can be found here also. More information about scientific integrity is published on the websites of TU/e and VSNU

Abstract

Since more and more vehicles are becoming increasingly connected, Artificial Intelligence based technologies play a significant role in improving the occupant's interaction with the vehicle, in addition to controlling driving-related activities. Smart voice assistants, in-vehicle infotainment systems, driver/occupant monitoring, safety, and many more will become a major driving for the development of AI-based technologies, which leads to increased efficiency and reduced human error. During recent years, the potential of AI is being utilized in autonomous vehicles with a focus on improving all driving and non-driving related activities, while minimizing the potential of human error. In addition to this, the implementation of AI will also be a key factor in optimizing the relationship between the driver and the occupants of the vehicle. One way of achieving this is by providing intuitive interaction and predictable behavior through in-vehicle information systems. Thus, the key goal of this master thesis project is to understand the scale of the impact of AI's implementation of Automotive Human Machine Interfaces. For this project, the systematic differences between various HMI's of the autonomous vehicle, with and without the implementation of AI, are being compared, and the user experience and acceptance of the occupants will be analyzed. Studying the user acceptance regarding the usage of AI-based HMI's is a key factor for this research, based on which the potential benefits and the corresponding caveats of implementing AI in Automotive HMI can be found and communicated.

Preface

This master thesis is the result of the research carried out at Eindhoven University of Technology, Eindhoven, Netherlands, in collaboration with the Fraunhofer Institute of Industrial Engineering (IAO), Stuttgart, Germany. The external supervisor at Fraunhofer serves as an advisor by providing the research problem and progress feedback, while the main supervision will happen internally within Eindhoven University of Technology (TU/e). Due to the COVID19 pandemic that hit in 2020, it was mutually decided between the two parties that the research is going to be conducted online, and all the scheduled meetings will now take place virtually. Even though this is the case, all of the steps and decision that went into the development of this thesis was not taken in isolation, and were influenced by the valuable feedback and discussions that took place with various researchers, professors, and students, across both the institutions on the whole. Due to this reason, the thesis was written using a third-person's point of view. The publications, guidelines and taxonomies that was used for building this thesis is provided at the end of this manuscript.

Acknowledgements

One quote that comes to my mind as I finish my master thesis is, **"what a year 2020 has been"**. Ever since the COVID19 pandemic started to take force in the early stages of 2020, all the thoughts and doubts came to my mind regarding whether I will finish my master thesis or not. Lock downs were imposed nationwide, universities shut, not being able to travel, especially to my home country of India, these thoughts kept on lingering in my mind for a long time. Now, as I reflect upon the experiences I had in the past year, I never would have thought finishing my master thesis in these dire times, even though being conducted all my study online. For this, first of all I would like to express my heartfelt gratitude and thanks to all, who risked their lives to keep the pandemic contained. In addition to this, without my university's swift action of changing the entire structure of education completely online, this feat would not be possible. Their constant updates about the pandemic, and the supports they provided, kept the flow of my thesis project smooth and without any technical difficulties. I would also like to thank my academic advisor, **Paula Verbeek**. She was one I approached whenever I faced problems during my entire master thesis program, and without her suggestions, I would have been extremely difficult for me to navigate through.

Without this master thesis, I would have been in the same state now, as I was in the beginning of this year, and for this I would like to express my heartfelt gratitude to **Frederik Diederichs** of Fraunhofer IAO, for providing me this wonderful opportunity to pursue my master this thesis in this topic. It was your belief and confidence in my abilities, made to navigate through my thesis phase and perform the intended tasks as effective as I can. It would have been nicer meeting you in person, but the innumerable conversations we had during our meetings felt as if it was happening in-person. I would also like to thank **Valeria Bopp-Bertenbreiter** of Fraunhofer IAO. Same as with Frederik, I really enjoyed conversing with you about my thesis, and your valuable feedback about my thesis made to me find out the points that I was searching for.

Also, I would like to mention another person, without whom I would not be here at this moment, my mentor, **Bastian Pfleging**. A simple thanks would not be enough when compared to the amount of insights and feedback he gave during my masters course. It was his expectations of wanting more, pushed to me to my limits during the masters course, which in turn elevated my mind regarding my specialization. I would also like to thank my esteemed committee members **Berry Eggen** and **Ion Barosan** for first accepting to be in my graduation committee and for taking your time in going through my thesis, and in addition, I would also like to extend my thanks to **Panos Markopoulos**. In addition to Bastian, the three of you have been a part of my masters journey and your teachings, insights and support will forever linger in my mind.

I would also like thank our department's doctoral researchers, **Debargha Dey**, **Melanie Berger**, and also **Christian Kunz** from Faunhofer IAO, for their valuable feedback and suggestions for improving the design and experiment of this thesis. Also, I would like to thank all of my colleagues and friends my specialization and masters course, who was there me to my any doubts I had and also provided me valuable inputs for the tasks that I was stuck to proceed. One thing I was very concerned during the pandemic was being alone and had no person to interact with, but thanks to my friends from India and from the ones I made during my masters course, I never felt alone even though being separated by a larger distance. I would also like to express my thanks to my bachelor professors, especially to my principal from my bachelors university, **K. Palanikumar** and to my mentor during my bachelors, **Ashwin Sailesh**, who constantly motivated me from the start of my bachelors till now. Without them and their guidance and motivation, I would not have been here in the first place.

Finally, none of this would have been possible, without the love, support, and belief, my parents **Suresh Krishnamurthy** and **Lakshmi Suresh**, had in me. Even though being separated from each other during this pandemic, their constant calls to me, felt as if I was in my hometown and it was their support and motivation that propelled me to reach this stage and will propel me to even greater heights. The automotive technology master course was indeed a new and surreal experience for me, and I would like to thank each and every one of them involved in this experience. At last, I would like to express thanks to everyone and my sincerest apologies to everyone that I have to failed to mention in this acknowledgement. As I finish the conclusion of this chapter in my book, I look forward to the paths and events that I encounter in the next chapter of my life and I am significantly excited for it. Once again from the bottom of my heart, thank you all, stay healthy and stay safe!!

Contents

List of Figures

List of Tables

List of Abbreviations

AI	-	Artificial Intelligence
AV	-	Autonomous Vehicles
BASt	-	Bundesanstalt für Straßenwesen
CES	-	Consumer Electronics Show
CTAM	-	Car Technology Acceptance Model
DARPA	-	Defense Advanced Research Projects Agency
EU	-	European Union
GDPR	-	General Data Protection Regulation
GPS	-	Global Positioning System
HMI	-	Human Machine Interface
ISO	-	International Standard Organization
IC	-	Instrument Cluster
IT	-	Information Technology
IVIS	-	In-Vehicle Infotainment System
JAMA	-	Japan Automobile Manufacturers Association
meCUE	-	Components of User Experience (modular evaluation)
MBUX	-	Mercedes Benz User Experience
ML	-	Machine Learning
MMI	-	Man Machine Interface
NASA	-	National Aeronautics and Space Administration
NHTSA	-	National Highway Traffic Safety Administration
POI	-	Point of Interest
SAE	-	Society of Automotive Engineers
SOTA	-	Software-Over-The-Air
TAM	-	Technology Acceptance Model
TLX	-	Task Load Index
TOR	-	Take-Over-Request
UEQ	-	User Experience Questionnaire
UI	-	User Interface
UTAUT	-	Unified Theory of Acceptance and Use of Technology
UX	-	User Experience
VR	-	Virtual Reality

Chapter 1

Introduction

Since vehicles are becoming connected, AI-based technologies play a significant role in improving the occupant's interaction with the vehicle, in addition to controlling driving-related activities. Smart voice assistants, in-vehicle infotainment systems, driver/occupant monitoring, safety, and many more will become a major driver for the development of AI-based technologies, which will lead towards increased efficiency and reduced human error [101]. During recent years, the potential of artificial intelligence (AI) is being utilized in autonomous vehicles with a focus on improving all driving and non-driving related activities, while minimizing the potential of human error [10]. In addition to this, the implementation of AI is expected to be a key factor in optimizing the relationship between the driver and the occupants of the vehicle. One of the potential way of achieving this is by providing intuitive interaction and predictable behavior through in-vehicle information systems.

Recently, almost all of the companies, especially the ones from the automotive industry, consider driving automation as a major driving force for the development of future cars, that can improve the driver's work efficiency and driving safety. As a result of this, the application of AI and machine learning algorithms in automotive industry is on the rise as these are the enabling technologies for driving automation, but also due to the shift in the mindset of the buyers to consider technology as a key factor, instead of considering how much horse power and torque the power unit can offer. Some of the user-expected technologies currently within a car include vehicle navigation and collision avoidance systems, automated vehicle guidance and braking, and the functionality of these technologies are achieved by incorporating several cameras and sensors that are capable of gathering accurate data from the vehicle surroundings.

Since the data compiled by the cameras and sensors are to be processed in real-time, autonomous vehicles require a very high level of computing software and algorithms to achieve this task, and this is where AI can provide its services [101]. This justifies the consideration of technologies as a key factor along with others for determining the purchase of vehicles, and therefore using AI for enhancing the driving and in-vehicle experiences will play a key role in transitioning towards the higher levels of autonomy (*SAE Level 3 and higher*) [22]. Ever since now, the topic of discussion regarding the application of AI in vehicles revolved around the topic of enabling self-driving technology. However, the fact that the much broader foundation of how AI has a disruptive impact on the automotive industry, has been overlooked until now.

1.1 AI and its influence in the Automotive Industry

Because of AI's rapid growth, many companies are beginning to identify the importance of integrating AI and machine learning into their organization and collecting large amounts of data for processing, one common method is through supervised learning [79]. According to the analysis done by Rick's cloud[1], the incorporation of AI within the automotive industry is expected to rise to 109% by 2025. In addition to this report, this trend is also being showcased in the study done by Deloitte[2] in their Moonshot project. According to this report, the growth rate for AI application in autonomous vehicles will increase, particularly in the deployment of software and services, followed by hardware implementation [33]. As a result, incorporation of AI is causing some disruptions within the automotive industry, some of the use cases are showcased in the following figure 1.1.

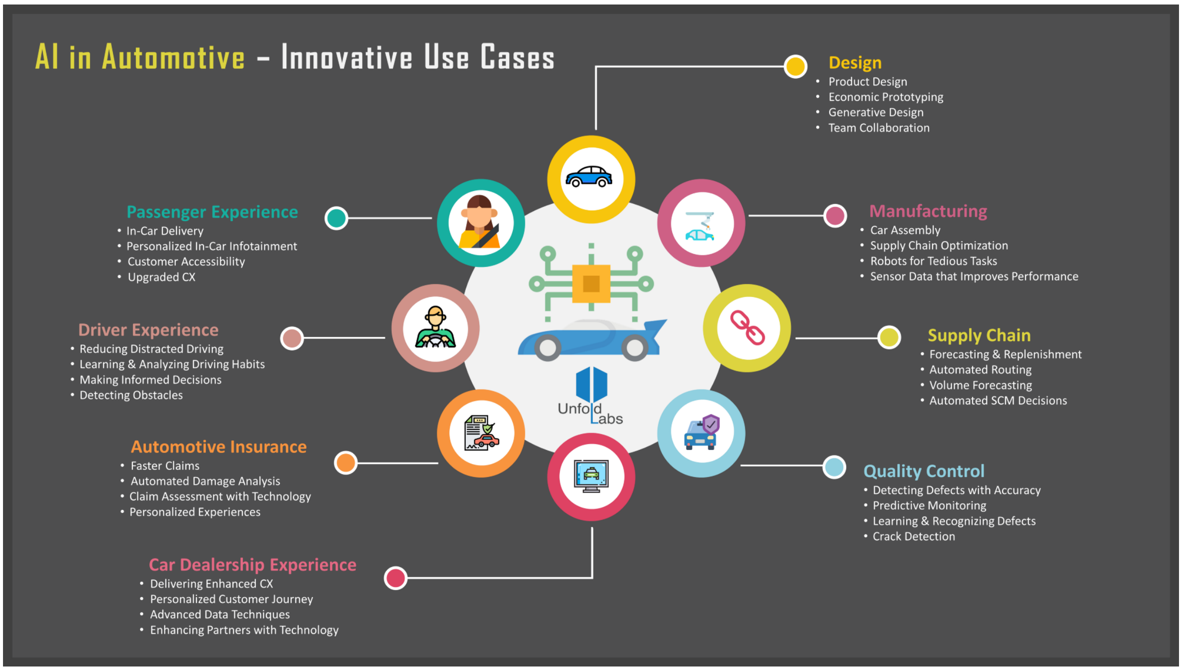


Figure 1.1: Disruptive Use-cases of AI
Source - https://bit.ly/3hStUn0

Almost all of the major car manufacturers like Daimler with their Mercedes Benz User Experience (*MBUX*), Hyundai with their *Smart Sense*, Audi's Man Machine Interface (MMI) Virtual Cockpit, Mitsubishi's *Maisart AI Technology*, and many more, in addition to the involvement of the tech companies such as Microsoft, *Watson* by IBM, Google, Samsung[3], LG[4], Nvidia, are realizing AI's impact on the services offered and are transitioning towards the development and nurturing of AI[5]. Furthermore, based upon the patent research done by Benčić et al [10], the recent trends of AI infusion in the automobile industry is researched

[1] https://rickscloud.com/intelligent-cars-ai-and-the-automotive-industry/
[2] https://bit.ly/3pBRt6I
[3] https://news.samsung.com/us/tag/digital-cockpit/
[4] https://bit.ly/2NqdBQh
[5] https://bit.ly/3fTje5E

and analyzed, where the authors classified the portfolio based upon the image processing and classification, vehicle routing and navigation, Human-Machine (or) Man-Machine Interfaces (HMI/MMI), and monitoring and control of vehicle systems, which is in-line towards the focus of improving the experience of the occupants of the vehicle. From the research, it was found that the large number of patents that were filed related to AI until 2016, belongs to the HMI/MMI systems, which largely discussed the data collection, transmission, and presentation techniques, and also due to the provision of an immediate way of providing time-critical information to the users [50, 104]. The automotive industry sudden interest in the incorporation of AI in their fleet will enable them to investigate its usage in other fields also, which was explained in the previous section, and also evident in the figure 1.1.

1.2 Gaps and Limitations

However, these AI-based patents are filed in such a way that they are not intended for the usage in automobiles exclusively. Instead, these patents also extend AI's functionality to other fields which can also incorporate some features that can also be applied for a vehicle. The reverse case is also possible since almost half of the companies that filed for patents in automotive AI are from IT companies [10]. In addition to this, due to the involvement of many companies for pushing many AI pilot projects, implementation of it in many user interfaces will lead to encountering some potential problems, and also the incorporated AI system is not universal which in-turn leads to compatibility issues.

Furthermore, there are major concerns relating to AI's security, privacy, and ethical issues. These highly uncertain behaviors of AI-incorporated system will lead to aggressive information hiding and also not being in-sync with the user's preferences [46]. Thus, it is essential to consider some barriers that may arise when incorporating AI in-order to provide an *universal and systematic framework* for AI incorporation in automotive user interfaces. Some of the barriers faced while incorporation of AI in HMI and some potential solutions are elicited below.

1. AI algorithms for HMI design needs to focus on ***transparency and explanation*** [83]. A balance should be laid between briefing the user of AI's operations and level of control it provides to the user. Sometimes, AI can make the system more predictable by handing the control to user more frequently, which makes system more transparent and less adaptable [34]. One potential solution is to allow the user to change the model fitting their needs, and should enable the user to acquaint with HMI's features through the provision of tutorials [46].

2. The AI system must be in a condition to explain the reason regarding why did it perform the tasks and how it did those tasks. This reasoning of AI is based on explanation generation and visualization techniques that enables the AI to build up its memory by accepting the data from the user. But here arises the potential problem of ***privacy and trust*** [96]. Sometimes, the system might force the user to provide data to the system for improving the personalization, which the user was unwilling to provide. A potential fix for this particular *privacy* issue is to decide which user information has to be made public and private. When *trust* is taken into consideration, it is essential that the HMI needs to prove the user that its adaptive behavior does in-fact improve the

interaction with the user. To research further in the field of trust with AI, it might be feasible to look into the AI's learning process and information filtering methods to find relevant information.

3. Also one of the potential barriers that AI-incorporate HMI face is that the AI should invest in more of a ***long-term interaction*** with the user rather than a short one, (.i.e) *training the algorithm* [66]. Systems should be designed to fit-in with the user's existing and desired work practices so that it can provide feedback to the user after the desired interaction and should determine which kind of feedback should fit the user's needs.

AI-incorporated HMI that are designed to address the problems that are stated above, was still found to violate the traditional principles of usability design [58], which will lead to the AI being inconsistent by reacting differently to the user commands. This causes confusion among users to decide whether to trust the system or not. Hence, it is of utmost importance to design the system that people can trust to engage with them effectively. As the increasing automation of the vehicles, affects the roles and expectation of the driver, it is also essential for the vehicle HMI, both AI and non-AI based, to be consistent across all the platforms, that addresses the above-mentioned issues, and also can allow them to be configured by the occupants, thereby facilitating towards an effective and seamless interaction between the occupants and the car [53].

1.3 Scope of the Thesis

The prime objective of this master thesis project is to understand *what impact does the introduction of AI will have on Automotive Human Machine Interfaces.* For this project, the goal is to explore the systematic differences between various automotive HMIs with and without the implementation of AI, by comparing the interaction choices offered by them and analyzing how they affect user experience and acceptance of the occupants. Studying the user acceptance regarding the usage of AI-based HMIs is a key factor for this research based on which the potential benefits and the corresponding caveats of implementing AI in Automotive HMI can be found and communicated. Since the major focus of this master thesis project is being laid on analyzing the acceptance and benefits of AI-incorporated automotive HMI systems, it is essential to provide a clear and systematic framework for conveying the differences between AI based and non-AI based HMI, leading towards a *consistent implementation.* Thus, based on the information addressed above, the research topic of this is coined as:

"To identify the systematic differences in the interaction choices offered by the Human Machine Interface (HMI) of the level 5 autonomous car incorporated with Artificial Intelligence (AI) over the conventional user interfaces, which can significantly improve the user experience and acceptance of the occupants during interaction"

1.4 Research Questions

In order to identify and analyze the impact that AI has on improving the interactions in automotive user interfaces, a detailed step-by-step approach is taken where, the comparison of the influence of Automotive HMI systems with and without AI has to done first,subsequently identifying systematic differences between them, and to analyze the user experience and acceptance of AI-based automotive user interfaces by testing and evaluating these HMI systems in vehicles with and without AI, and highlighting the benefits of AI-based HMI over the conventional user interfaces. Thus, owing to the thesis research topic mentioned above, the research questions that aims to address this has to be coined based on the details addressed above, and the resulting research questions are given below.

1. How does an AI-based HMI of an autonomous vehicle which allows the occupants to interact and customize its interface result in an improved user experience and acceptance of the system?

2. How well do the occupants of an autonomous vehicle understand and anticipate the behavior of the system based on the context that was presented to them?

As this project identifies the systematic differences between AI and non-AI based HMI and communicating AI based HMIs benefits and caveats to the public, it is essential that the comparison and testing of prototype interfaces should be targeted towards the user, and should also involve the user during the design process. This is critical for analyzing the acceptance and user experience of the interface prototype, which significantly leads to communicating the benefits of AI based HMI. So, to facilitate the research moving towards the direction of user-centered design, the key activities to be performed during this master thesis project [87] is by arriving at the interaction logic concepts for the AI-based HMIs first using UI/UX rules, which in turn leads to the creation of interactive prototypes and wireframes. Once done, the models are built and/or extended which are then deployed for user study and evaluation.

1.5 Research Methodology

Until now, there are multitude of UX methods to chose from, for evaluating the prototype to measure it's user experience and acceptance. The UX evaluation methods that are chosen for constructing the testing of the prototype according to this project context, is elaborated below in order.

1. Identifying all the research studies about the research question, and evaluating and analyzing them is done through a ***systematic review approach*** [39]. This approach is effective for summarizing the systematic differences that are to be identified. Also through this approach, any gaps within the area of the current research can be identified. Since this project prioritizes theoretical approach more, it is essential to manage the collected literature using references managers like ***mendeley***.

2. Another key factor for improving the user experience is to understand who the users are and what are they trying to accomplish while using this interface. Thus, it is essential to identify the user's needs and expectations, which can be done by eliciting ***user requirements*** by creating user profiles, contextual inquiry and even task analysis.

3. Many types of user research methods are available for the development for the prototype. In addition to the systematic review approach, the current trends for this project is analyzed by conducting surveys and market research, followed by gaining insights from the user through ***brainstorming sessions*** and by interviewing ***focus groups***.

4. For an interface to enable a seamless interaction between the user and the HMI, it is essential first to refine it to a position that it can be suitable for implementation and deployment. Thus, testing of the interactive interfaces is done by creation of online mock-ups and wireframes using ***quick and dirty prototyping***, and subsequently integrating it in an online testing tool for performing user-centered studies.

5. Finding the usability problems that are associated with the exiting interfaces by involving experts to review the interface based on well established usability heuristic principles, which in turn is called as ***heuristic evaluation*** [58].

6. Before deploying the final prototype to the eyes of the public, it is essential first to understand how the user interacts with it by performing a ***usability test***. By conducting a usability test, some of the design problems within the prototype is identified, along with potential improvement solutions, and learning about the target user's behavior and preferences. By considering the project boundaries, it is feasible to perform a ***remote usability testing***[6] by leveraging online communities [40].

Risks and Foreseeable Side Effects

There are no foreseeable risks, discomforts, or inconveniences associated with participating in this experiment. Interviews, observations, and studies conducted, will be focused exclusively on the usage and experience of using the prototype(s). Additionally, the tasks that will be asked from participants in order to structure evaluations and tests, will be in such way that they do not deviate from regular activities in the specified context, and will not involve any complex interactions that might cause some discomfort. Also, the variables provided in the questionnaire includes only low risk information, with results being only going to be presented in an aggregated form. Participants will not be exploited, and none of their data would place participants at risk of criminal or civil liability.

Data Protection and Storage

The collected data will be coded and allocated a randomized number. The coded data will be stored locally and temporarily on the password-protected devices of the researchers. Long-term storage will happen on a password protected institutional repository at the Eindhoven University of Technology or a provider which has data protection clearance by TU/e (e.g. SurfDrive/ResearchDrive). The duration is, as outlined in guidelines for good academic practice (e.g., Netherlands Code of Conduct for Research Integrity: 10 years). The audio/video recordings taken during this study, will be stored in the local password protected devices of the researchers for 5 years and in a long-term storage (as stated above) for 10 years.

Data Confidentiality

All personal data collected during the study will be processed confidentially and test subjects will never be recognizable in publications, academic material or any other mean. For illustrations in publications and other educational/academic material the researchers will substitute the participants in pictures / videos. As an alternative, a separate, explicit consent (Article 6, paragraph 1, point a EU GDPR) will be asked from (selected) participants. In case of third

[6] `https://www.nngroup.com/articles/unmoderated-user-testing-tools/`

parties being involved in the research and analysis process, a data privacy / processing agreement will be made before data is shared. The consent form will include asking for consent to reuse the recorded data also for future investigations (GDPR Art.13(3),Rec. 32 and WP259 rev.01pg.10). All people involved in the experiment will be instructed on the importance of data privacy and security, to maintain confidentiality, and are required to follow procedures as outlined for instance in the EU GDPR regulations.

For simulator and real-world studies, we will explicitly instruct and remind the participants that they can withdraw at any point in time and urge them to immediately quit in case they experience any sign of motion sickness. Finally, no individual results will be published, as conclusions will be made from the entire cohort's data. The results of this study will be disseminated in scientific conferences and published in conference proceedings, scientific research journals, project reports, student theses, and standard press and social media (advertising the actual research papers).

1.6 Evaluation

With the main thesis goal of analyzing the impact that the AI has on automotive HMI, the question of how to measure the user experience and acceptance was imminent. As Don Norman and Jakob Nielsen defines, "User experience" encompasses all aspects of the end-user's interaction with the company, its services, and its products. So, more emphasis should be laid on creating products, in this case, automotive user interfaces, that provide meaningful and relevant experiences to users, and to subsequently evaluate these products to determine the user's perception while using the product. For measuring the user experience of the product, there exists a multitude of methods and standardized questionnaires, which consists of series of invariable group of questions relating to the product's qualities where the participants are expected to provide answers to. Among the standardized questionnaires that are available for measuring the user experience, the three most recognized questionnaires[7] are given below.

1: Attrakdiff

Developed by Marc Hassenzahl, the Attrakdiff survey[8] is used to measure the pragmatic and hedonic qualities of the product, which consists of 28, seven-step items whose poles are opposite adjectives to each other,used for measuring pragmatic usability, hedonic simulation, hedonic identification and attractiveness. These seven-step response items are developed based on a 7 point likert scale ranging from -3 to +3, with 0 as a neutral score.

2: User Experience Questionnaire (UEQ)

Like the Attrakdiff questionnaire, the User Experience Questionnaire[9] (UEQ) developed by Laugwitz et al. [43], is also used to measure the pragmatic and hedonic quality of the product, using 26 items rated on a 7 point likert scale. In UEQ, a more comprehensive immersion of user experience is measured, which addresses both the classical usability aspects (efficiency, perspicuity, dependability) and user experience aspects (originality, stimulation).

[7] https://measuringu.com/pragmatic-hedonic/
[8] http://www.attrakdiff.de/sience-en.html
[9] https://www.ueq-online.org/

3: Components of User Experience (modular evaluation) meCUE questionnaire

The biggest of the other two questionnaires, meCUE is a standardized scale developed based on Thüring and Mahlke's Components of User Experience model. Compared to Attrakdiff and UEQ, the meCUE questionnaire comprises of 33 items with five separately validated modules that relate to the perception of different product characteristics (usefulness, usability, visual aesthetics, status, commitment), to users emotions (both positive and negative emotions) and to consequences (product loyalty and intention to use). The fifth module allows for a global assessment of the product[10]. Ratings for the items are recorded on a 7 point agreement scale ranging from strongly disagree to strongly agree.

Even though the use of UEQ as a standardized questionnaire for evaluating the user experience is on the rise since 2017 [18], for this thesis, Attrakdiff is chosen as the standardized questionnaire for evaluating the user experience, since it was the most used till now compared to the other two questionnaires, and also due to being more reliable than UEQ and meCUE [42]. In addition to being categorizing the product as being useful and usable, the product needs to be accepted by the users for daily use. So, the user's motivation to use the product, is measured using the ***Technology Acceptance Model (TAM)*** proposed by Fred Davis [15], where the external variables that affect the user's motivation such as usefulness and ease of use, are used a construct for determining the user's attitude and intention to use to the product. However, this iteration of the model only moderately addressed the mental effort of the user's, Venkatesh extended the exiting model to ***TAM2*** [97] and ***TAM3*** [98] by adding more variables to the existing constructs like subjective norm, job relevance, output quality, and results demonstrability. All the above 3 models in addition to ***Unified Theory of Use and Acceptance of Technology (UTAUT)*** model, are used to address the acceptance of technology in general, but when vehicles are taken into the picture, the constructs present in these scales have to be modified slightly. So, in order to fit the construct of these models in the domain of car, Osswald et al. [63] proposed the ***Car Technology Acceptance Model (CTAM)***, where the constructs of this model adapted from all the previous TAM scales, are shown in the figure 1.2 below.

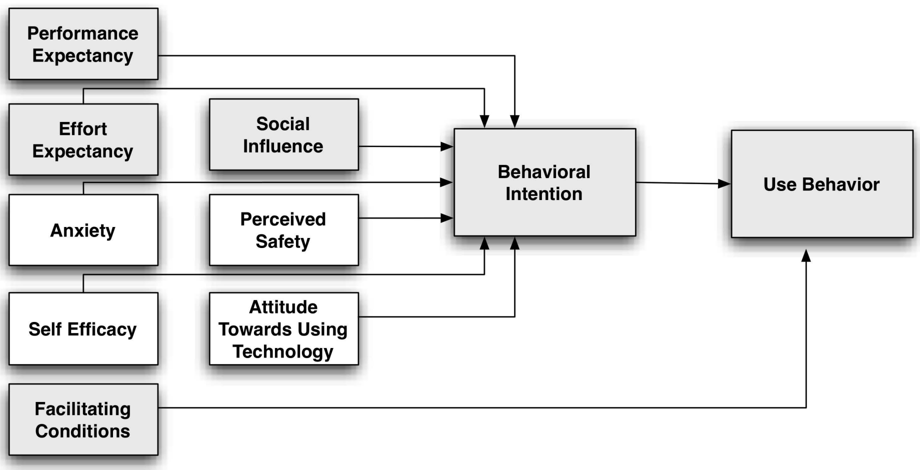


Figure 1.2: Constructs of the CTAM scale [63]

[10] http://mecue.de/english/home.html

The detailed explanation of these constructs and the underlying descriptions are provided in the appendix A for reference. Since the acceptance of AI-based automotive user interfaces are evaluated in this thesis, the CTAM scale is chosen for measuring the acceptance of the product. As shown in the figure 1.2, the scale consists of nine constructs in total, and the variables in these constructs are measured based on a 7 point likert scale ranging from -3 to +3, with 0 being a neutral score.

1.7 Thesis Outline

This thesis report consists of 8 chapters in total in addition to this, and it serves as the official guide representing the different steps involved in this thesis' span. The theoretical groundwork for this thesis is laid in the chapter 2, where the background of this research project and the subsequent motivation has been provided, which can later help to impose stronger assertion on the research questions. In chapter 3, the guidelines based on which the system has to be designed/updated are drafted, and the ways these guidelines can be used to support the user's understanding of the system is identified and coded. Once these are done, it is necessary to identify and reflect some the inconsistencies present in the existing interfaces, these are which are illustrated in detail in the chapter 4. In addition to this, methods employed to arrive at the thesis' use cases are also defined in the 4^{th} chapter. The following chapters 5 and 6 provides a brief overview regarding the steps involved in the development of the HMI under consideration, and the subsequent testing of this prototype along with the results is highlighted in the chapter 7. Finally, the entire overview of this research along with its gaps, and potential work to be carried out in the future is provided in the chapter 8.

Chapter 2

Project Background

As more and more vehicles are becoming increasingly connected, artificial intelligence-based technologies play a significant role in improving the occupant's interaction with the vehicle in addition to controlling driving related activities. Smart voice assistants, in-vehicle infotainment systems, and many more will become the major factor of Artificial Intelligence based Automotive Human Machine Interface, which leads to increased efficiency and reduced human error. So, this thesis aims to analyze the impact the AI-based HMI has in Human-HMI interaction. Before delving deep into analyzing the impact of AI-based HMI, it is important to address and discuss the fundamentals behind the research, which is based on the context of autonomous cars, its user interfaces, and the way AI is incorporated in them.

To get a clear idea of the types of AI and the ways they are being incorporated in autonomous vehicles (AV), the thought process went into the implementation of AI in autonomous vehicles has to be identified, and hence, the first part of this chapter aims to illustrate the history behind the development of autonomous vehicles, followed by addressing the state of autonomous vehicles right now and in the future, finally concluding with how AI has been brought into the development of AVs. Due to the vision towards the continuous development of cars to achieve perfection, across its development history, many standards and guidelines are also being developed and updated continuously to match the current trend. These are also explained in this chapter. The chapter then concludes by providing an overview of different evaluation methods that can be used for development and analysis of these technologies.

2.1 Autonomous Vehicles and its History

As Bel Geddes quoted in his book, Magic Motorways [28], "***These cars of 1960 and the highways on which they drive will have in them devices which will correct the faults of human beings as drivers***", such was the kind of vision he envisioned for the future, as he first proposed a concept of smart highways where the cars will drive automatically. This presentation by Bel Geddes at General Motor's exhibit Futurama in 1939, was a milestone for autonomous theorists that made them to pour significant hours of development to make his vision of autonomous vehicles come true. With only 19 years gap between the pioneering presentation, General Motors along with the cooperation from Radio Corporation of America, turned the vision of Geddes into a reality, by a paving a 400 foot, electronics and sensors embedded, highway of the future[1]. For experience the so called "autonomousness" of the vehicle, cars mounted with a huge sensor array are made to drive over these roads, where it picks signals from the electric cables running underneath to gather information such a lane/road obstruction, allowing the car to autonomously brake based on the information.

Although this vision and significant development of autonomous cars, were novel at that time, it was simply impossible to achieve automation on all types of roads. Robert Fenton and Karl Olson of the Ohio State University paved a vision of an ***Electronic Highway*** [25], where they posit a concept of a dual-mode system, where the car can operate autonomously in such called electronic highways, but has to switch-over to manual mode on non-automated roads. As most of the autonomous facilities at that time relied heavily on infrastructure built into the roads directly, this concept was significant, however, not so efficient as this highway autonomy were extremely harder to expand upon and demanded a huge sum of money. After this point, it was decided that only way to move forward is to make the car to make decision on its own rather relying on sensors and coils embedded within the roads, This in turn led to development of these so-called smart cars which will be capable of making its own decisions,

Out of all the other researchers delving into this newly coined autonomous vision at that time, researchers from Stanford envisioned the idea of the smart car as a platform for lunar rovers. Being initially built as a four wheel cart as shown in the figure 2.1, mounted with a video camera for navigating purposes, developed into a first instance of an autonomous vehicle, containing greater intelligence and image processing capabilities, and it was the first one at that time to cross a room filled with chairs without human intervention[2]. However, this was cart was not considered to be the first stand-alone autonomous vehicle, and this feat was achieved in 1977 by the researchers at the Tsukuba Mechanical Engineering Laboratory, Japan, where the car was able to employ machine vision and signal processing to make its own decision. The car acted autonomously based on the information gathered from the surroundings, instead of being reliant on the road infrastructure as was in the earlier days.

Figure 2.1: Lunar Autonomous Cart by Stanford

[1]https://bit.ly/3pB4FbU
[2]https://web.stanford.edu/~learnest/sail/oldcart.html

2.2 Incorporation of AI in Autonomous Vehicles

The autonomous behavior exhibited by the Stanford lunar rover cart, was the first step taken to bring these cars into the region of ***pervasive computing***, also called as ***ubiquitous computing***[3], which means by making these cars communicate with other at any time, in any place, and in any data format across a wide band of network arrays. After the successful run exhibited by the researches at Japan, the employment of signal processing and machine vision in the cars that can enable them to make decisions on its own, paved the way of incorporating AI to these cars. During the mid 80s, interest in application of AI into the autonomous cars peaked, due to the autonomous cars developed by a German Aerospace Engineer, Ernst Dickmanns, whose pioneering work in autonomous vehicles enabled Europe to launch the biggest research and development project on driverless cars called the ***PROMETHEUS*** project[4], where in collaboration with Mercedes Benz, their two robot prototypes, VaMP and VITA-2, drove autonomously for more than 1000 miles at speeds of nearly 130 kilometers per hour along the french highway in 1994 [19].

Figure 2.2: Autonomous Road Vehicles under PROMETHEUS project : Source - politico.eu

The successful run produced by this project enabled a continuously evolving research in the field of autonomous vehicles around the late 20^{th} century, and was propelled even further at the start of the 21^{st} century, due to the DARPA's grand challenge for autonomous vehicles[5]. From these challenges till now, AI has been revolutionizing the automotive industry

[3] https://bit.ly/38NI6dq
[4] https://en.wikipedia.org/wiki/Eureka_Prometheus_Project
[5] https://www.wired.com/story/darpa-grand-urban-challenge-self-driving-car/

in more significant ways than imagined, as it is now easier to gain critical insights into the cars, roadways, peoples, and also the process, due to abundance amount of data set collected and processed. These new trends enable the manufacturers to delve deep into the AI spectrum where they can play with various types of AI such as investing in ***machine learning*** which helps the AI-build to improve methodically over time, and also in so called ***deep learning*** which offers new and improved data to the researchers by means of neural networks[6]. Even though there are a significant amount of research and development involved in incorporating AI in to autonomous vehicles and also in other use cases, the implementation of it has only increased slightly recent days. Based on a survey conducted by Capegemini research institute in January 2019, only 10% scaled implementation of AI was found at that time, and the number of AI initiatives that has not been implemented yet, stood at 39%[7] at the start of 2019.

The main reason for this is the level of hype that has been placed upon these AI systems, such as being able to solve everything and expecting the pilot project to work for the first time. Scaling of these systems will significantly decrease if it does not deliver according to the set hype[8]. In addition to this, skill also play a huge role in development and deployment of these AI-based systems. According to Andrew Ng, founder of deeplearning.ai[9], without motivating cross functionality between various departments and instead being relying solely upon some machine learning engineers, the AI pilot projects will not carry over to the final stages [56]. In addition to these setbacks, the problem of ***selection and scaling*** of use cases for which AI has to be adopted, is still a major issue amongst all the developers, since more and more employers, as shown in the figure 2.3, are launching pilot AI projects that addresses their own needs and problems, Although the battle is fierce, it enabled the companies and research institutions to broaden their level of understanding of AI and where it can bring benefits, thereby leading to more selective AI scaling engagements in the future.

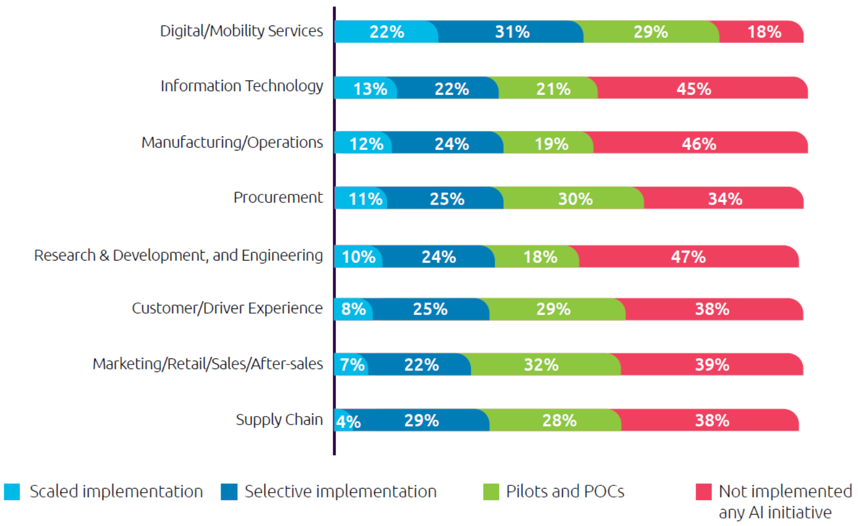


Figure 2.3: State of AI implementation in the Automotive Industry : Source - Capegemini

[6]https://www.netapp.com/media/8730-e-book-ai-in-automotive.pdf
[7]https://bit.ly/2L5ukdP
[8]https://bit.ly/2LbziG4
[9]https://www.deeplearning.ai/

2.3 Role of AI in Connected Vehicles

As mentioned in the beginning of this chapter, the vision of an autonomous vehicles leads to the vehicle to make decision on its own, instead of being completely reliant upon an external architecture. Also, in addition to the vehicle making decision on its own, it should also listen and consider the decision taken by other similar vehicles that are in the vicinity. Making these talk with each other will make the rides in these vehicles ***smoother and safer***, as the cars now can receive and share important safety and mobility related information between each other (if requested). This led to an significant increase in the platform called ***connected cars***, where the information was made to be transmitted either through GPS or by wireless communication between the vehicles, and also to mobile devices, smart phones, and many connected appliances. One such example for this is the ***digital cockpit concept***[10] from Samsung, where the car enables bi-directional communication to the user's home appliances through their SmartThings platform.

With the vehicles becoming more and more connected in the present, these vehicles are to be deployed in such a way that it not only does convey driving and context related information to the driver for alerting them, but also to significantly enhance both the vehicle's and the occupants' driving behavior[11]. Being considered a part of intelligent transport systems, these vehicle-2-vehicle communication technologies[12] are being developed as a part of connected cars realm, whose main core function is to improve automotive safety and increase the riding efficiency. So, it is of core necessity of these systems to incorporate AI in their programs, and most importantly, define the role that the AI has to perform within them. The global AI survey conducted by Arm revealed that, the user's perceived three distinct roles that AI has in certain applications[13], which can be seen in the figure 2.4 below.

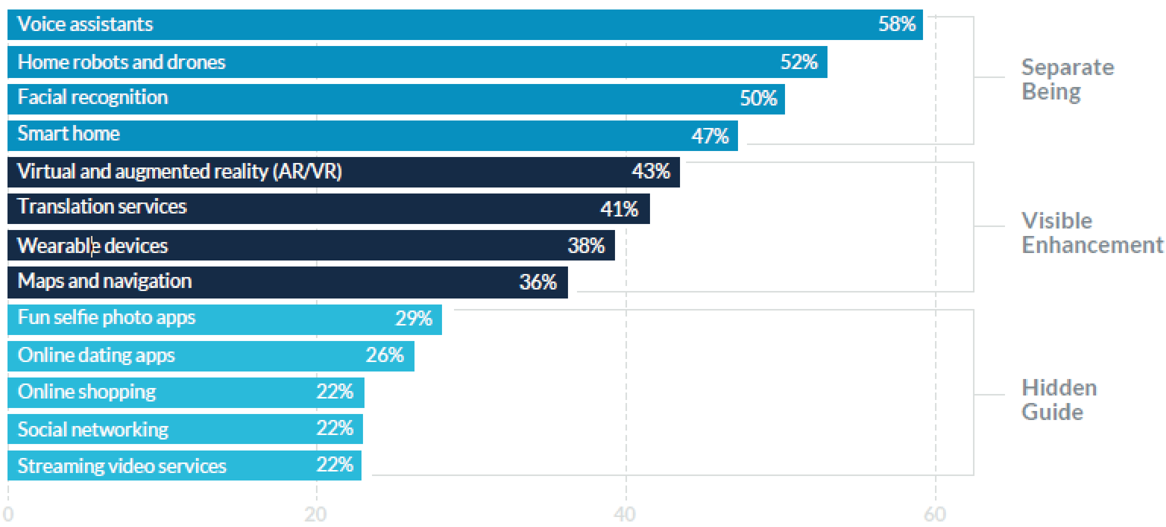


Figure 2.4: Essential Roles of AI and its implications : Source - arm Blueprint

[10]https://news.samsung.com/global/digital-cockpit-drives-the-future-of-connected-cars
[11]https://www.digi.com/blog/post/what-is-connected-vehicle-technology-and-use-cases
[12]https://www.nhtsa.gov/technology-innovation/vehicle-vehicle-communication
[13]https://bit.ly/396zVsT

The first role was by considering AI as a ***separate entity***, which means that the user's can interact with the system in a manner as it was alive. It can also be considered to fulfil the role as an assistant who sole purpose is to ***enhance or change*** the way the user's interact with the world. Finally, the AI can also be made to fulfil the role of a ***hidden entity***, which merely serves as a guide for the user to achieve their intended interaction, and also to assist and improves the quality of the program without the user noticing it. Thus, the major challenge for the automotive sector is to understand what these connected cars could do to secure a significant consumer base, and they key to meeting this challenge is to enhance and update the system and security of these over software-on-the-air (SOTA) updates, at any time and at any place, thereby leaning towards the ***continuous integration and continuous delivery*** [99] model.

2.4 Levels of Automation

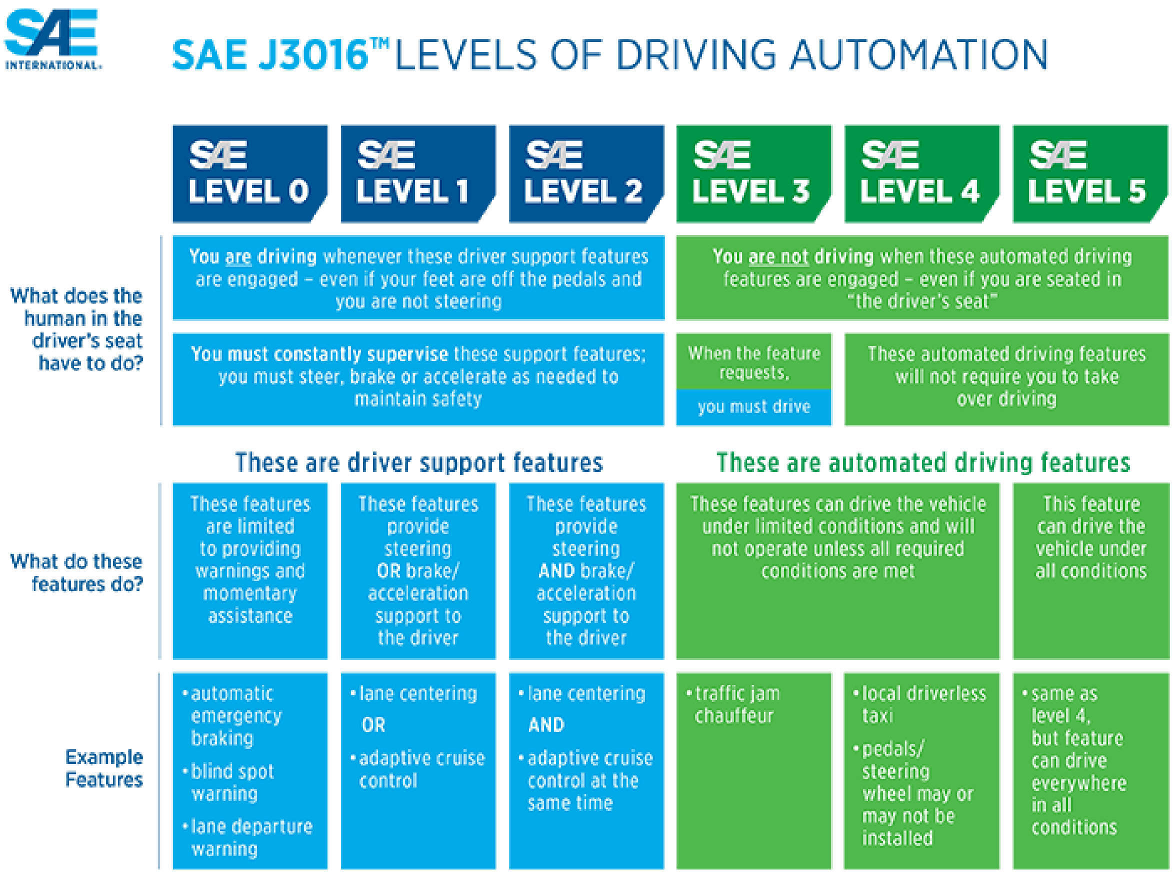


Figure 2.5: Levels of Automation provided : Source - SAE

The three roles of AI in a connected car mentioned above, will also come into play depending upon the type of activity that are being performed in the car (not)during driving, and most importantly, the level of automation that has been present in the vehicle. As connected cars and autonomous vehicles gain importance and traction, a common taxonomy representing the levels of automation by considering the technical and legal aspects of automation, has to be formed. Multiple international bodies like the German Federal Highway Research Institute (Bundesanstalt für Straßenwesen (BASt)) [27], and the United States' National Highway Traffic Safety Administration (NHTSA) [94] released a classification system, but the taxonomy released by Society of Automotive Engineers (SAE) at 2014 is considered to be a standard used for representing the levels of automated driving[14]. The visual chart generated by SAE representing the levels of automation is shown in the figure 2.5. The narrative definitions of levels of automation accessed from the SAE excerpt [36] is provided below.

SAE Level 0 - No Automation

The full-time performance by the human driver of all aspects of the dynamic driving task, even when "enhanced by warning or intervention systems. In this level, only the human driver is responsible for steering execution along with acceleration/deceleration.

SAE Level 1 - Driver Assistance

The driving mode-specific execution by a driver assistance system of "**either** steering **or** acceleration/deceleration" using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task. Here, both the human driver and system is responsible for controlling the car.

SAE Level 2 - Partial Automation

The driving mode-specific execution by one or more driver assistance systems of **both** steering and acceleration/deceleration using information about the driving environment and with the expectation that the human driver performs all remaining aspects of the dynamic driving task. Even though the driver has to monitor and control the environment at this level, the execution of the steering along with acceleration/deceleration is performed by the system.

SAE Level 3 - Conditional Automation

The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene. In this level, the driver is a necessity but not required to monitor the environment; however, the driver must be ready to take over the control of the vehicle whenever prompted.

SAE Level 4 - High Automation

The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task even if a human driver does not respond appropriately to a request to intervene the car can pull over safely by guiding system.

SAE Level 5 - Full Automation

The driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver.

[14] `https://bit.ly/3507Mm2`

2.5 AI for improving the Driver and Passenger Experience

Since the advent of AI incorporation in autonomous vehicles, especially in the realm of connected cars is increasing, successfully exploiting and incorporating these in the automotive industry will depend on the ability of the manufacturers and researchers to include key concepts from multiple AI disciplines. As said in the section 2.2, in addition to incorporating machine and deep learning into AI development, ***data science*** is also a key factor, where it can be used to collect and analyze large heaps of data with a main intent to arrive at a hypothesis that the machine learning programs can work upon. Involving these three core parts of AI in the automotive development along with the inclusion of cognitive computing and chatbot development, should make this a necessary ingredient for successful AI exploitation.

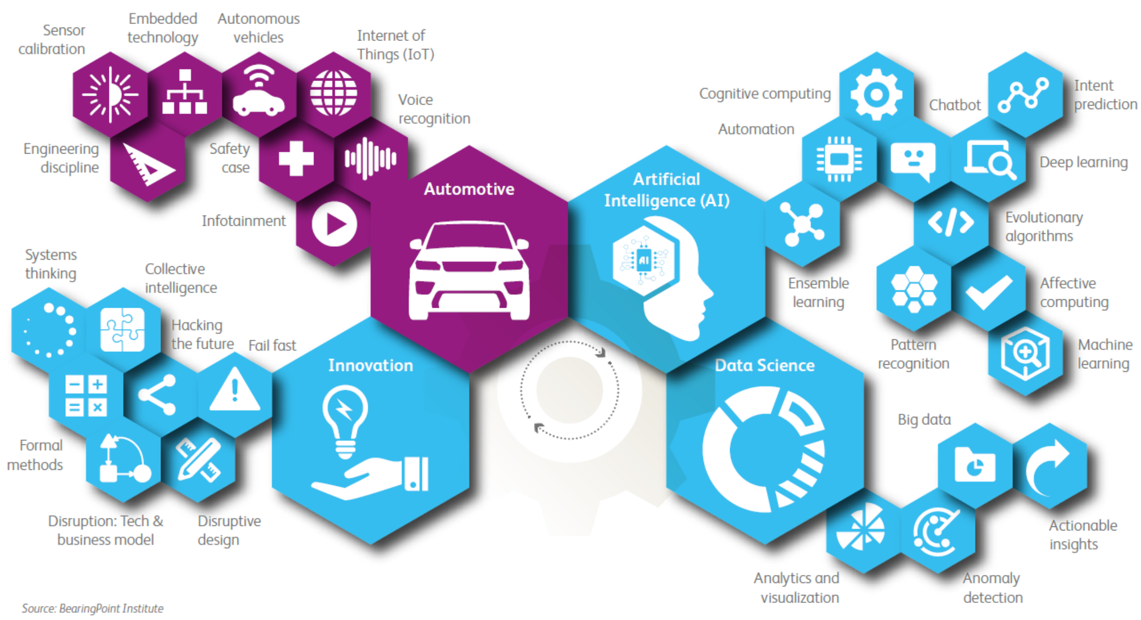


Figure 2.6: Different AI approaches that can be exploited : Source - BearingPoint Institute

As seen in the figure 2.6, exploiting AI into the automotive domain (illustrated by a purple color in the diagram) will effectively lead the path towards a car that is capable of achieving true autonomy; however as mentioned in the section 2.2, the implementation of AI in the automotive industry has only grown marginally till 2019. In spite of the slow scaling of AI, companies and investors alike in the automotive industry are looking to boost their AI capabilities by investing in startups where AI is considered to be key part of the company's business model. As per the market research done by Capegemini[15], the amount in millions invested by auto companies in these AI-led startups is the ***highest*** for the ones involving in driver and passenger experience, followed by the investment in mobility services. By making significant advancements in safety and reshaping the tech through disruptive innovations, will re-tone the way in which riding can be experienced. Exploitation of the above mentioned AI solutions effectively, will bring the vision of a perfect autonomous car closer to reality, and as evident from the Capegemini survey, automakers are beginning to cash in on the solutions

[15] https://bit.ly/2L5ukdP

that can enhance the in-vehicle experience of the drivers and the passengers of the vehicle. Even though the technology that went into improving the user experience might render the driver of the vehicle obsolete, these can also be adopted for a shared autonomous future. Some of technology that the automakers currently research and inspect for improving the experience[16] is shown in the figure 2.7 below.

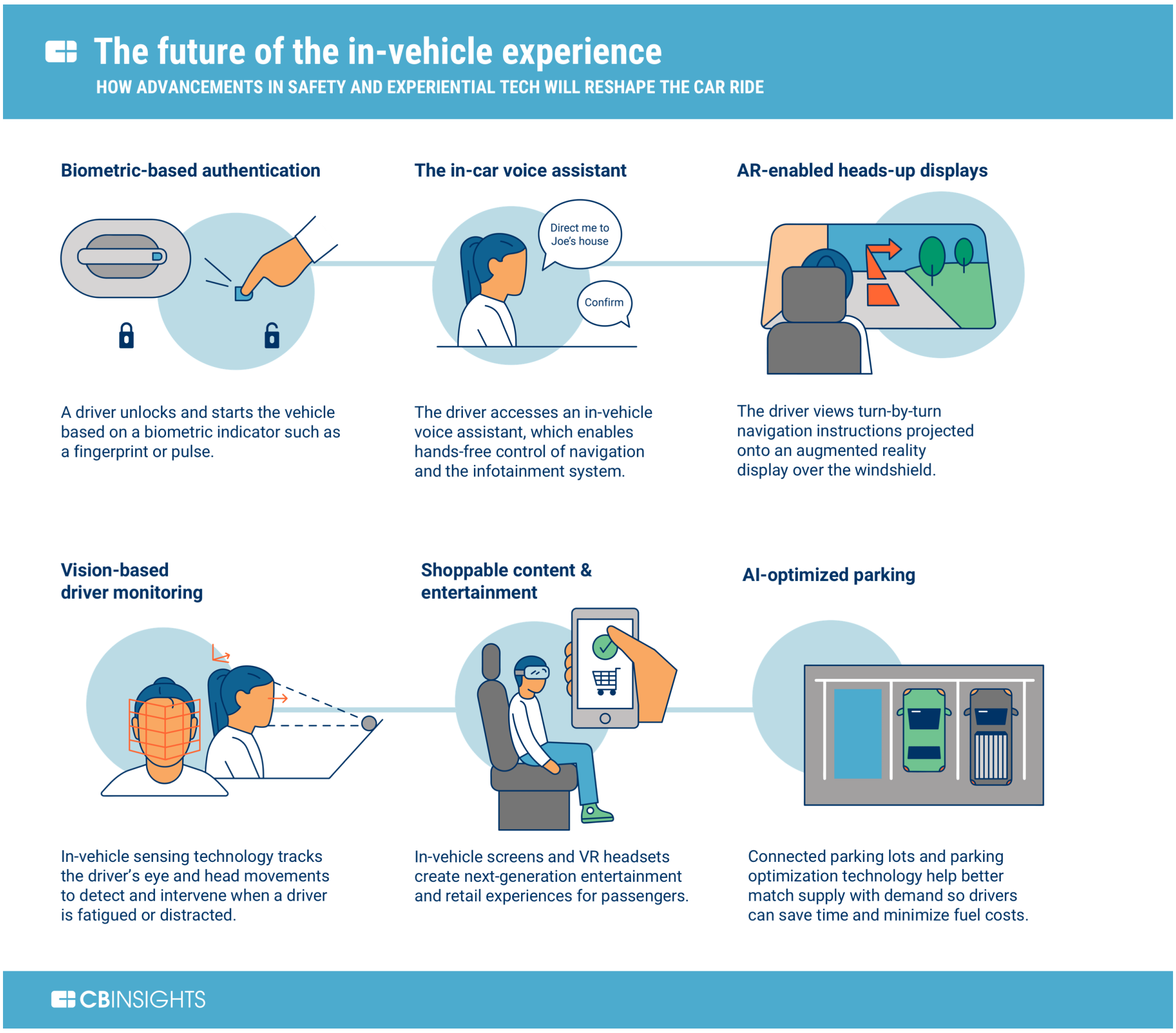


Figure 2.7: Technologies involved in improving the driver and passenger experience : Source - CBInsights

Each and every actions that can be performed by the passengers in a vehicle, like from entering into the vehicle and starting the ignition, to parking and getting out of the vehicle, is monitored for improving the experience. Out of all these technologies, the in-car voice assistants are and has been significantly researched and improved upon. Even though being commonly implemented in vehicles, especially in cars, these in-car voice assistants are far from perfect. These assistants should support ***contextual understanding*** which in hindsight enables smooth two-way interactive communication and disables the need for the user to

[16] https://www.cbinsights.com/research/report/in-vehicle-experience-technology-future/

constantly repeat their query. Another important point to remember when considering conversational AIs is that it needs to adapt to tomorrow's needs and expectations, and should keep the conversations on track. This can be made possible through machine learning, which allows the application to make smart decisions and complex interfaces, while the rules ensure the system maintains a consistent and correct personality. Companies like Cerence[17] are pioneering in speech technologies to achieve this, where their ***dragon drive*** proves to be an AI assistant with expanded conversational and cognitive capabilities, with the ability to allow the users to ask with no wake-up phrase or button press.

In addition to in-car voice assistants, another use case than can create a significant additional revenue streams to the product is through in-car commerce and on-demand connected services. To keep the passengers occupied and entertained, these automakers are pouring a lot of time in investing in immersive in-vehicle experiences. By the application of these, the passengers can now reserve and/or place an order of what they want within the car itself. The first iteration of this was soon made available by General Motors. Using IBM's Watson technology[18], they created a platform called ***OnStar Go***, where it features a custom features and content from specific market partners[19]. By learning the daily behavior of the users, IBM Watson can suggest a couple of services that the users tend to perform at that time. Also, more and more automakers like Tesla, and the concept Sony car presented in CES 2020, has support for multiple streaming services like Netflix and Hulu, which can be viewed-on-demand by the user's while driving. Streaming services aside, support for playing games are also becoming more and more present in cars, which is the clear case in the Audi's ***Holoride***[20] presentation at CES 2019. Their vision of transforming car into virtual theme parks was the main attraction of their presentation, where they used a motion synchronized Virtual Reality (VR) that enables the passengers to play games and watch movies whose main purpose is to combat motion sickness experienced by the passengers.

Out of the other technologies mentioned in the figure 2.7, the use of bio-metric based authentication is gaining a significant amount of traction in improving the security of these vehicles. Since the advantages of incorporating biometrics in the field of automobiles are being conveyed almost everytime, automakers are trying to implement the biometrics in the cars, the same way as present in mobile and laptop devices. Automakers like Hyundai and the new 2020 Mercedes S-class[21] contains a mandatory biometric authentication, only after which the ignition of the car can be turned on. In addition to employing biometric technology, the new Mercedes S class also contains an AR enable dynamic heads-up-display which displays glowing lights under the car in front to help the driver maintain a safe distance. The system projects red lines on the edge of corners, and it even displays a plethora of computer-generated images to help the driver safely navigate a roundabout. Utilizing technologies similar these will help in achieving the vision of a vehicle that can significantly impact and improve the driver and passenger experience.

[17] http://cerence.com/cerence-products/cerence-drive
[18] https://www.ibm.com/in-en/watson
[19] https://bit.ly/3rJPTRT
[20] https://www.holoride.com/
[21] https://www.mercedes-benz.com/en/vehicles/passenger-cars/s-class/

2.6 AI and Automotive User Interfaces

Even if the car is equipped with different technologies that can enhance the driver and passenger experience, it is therefore necessary to consider the implication that the drivers and passengers might have when experiencing these for the first time. As these technologies are constantly being implemented and tested in vehicles, it is therefore suffice to say that the vehicles, especially cars, are now being considered as an infotainment platform, packed with features and services that the user's might expect. In addition to laying emphasis on safety, the future applications and interfaces need to be designed in such a way that it should enable communication, work and entertainment that can be offered to the occupants of the vehicle. From the report of the Dagstuhl Seminar, Riener et al. [77] summarized three core challenges these developed interfaces could face in the time of automation.

1. Due to the inclusion of shoppable-content and entertainment services within the car, it demands the attention shift from the passengers (drivers too in case a fully autonomous vehicle) towards viewing these services and other non-driving related activities that the user's tend to perform for some time and/or even for a longer time. When highly automated vehicles are taken into consideration, more focus has to be laid towards the areas where these interactions can be improved, and conversely focusing less towards the driver's attention on the road. In addition to improving the areas of interaction for a single user, these interactions might also include other people within the vehicle or entirely from a separate vehicle itself. Hence, designing these user interfaces to support all kinds of interactions and to be compatible with every other vehicles in the surrounding is the key factor for transforming vehicles as an interaction medium.

2. Most of the automotive critics firmly believe that more autonomous the vehicle becomes, then more disengaged the driver will be from the driving task. This might lead to serious safety hazards, such as accidents reported by automaker Tesla due to their Auto Pilot system[22]. Due to the plethora of non-driving related activities that occupant's of the vehicle can perform [67], it might become a challenge for some occupants to shift their attention suddenly to the primary driving task in case of a take-over-request (TOR) is invoked. Thus, the interface and system designers should design a system in such a way that it can create a balance between the primary and secondary tasks, and should facilitate effective re-engagement between these two at any context.

3. Also revealed in the global AI survey conducted by arm[23] is that, among the wishlist of what the users expect when AI was brought into the automation context, traffic control systems that were designed to modulate the vehicle flows to ease congestion, topped the wishlist. As mentioned in the beginning of this chapter, even though old, the dual-mode system proposed by Robert Fenton and Karl Olson [25] in 1969, still holds true in the present. Which means that, during traffic situations there might come a time where information from a fully autonomous vehicle has to be transmitted to a non-autonomous vehicles to keep the drivers of those cars in a information loop, and the reverse should be considered. So more emphasis has to be placed on how to convey these information in a mixed traffic scenarios, as a scenario where all the cars corresponding to the SAE automation levels will coexist in the same space.

[22] `https://bit.ly/382J06t`
[23] `https://bit.ly/396zVsT`

Designing the user interfaces that addresses these three challenges is considered as a milestone, as the systems addressing these should be capable of enhancing existing systems and to develop new disruptive interfaces, thereby improving its performance. However, research revealed that interacting with these AI-based user interfaces could trigger some contradictory behavioral responses from the users [80]. This can be associated with the uncanny valley hypothesis [55] which states that the user's familiarity with the system fades as the system tries to become more like human but fails to represent this appearance by any means. So, it is essential to fit the levels of interactions provided by the system with the cognitive style (mental model) of the occupants which can be done by enabling the occupants to interact with the vehicles through different modes, and the same can also be said for the system communicating the information back to occupants. The provision of these type of ***multi modal interactions*** as shown in the figure 2.8[24], will make enable interfaces to be inherently flexible, and to especially provide an ideal interface for accommodating both the changing demands encountered during interactions and also the large individual differences present in the population [68]. Multimodal interfaces thereby provides the necessary advantages of being robust in nature and aims to reduce the driver's perceived workload and distraction.

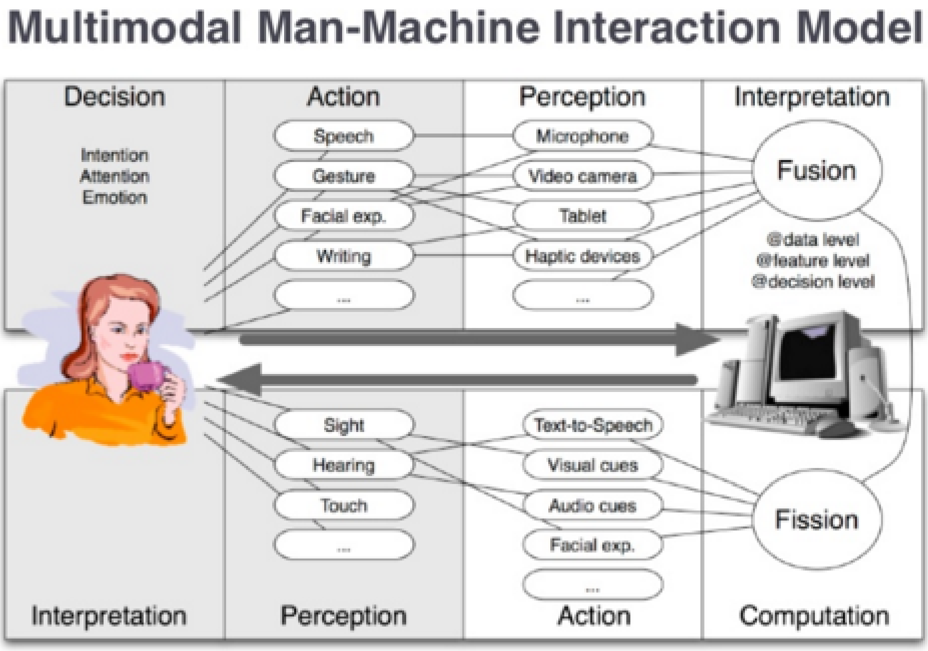


Figure 2.8: Multi Modal Interaction in a MMI : Source - Arizona State University

Using multimodal interaction, the users can now input their interactions to the system by either using the touch input or by invoking the voice command or through gestures, and the system can seamlessly combine these inputs, processes them, and conveys the requested information back to the user also through a combination of different modes, such as like a visual cue through a display or via a voice assistant or by means of haptic feedback. The way these information conveyed to the occupants can be enhances by combining these three output modes in such a way that it suits the information being conveyed.

[24] https://bit.ly/3rL5CjN

2.7 Human Factors in AI-based User Interfaces

Due to the complex interaction sequences and machine intelligence offered by these interfaces, a shift from a mere conventional one way interaction system to having a much more evolved and complex cooperative human-machine relation is evident, and for this purpose, the incorporation of human factor principles and standards in the early stages of AI-based user interfaces development is a necessity and will lead towards a fully functioning and more usable product. Wickens et al. [102] suggests that, the breakdown in the interactions between humans and the interfaces with which they work, is the core of human factors, and they define this concept as "***a study of factors and development of tools that enhances the performance of these systems, increases safety and user satisfaction which will lead to the achievement of the product's goals***". Thus, the principles of human factors has to applied in such a way that it focuses entirely on the thinking and knowledge aspects of these systems, irrespective of whether the tasks are being carried out by a human or through virtual agents.

Since user interfaces are being considered here, the way that the user's perceive the information being conveyed from these displays, how they process and store these information in their memory that enables them to trace back to their mental model, and how they relate these perceived information to human attention, needs to be analyzed. As shown in the figure 2.9, Wickens et al. [102] define these displays as a human-made artifacts designed to support the perception of the system variables and facilitating further information processing. Similar to participants exhibiting changes in their behavior during interacting with the system, there is not an instance of a perfect display that is best suited for all the provided tasks. So, in order to meet a common ground, it is efficient to map these displays based on a specific set of principles and guidelines that addresses human perception and information processing.

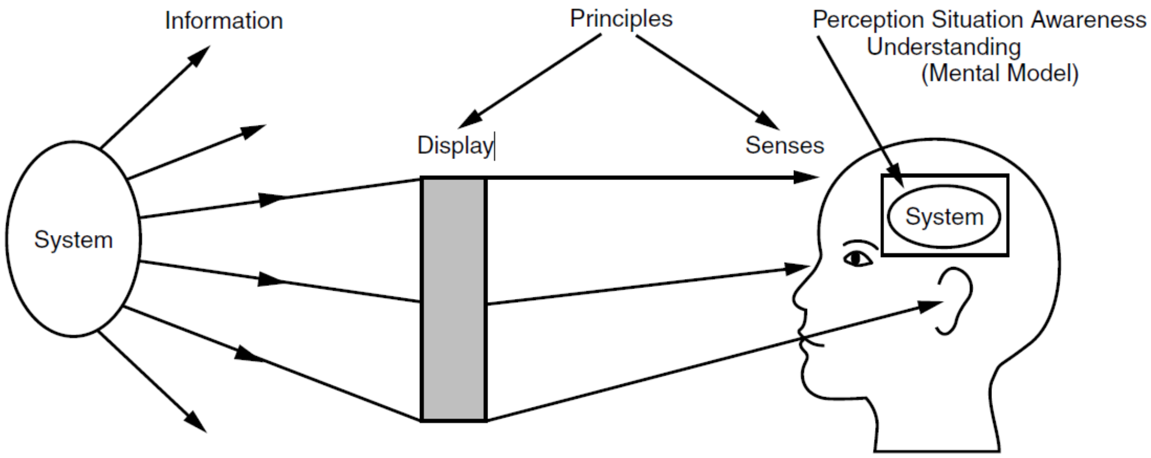


Figure 2.9: Key components in Display Design - Retrieved from [102]

However, while designing for interactive interfaces, a multitude of guidelines elicited by different authors involved in the human factors research, in addition to the plethora of taxonomies published by international standards and government bodies alike, exist for each and every specific use case that are taken into consideration. Be it the criteria and verification procedures on driver interactions by Alliance of Automobile Manufacturers, to European Statement of Principles for the Human Machine Interface, and to in-display guidelines from Japan Automobile Manufacturers Association (JAMA), there exist a couple of similar albeit different set of guidelines in addition to the standards set by ISO, NHTSA and SAE. Furthermore, many scientists like Wickens et al. [102], Shneiderman [85], Ross et al.[78] and Stevens et al. [89], suggested their own set of guidelines of improving the display to suit the human factors needs. One major problem with the presence of multitude of these display guidelines is that some of them will be in contradiction with other, and also the gives rise to problem of knowledge gap. Here arises a need to analyze and validate these guidelines in such a way that it can be used to match the needs and expectations that these displays can deliver upon. Hence for this thesis, the guidelines that are used for designing the display is adapted from the literature review done by Normark et al. [61] along with industry approved Microsoft Style Guidelines[25]. From the thorough literature review done on various standards and guidelines, Normark et al., classified these guidelines on basis of compliance with the existing standard, being referenced in other peer publications or being technologically dependent.

Displays that are designed based on the above mentioned guidelines are bound to have some usability problems which has to be addressed at the starting stages of design itself. In order to find these usability problems, Jakob Nielsen [58] proposes a set of 10 usability heuristics, which are then used as reference for evaluators to find the issues present in the prototype. These 10 usability heuristics[26] are provided below.

1: Visibility of the System Status

The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time. When users know the current system status, they learn the outcome of their prior interactions and determine next steps. Predictable interactions create trust in the product as well as the brand.

2: Match between system and the real world

The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order. When a design's controls follow real-world conventions and correspond to desired outcomes (called natural mapping), it's easier for users to learn and remember how the interface works. This helps to build an experience that feels intuitive.

3: User control and freedom

Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process. When it's easy for people to back out of a process or undo an action, it fosters a sense of freedom and confidence. Exits allow users to remain in control of the system and avoid getting stuck and feeling frustrated.

[25] https://docs.microsoft.com/en-us/style-guide/welcome/
[26] https://www.nngroup.com/articles/ten-usability-heuristics/

4: Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform and industry conventions. Users' experiences with those other products set their expectations. Failing to maintain consistency may increase the users' cognitive load by forcing them to learn something new.

5: Error prevention

Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions, or check for them and present users with a confirmation option before they commit to the action.

6: Recognition rather than recall

Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g. field labels or menu items) should be visible or easily retrievable when needed.

7: Flexibility and efficiency of use

Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions. Flexible processes can be carried out in different ways, so that people can pick whichever method works for them.

8: Aesthetic and minimalist design

Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility. This heuristic doesn't mean you have to use a flat design — it's about making sure you're keeping the content and visual design focused on the essentials. Ensure that the visual elements of the interface support the user's primary goals.

9: Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution. These error messages should also be presented with visual treatments that will help users notice and recognize them.

10: Help and documentation

It's best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks. Help and documentation content should be easy to search and focused on the user's task. Keep it concise, and list concrete steps that need to be carried out.

2.8 Summary

In this chapter, the technologies that has been used in the first iteration of an autonomous vehicle to the technologies like AI being used in the current generation of vehicles, for the sole purpose of improving the driver and passenger experience, is discussed in detail. The way AI has been incorporated in the automotive user interfaces and how it impacts the interactions offered and improves the user experiences, is explained in detail. Finally, the set of guidelines that can be used to develop these displays and to the ones that are subsequently used to improve these displays by identifying the usability problems, is provided. In the coming chapters, a more detailed look of how incorporating AI in automotive user interfaces improves the interaction and experience of the occupants is discussed in more detail.

Chapter 3

HMI Guidelines for Human-AI Interaction

The arrival of autonomous vehicles in the market is fast approaching, and still the evidence that suggests the benefits of AI based user interfaces and its ability to be consistent across all the systems are scarce. In addition to this, as many organizations proceed with the development of AI-based HMI using different set of guidelines, tools and techniques [34, 35, 60], it is certain that many inconsistencies and errors between various systems in the market are bound to occur, because of the system's continuous change in its architecture by learning over time, and also due to some incorrectly/poorly understood behaviors of the user. As a result of these alarming inconsistencies (e.g. variations in interaction styles [49]), in addition to the occurrences of both simple[1] and major errors (leading the system ineffective to respond to the user's interaction[2]), designers will face a major problem and challenge of integrating AI services into their new apps, system, and devices. Based on the research done by Dove et al. [21], UX designers and researchers still faced problems while integrating ML and AI based technologies into their devices, due to their lack of knowledge in its technical underpinnings and also due to their lack of understanding in the technology.

Thus, it is of absolute essential to make the system designed not only being easily understood by the designer and the user, but also making the system trustworthy. This can be only be achieved by providing a consistent explanatory aspects (guidelines) thereby guiding the user towards the necessary steps and actions to be taken in designing a consistent interface that is both understood by the community and being trustworthy [29]. One thing to note that, the main objective of drafting these guidelines for HMI is to be consistent across most of the platforms. However, the ML and AI systems being developed on the basis of these guidelines, has to collect more and more data, where there might be an occurrence of the data being unused and/or being inaccessible [7]. Therefore, the guideline should also be drafted in such a way that it ensures the reuse and retrieval data will not be subjected to some of the common mishaps explained in the previous sentence. So, this chapter addresses the process of obtaining the specific set of guidelines for Human-AI interaction with the automotive HMI, and categorizes the guidelines in way that reflects the user's mindset, involvement and their knowledge outcomes while interacting with the system designed with the guidelines [24].

[1] https://bit.ly/3gqYUds
[2] https://www.wired.com/story/tesla-autopilot-why-crash-radar/

3.1 Selection and Codification of Guidelines

The guidelines that are necessary for addressing the inconsistencies and problems mentioned above, are then derived from the vast range of sources available (academic journals, online articles, patents, design guides, etc.) which were published over the past 20 years. But, in order to keep this set of guidelines relevant and up to date to the current trend, most of the sources are selected in such a way that these were published at or after 2010 (in case of patents, the ones which are submitted for patents or has a pending status as of from 2010). However, an exemption was made for some of the classical interaction guidelines, as the newer articles used the classical articles as a benchmark/standard in their publications.

The selection criteria of these resources was based on the relevance to the core research question of this thesis, meaning, articles, patents and blogs were searched in such a way that it poses a significant amount of relevance towards AI and ML, interacting with automotive HMIs, AI-based intelligent and adaptive recommender systems, the user's trust in these AI-based systems, knowledge and subsequent understanding of these systems, and many more. As said in the beginning of this chapter, the information regarding interaction with AI-based system is scarce, particularly human interaction guidelines that improves the interaction with the HMI are too-scarce [2]. This due to the fact that some of the articles and publications generally contain relevant information that are not directly elicited or classified as suggestions or guidelines. As a result, similar design guidelines that were used in other intelligent and/or automated systems, also in other commercial products other than vehicle's HMI, is considered for curating these guidelines.

Before selecting the guidelines, all the selected literature was read through briefly in order to examine the relevancy and validity of the article with respect to the thesis' core topic, and also to identify the key insights of the article. After rejecting some articles that failed to meet the selection criteria mentioned above, a little more than 25 published articles, 7 patents and some more online blog posts and articles, were read again thoroughly to curate the necessary guidelines. All the relevant statements that was encountered during this systematic review of the articles were listed in a spreadsheet. These 171 insights are first grouped together as shown in the figure 3.1, based on its relevance towards to the core research topic. The grouping and clustering of these insights was done using an online collaboration tool called MURAL.

The next step involves grouping of these similar insights as a single group, which was done asynchronously using ***Affinity Diagramming*** method as shown in the figure 3.2, which refers to organizing similar and relevant insights into distinct clusters[3]. From this figure, it can be seen that the initial clustering of these insights yielded 13 groups in total. In order to avoid repetition in the final guidelines, similar insights within the group cluster are identified and linked with each other, thereby leading to the formulation of 17 specific category guidelines. Ethical guidelines ([37, 51]) kept separate from these guidelines since these are standardized and can be applied to any design guidelines that are developed or that are in-development. It is to be noted that these guidelines are not considered as the functional requirement for development, but as a standard, used for guiding the concerned people into developing the devices that improves the Human-AI interaction.

[3] `https://www.nngroup.com/articles/affinity-diagram/`

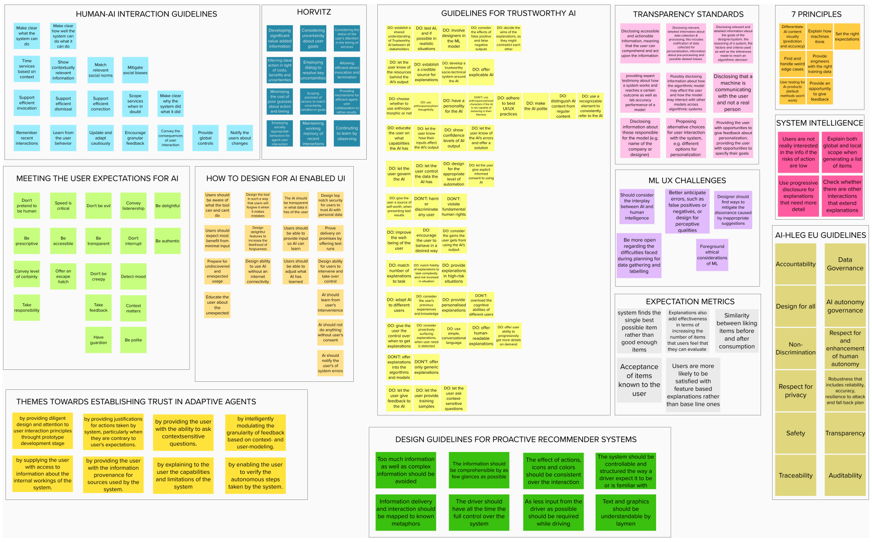


Figure 3.1: Clustering of Insights due to relevance

Figure 3.2: Clustering of Insights into distinct groups

3.2 Final Set of Guidelines

The resulting 17 specific categorical guidelines along with its sub-statements and its subsequent references are listed below. These set of guidelines and sub-statements are curated based on the article ***"Component Design Guidelines"*** published by Nathan Curtis[4].

[4] `https://medium.com/eightshapes-llc/component-design-guidelines-eca706100e7c`

1. **Explain the user with respect to the capabilities that the AI framework has**
   (a) Give declaration about how a framework functions and arrive at a specific result just along with execution of a model with accuracy [2, 4, 47, 83]
   (b) Utilize progressive disclosure for clarifications that the user might feel overwhelming and provide it to them in a readable and understandable format [29, 71, 83]
   (c) Disclosing relevant and detailed information about the goals of the designer/system, the reasoning of a system, the factors and criteria used as well as the inferences made to reach an algorithmic decision [2, 29, 35, 92]
   (d) Explain both the understanding of the entire system as a whole (global) and an individual instance (local) from the generated list of suggestions [5, 45, 69, 74, 83]
2. **Identify, educate and mitigate the potential errors and unexpected edge cases that might occur during user interaction**
   (a) Support predictive maintenance by anticipating errors, such as false positives or negatives, or design for perceptive qualities [2, 21, 45, 103]
   (b) By providing justifications for actions taken by system, particularly when they are contrary to user's expectations [29, 74, 83]
   (c) Should find ways to mitigate the dissonance caused by inappropriate suggestions by proposing alternate choices for enabling efficient personalization [9, 21, 29, 83]
3. **AI should be transparent in what data it has of the user and how it can be accessed by the user**
   (a) Disclose accessible and actionable information, so that the user can comprehend and act upon the information [21, 45, 71, 83]
   (b) Disclose relevant, detailed information about data collection and processing. This includes notification of data collected for personalization, information about pre-processing and possible data set biases [21, 29, 71, 83]
   (c) Make sure the system is predictable by giving control to the user by making the system transparent (adaptivity) [2, 29, 88]
   (d) By providing the user with the information regarding the roots or origins of the sources used by the system [45]
   (e) By supplying the user with access to information about the internal workings and the output of the system [29, 83, 90]
4. **Enable the user to provide feedback by indicating their preferences during regular interaction with the AI system**
   (a) Intelligently modulating the quality of the feedback received from the user based on context and user modelling [2, 29, 45, 74, 90]
   (b) Let the user provide training data to the system through which the system can learn from the samples provided [29, 72, 83]
   (c) Providing the user with opportunities to give feedback about personalization [2, 72, 83, 90, 92]

5. **Maintaining the working memory of recent interactions by learning continuously from user behavior**
   (a) Consider the user's previous experiences and knowledge regarding their interaction with the system [47, 52, 76]
   (b) AI should learn from user's intervenience [2, 21]
   (c) Provide the user with opportunities to specify their goals. needs and recommendations [45, 83]
6. **Provide personalized explanations**
   (a) Match the fidelity of the explanations based on the task complexity and risk involved in the situation [21, 45, 70]
   (b) Users are more likely to be satisfied with feature based explanations rather than base line ones [4, 45, 47]
   (c) System finds the single best possible item rather than good enough items for efficient personalization and user profiling [92, 93]
   (d) Consider the gains the user might experience when interacting with the system [45, 75]
7. **Develop and show significant, value-added and contextually relevant information**
   (a) Provide the user with the ability to ask context-sensitive questions [2, 29, 35]
   (b) Too much information (global scope) as well as complex information should sometimes be avoided [5, 45, 74]
   (c) Establish a credible source for explanations [74, 92]
8. **Update and adapt to different users cautiously**
   (a) Consider proactively surfacing explanations, when user need is detected [2, 35, 45, 73, 74, 91]
   (b) Effectively switch between different topics and tasks [73]
   (c) Should keep up with the user's pace [21]
   (d) AI need to know how to respond if the user is trying to communicate in a language you cannot communicate back in [21, 35]
9. **Match relevant social norms**
   (a) Taking natural pauses and using varied language to show the understanding of the request is expected for courtesy [5, 45, 74]
   (b) *Be polite* - The system needs to understand the different types of courtesies and how to respond. Equally it should respond politely but firmly when the user becomes abusive to avoid normalizing abusive interactions [37, 65, 84]
   (c) Don't violate fundamental human rights [37, 51]
   (d) Encourage the user to behave in a desired way while interacting with the system [51, 90]

10. **Mitigate social biases**
    (a) AI should not harm or discriminate any user of any demographics [2, 37, 51]
    (b) Don't pretend to be human. Let the users know that they are interacting with a system rather than mimicking to be a human [69, 76]
    (c) Employing socially appropriate behaviors for agent-user interaction [37, 69, 90]
11. **Support efficient AI invocation, dismissal and correction**
    (a) Allowing efficient direct invocation and termination [2, 35]
    (b) Providing mechanisms for efficient agent-user collaboration to refine results [38, 96]
    (c) When the system fails, user needs a way out. There should always be a way to decline/undo any suggestions or changes system makes [21, 76]
    (d) By enabling the user to verify the autonomous steps taken by the system [29, 83, 90]
12. **Scoping precision of service to match uncertainty, variation in goals**
    (a) Engage in disambiguation or gracefully degrade the AI system's services when uncertain about a user's goals [2, 3, 35, 74]
    (b) Use simple, conversational language, that can be easily understood by the user [45, 70, 90, 95]
    (c) To avoid uncertainties, distinguish AI content from regular content [21, 35, 83, 105]
    (d) Use a recognizable element to consistently refer to the AI [105]
    (e) Inferring ideal action in light of costs, benefits and uncertainties [30, 69]
13. **Be accessible**
    (a) Provide global controls by making the system universally accessible [2, 35, 95]
    (b) Offer the user, the ability to progressively get more details on demand [29, 88]
    (c) The system should be controllable and structured the way a driver expect it to be or is familiar with [35, 47, 95]
14. **Let the user govern the AI**
    (a) Give the users the control to decide when to receive explanations [4, 47, 65]
    (b) Users should be able to adjust what AI has learned [21, 29, 90]
    (c) Allow the user to control the data holded by the AI [21, 88]
15. **Don't interrupt services and time them based on context of the interaction - *Speed is critical***
    (a) Consider the status of the user's attention in the timing of services [2, 4, 93]
    (b) As less input from the occupants as possible should be required while travelling [5, 74, 95]
    (c) The information should be comprehensible by as few glances as possible [20, 81]
    (d) Should minimize the cost of poor guesses about action and timing of the services offered [81]

16. **Make AI delightful**
    (a) Design delightful features to increase the likelihood of forgiveness [38, 65]
    (b) Design the tool in such a way that users will forgive it when it makes mistakes [65]
17. **Security and Privacy matters**
    (a) Splitting the information storage to private/public, based on which the confidential info can be stored private and rest of the user profiles can be made public [29, 54, 83]
    (b) Let the user give explicit informed consent for using the AI [37, 96]
    (c) Should use a fail safe to monitor all the changes and suggestion the system makes [21, 54]
    (d) To tackle the challenges in transparency and explainability, system should document both decisions they make and whole process that yielded those decisions to make it more traceable [29, 47, 83]

The above 17 guidelines are drafted in such a way that it significantly reflects the common aspects that the researchers and developers should consider in order to develop a consistent AI-based user interface across all the systems. In the scope of this thesis, these guidelines are going to be used initially to perform a systematic comparison between various interaction choices offered existing automotive HMI's, which are explained in detail in the coming chapters.

3.3 Supporting the understanding of the User

Even though the above 17 guidelines serve a guiding block for the developers and researchers to provide a consistent AI-based interface among all platforms, the biggest task/milestone for the developers is to design the algorithms and interaction in such a way that it enables the user to understand the system easily and effectively. In order not to make the AI/ML system opaque, it becomes a necessity for these systems to satisfy a certain ***system qualities***. In the work done by Eiband et al. [24], they grouped some of the commonly occurring traits of these intelligent AI/ML-based HMI systems like *Accountability, Debuggability(end-user debugging), Explainability, Intelligibility, Interactivity, Interpretability, Scrutability, Transparency*, together into a group and termed them as system qualities.

Even though using these system qualities in designing these interfaces is considered as an essential factor in supporting consistent Human-HMI interaction, there exist a lot of contradicting and diverging assumptions among different developers of how to support user interaction using AI across all platforms. Therefore, it is important to address, or in other words to trace the mental model that the user develops while interacting with the system [100]. The more efficiently the users develop a mental model of the system, the more effectively they can predict and anticipate the behaviors of the system, which in turn increases their understanding towards the system [13, 41]. To achieve this, certain assumptions has to be made by the researchers to predict the behavior of the user for classifying the system qualities for supporting the user's understanding of the system. This can be done effectively based on the framework developed by Eiband et al., which consists of 3 categories that can identify and differentiate of how the user's interact with an AI-based HMI. The conceptual framework taken from [24] is shown in the figure 3.3 below.

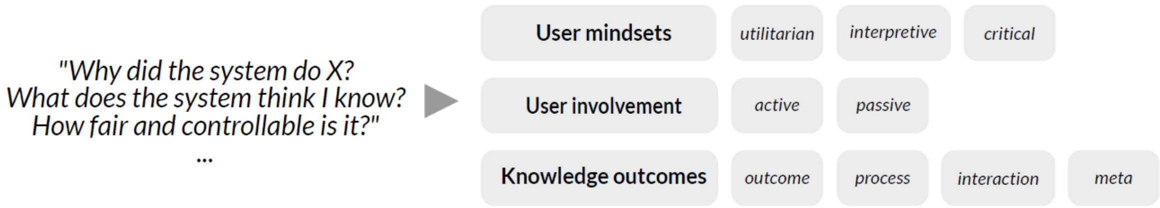


Figure 3.3: Framework (retrieved from [24]) used for supporting the user's understanding with intelligent systems

Figure 3.3 shows 3 categories namely ***Users Mindsets, User Involvement***, and ***Knowledge outcomes***, which are subsequently used for structuring the user's understanding. Based on the information gathered from the user, the above 3 categories can be used a basis of structuring the user's questions for providing effective solutions to those questions developed by the user [41].

1. ***User Mindsets***

   Identifying the mindset of the user is essential for the developers to help to understand what the user seeks to know when they are interacting with the HMI. Three types of mindsets that the user's exhibit during system interaction are given below.

   (a) **Utilitarian Mindset**, which aims to control the system's behavior to reach a particular goal

   (b) **Interpretive Mindset**, which interprets the system's actions based on the user's perception and experience with the system

   (c) **Critical Mindset**, which deals and reflects with the normative and ethical aspects of the system

2. ***User Involvement***

   As Kulesza et al. mentioned [41], the nature of interaction between the system and the user, and how it is supported and manifested through all designs, should be defined during design, so as to make the user feel more involved in interacting with the system. Thus measuring the user involvement is a key factor used for identifying how the users develop and build their mental model while interacting with the system. Two ways in which the interaction can happen between the user and the system are given below.

   (a) Interaction between the User to the System (**Active User Involvement**)

   (b) Interaction between the System to the User (**Passive User Involvement**)

3. ***Knowledge Outcomes***

   While interacting with the system, it is evident that the users gain some knowledge on the underpinnings of the system; however, these knowledge gained might differ from one user to other. To account for this issue, Eiband et al., characterized these different knowledge gains into 4 categories, which are listed below.

(a) **Output Knowledge** targets the individual output instances produced by the system, whereas the **Process Knowledge** identifies the reasoning behind the individual output instance

(b) **Interaction Knowledge** is used to identify how to operate the system in an interactive environment, and the **Meta Knowledge** is used to describe this interaction knowledge in a more generalized environment.

The ways with which the developers and researchers can support the AI/ML HMIs based on the 17 guidelines, is done by mapping these 3 categories in the framework mentioned above to each guidelines, that makes it efficient to find solutions to the potential user's questions that arise during their interaction with the system. The mapping of these guidelines is given in the table 3.1 below.

Table 3.1: Technical Guidelines coded based on the 3 categories to support and structure user's understanding

ID	**Guidelines**	**User Mindset**	**User Involvement**	**Knowledge Outcomes**
1	Explain the user with respect to the capabilities that the AI framework has	Utilitarian	Passive	Output Process Interaction Output & Process
2	Identify, educate and mitigate the potential errors and unexpected edge cases that might occur during user interaction	Utilitarian or Critical	Active and Passive	Process Process Interaction
3	AI should be transparent in what data it has of the user and how it can be accessed by the user	Interpretive or Critical	Active and Passive	Process Meta Interaction Meta Meta
4	Enable the user to provide feedback by indicating their preferences during regular interaction with the AI system	Utilitarian	Active	Interaction Process Interaction
5	Maintaining the working memory of recent interactions by learning continuously from user behavior	Utilitarian	Active	Meta Output Interaction
6	Provide personalized explanations	Interpretive	Active	Interaction Output Output & Process Process
7	Develop and show significant, value-added and contextually relevant information	Interpretive	Passive	Interaction Output & Process Meta

8	Update and adapt to different users cautiously	Utilitarian or Interpretive	Active	Interaction Interaction Interaction Interaction
9	Match relevant social norms	Critical	Passive	Meta Meta Meta Meta
10	Mitigate social biases	Critical	Passive	Meta Meta Meta
11	Support efficient AI invocation, dismissal and correction	Utilitarian	Active	Interaction Interaction Process Output
12	Scoping precision of service to match uncertainty, variation in goals	Utilitarian or Interpretive	Passive	Process Interaction Process Output Meta
13	Be accessible	Utilitarian or Critical	Active	Meta Process Interaction
14	Let the user govern the AI	Critical	Active	Interaction Interaction Interaction
15	Don't interrupt services and time them based on context of the interaction - Speed is critical	Interpretive	Passive	Interaction Output Output Interaction
16	Make AI delightful	Critical	Passive	Process Process
17	Security and Privacy matters	Utilitarian	Passive	Process Output Meta Process

3.4 Remarks

On the whole, using these coded guidelines in designing the system enables the researches to draw connection between the system qualities across existing and yet to come products, thereby arriving at a clear goals and underlying assumptions. As said in the beginning, these guidelines are a first step into developing a more comprehensible AI-based HMI, and are subsequently used as a measure of comparison between various existing automotive HMI's available, and finally to arrive at the necessary requirements for designing the prototype.

Chapter 4

Understanding the Systematic Differences and User Needs

Nowadays, most of the automobile manufacturers and other third party companies, design the user interfaces for the vehicles in such a way that these are mostly used to bring the full potential of the various functions offered by the vehicle, and also to utilizes the design elements of the previous UI version, so that users would not get confused while interacting with it. Even though this might be useful for the case of returning customers, the same cannot be said for newcomers as they are completely unfamiliar with the new UI in the car, and they expect some functions to be same as they experienced in their old vehicles. This problem of inconsistent UX exists in all kinds of products, in addition to automotive HMI. A recent industry survey yielded that maintaining consistency with the UX has become a global challenge[1]. This mainly due to concept of scaling in the industry (even in the automotive HMI stand), where the companies release it's first version and provide updates to its UI consistently, so that it remains in the market for a long time, without thinking for the need of a complete design overhaul [8].

Due to this product scaling, the collaboration and communication between various stakeholders becomes limited, thereby leading to lack of shared vision between them when involved in the product development phase. In addition to facing the problem of scaling and innovation, inconsistencies in the system often lead to complicated workflows and also places an increasing amount of load on the developers. Thus, there is a need for eliciting the requirements that the designers and developers can follow to produce a consistent user interface across all the systems by adhering to guidelines mentioned in the chapter 3. Most of the requirements for the prototype can be arrived easily by doing some market analysis and interviewing the users; however, there exist some instances where the inconsistency problem still persists within the design and development [62], and the same can be said for various automotive user interfaces also. Therefore, according to this thesis scope, it is of essential to find the inconsistencies within various existing automotive HMI's that are available or in development, along with some patents that are filed by the industry that either published or has a pending status. These are achieved by using the guidelines described in the section 3.2 of the chapter 3, as a measure for performing the systematic comparison, thereby arriving the requirements which can potentially be used to solve these inconsistency problems.

[1] https://uxplanet.org/improving-ux-consistency-1dcaadbec783

4.1 Systematic Comparison of the Automotive HMIs

In order to get a holistic picture of the core thesis topic and to find the inconsistencies in the types of design and interactions, a systematic comparison of ***various interaction choices*** offered by the system is done within various automotive HMI's and patents in order to arrive at the functional and design requirements of the user interface. Here, the HMI's of some of the cars exiting in the market and release on or after 2016, and also the ones which employ and offer a significant amount of AI for improving Human-HMI interaction, were considered for this systematic comparison, and the same goes for the patents that are filed by the companies. This selection criteria resulted in a total of 14 viable automotive HMIs that exist in the market, along with 8 patents. In order to avoid ambiguity, the systematic comparison was done separately for existing HMIs in cars and for patents, and the results were then combined together to elicitate the requirements. The hypothesis based on which the systematic comparison is conducted and its expected outcomes are listed in the table 4.1.

Table 4.1: Research (RH) and Null (NH) Hypothesis for the systematic comparison

Code	**Hypothesis**
RH_{SC}	The systematic comparison between the existing HMIs and patents yielded several inconsistencies in regards to their design, interactions and internal workings of the system
NH_{SC}	The systematic comparison between the existing HMIs and patents yielded in an overall consistent design choices, interactions offered and internal workings of the system

Expected Outcomes

1. The interfaces under comparison are designed in such a way that either they do not even incorporate any guidelines or moderately incorporate some of the guidelines for Human-HMI interaction.

2. More than half of the HMI's of the vehicles under consideration has the technical guidelines moderately applied in their architecture, instead of solely being designed upon them. For patents, the reverse of this outcome is expected.

3. It is expected for some very minor population of the HMIs, not have been incorporated with some of the technical guidelines.

4.1.1 Approach

The systematic comparison between the existing automotive HMIs available in the market is performed not directly acquiring the car, but using a platform developed by ***Screens studio***[2]. This platform is a database containing videos of every possible Human-HMI interaction and every features and settings that the car has and has to offer. The navigation structure in their database in structured in such a way that it reflects the same type structure found within the

[2] `https://screens-studio.com/`

car's HMI, and this structure also makes the user to find a certain settings and feature faster than interacting with the HMI in-person.

4.1.2 Systematic Comparison - Existing automotive HMIs

For this systematic comparison, HMIs from 14 existing cars around the market, which were revealed or release on or after 2016, are considered. The percentage of the guidelines that has been fully applied, moderately applied and not applied is shown in the table 4.2 below. The entire systematic comparison table can be found in the appendix B. For coding purposes, the names and brands of the cars are not revealed in the table found below and in the appendix; instead they are coded by the number representing the order it was reviewed and analyzed.

Table 4.2: Systematic comparison of existing Automotive HMIs

	Strongly Applied	**Moderately Applied**	**Not Applied**
C1	41.2%	58.5%	0.0%
C2	11.8%	82.4%	5.9%
C3	47.1%	47.1%	5.9%
C4	29.4%	70.6%	0.0%
C5	47.1%	52.9%	0.0%
C6	11.8%	58.8%	29.4%
C7	47.1%	52.9%	0.0%
C8	47.1%	52.9%	0.0%
C9	58.8%	41.2%	0.0%
C10	47.1%	52.9%	0.0%
C11	41.2%	58.8%	0.0%
C12	35.3%	64.7%	0.0%
C13	11.8%	47.1%	41.2%
C14	11.8%	88.2%	0.0%

4.1.3 Systematic Comparison - HMI Patents

As mentioned in the chapter 3, some of the guidelines were already drafted based on some selected patents that are relevant to the selection criteria. For performing the systematic comparison on the HMIs whose patents are either granted or still pending, a more comprehensive selection has to be done, and this was done through the process of ***Patent Landscape Analysis***[3]. Similar to existing automotive HMIs, the patents are selected on the basis of either

[3] https://www.ipcheckups.com/blog/patent-landscape-analysis-how-to-5-steps/

for the patents are approved or still pending as effective from 2010. The total systematic comparison of the 8 HMI patents based on the technical guidelines is elicited in the table 4.3 below. Refer appendix B for the full comparison table. The patents here are also not expressed by the industry that hold/filed the patent, but through the order it was compared.

Table 4.3: Systematic comparison of HMI Patents

	Strongly Applied	**Moderately Applied**	**Not Applied**
P1	35.3%	47.1%	17.6%
P2	47.1%	23.5%	29.4%
P3	47.1%	29.4%	23.5%
P4	58.8%	35.3%	5.9%
P5	52.9%	17.6%	29.4%
P6	52.9%	23.5%	23.5%
P7	58.8%	23.5%	17.6%
P8	58.8%	11.8%	29.4%

4.1.4 Observations

From the systematic comparison performed on the 14 existing automotive HMIs (refer table 4.2), almost 50% of the technical guidelines are only moderately applied in 11 out of 14 HMIs, thereby leading with the assumption that only the guidelines that are relevant to the construction and architecture of the HMI are being moderately considered, and rest of the guidelines might be merely kept as a statement of reference, which can further be addressed at the later stages of the software life-cycle. There might also be another case where some of guidelines are strongly taken into consideration at the beginning of the design phase itself, so as to not touch them at later parts of the product life cycle.

The later case also poses a significant problem because, as the products are being updated continuously, the core guidelines that are strongly applied at the beginning, might not get any updates and becomes redundant. Taking a look into the car C9, which by comparison showed that up to 59% of the technical guidelines are strongly applied, but the percentage of moderately applied guidelines is also significantly high. Since the products are being scaled nowadays [8], the above said problem might also occur in the HMI of the car C9 (since the car had only being launched just recently), and also for the rest of the cars.

On the other hand, performing the systematic comparison on some of the patents yielded the results, which is reverse of the outcomes that has been found while doing comparison on the existing automotive HMIs, which in this case is an expected outcome. From the table 4.3, it is clearly evident that the patents are filed by the specific industry in such a way that it addresses the commonly found inconsistencies found in the HMI (this is due to 8 out of 9 patents clocking up to or more than 55% of the guidelines that are strongly applied), but it is still is an uncertainty whether these will be addressed at the product launch itself, or at the later stage due to product scaling. Finally, the biggest difference in terms of comparison is that, except for a small population of existing car HMIs, almost all of them had the guidelines that are either fully mapped or moderately mapped to their construction, whereas all of the patented HMIs did not even map some of the technical guidelines during development.

It is natural to assume that the automotive HMIs that have been patented and not out on the markets will have all the guidelines either fully or moderately mapped to them during development. This was not the case based on the results that yielded from the systematic comparison. Also from the tables 4.2 and 4.3, the mapping of guidelines across all the HMIs are extremely inconsistent, thereby satisfying with the research hypothesis. This problem will also exist prominently among the users of these automotive HMI, since in spite of it being developed by mapping all the technical guidelines, the needs and expectations, along with their mental models, will be different for different users. Thus, to minimize the problem of inconsistency in user interfaces, mapping the technical guidelines along with the consideration of the needs and expectation of the user would be a viable solution.

4.2 Understanding the User's Needs and Expectations

The sufficient guidelines that the developer has to follow to ensure an efficient interaction between the human and AI-based HMI is elicited and compared, and the resulting inconsistencies from that comparison can be further addressed by addressing the needs and expectations of the user. Identifying the user's needs and expectations, is in itself is not a simple task, and the way these can identified is gathering their opinions about AI-based HMI and AI in general through their eyes. For this purpose, a survey with $N = 50$ participants (26 Male, 24 Female) was conducted, which was aimed to identify the opinions of the people about AI, and also their impression about its internal workings.

For this survey, the participants were selected based on age (max 30 years) and also was temporarily restricted to the area within Europe. Since there were not enough responses, around 5 responses were collected from India. Since this survey was planned to be distributed only within Europe, the 5 participants from India are selected in such a way that their responses will not pose a higher significant difference when analyzed. The survey totally consisted of around 55 questions and was divided into 3 blocks. The entire summary of the survey's questions are provided in the appendix C.

4.2.1 User's opinion about AI and its usage

Before diving deep into identifying the needs and expectations of AI-based HMI under consideration, knowing what the public think of AI in general, their opinions and how the perceive while interacting with those technologies in their day-to-day life, poses a significant importance in identifying their needs and expectations. Hence, it is essential for the survey to find out the public's opinions on the following research questions. Since, the thesis involves around automotive HMIs, some general questions were also asked about self-driving cars and their knowledge about it, in addition to some common driving related questions. One important thing is that these questions are designed in such a way that these are easily understandable, unbiased, and concise.

1. Does everyone know and understand what Artificial Intelligence is?
2. How often do people think that they interact with AI-based systems and programs daily?
3. Does the concept of AI make people scared and anxious about using them?
4. To where does the people think that AI is headed in the future?

From the survey, it was somewhat surprising to see that even though almost 80% of the participants did not take any classes or courses about AI, they possessed an substantial amount of technical background (recording a score of ***2.98*** out of ***5***) regarding AI in general (refer figure 4.1 for scores distribution). This due to fact that, the participants kept on hearing news about AI and its several use cases through the media several times a month or even a week, which is represented in the figure 4.2 below.

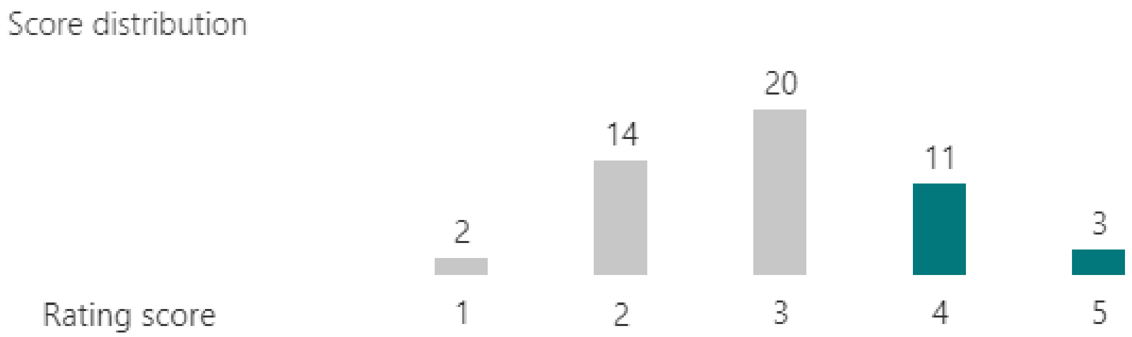


Figure 4.1: Score distributions for the AI technical background ratings

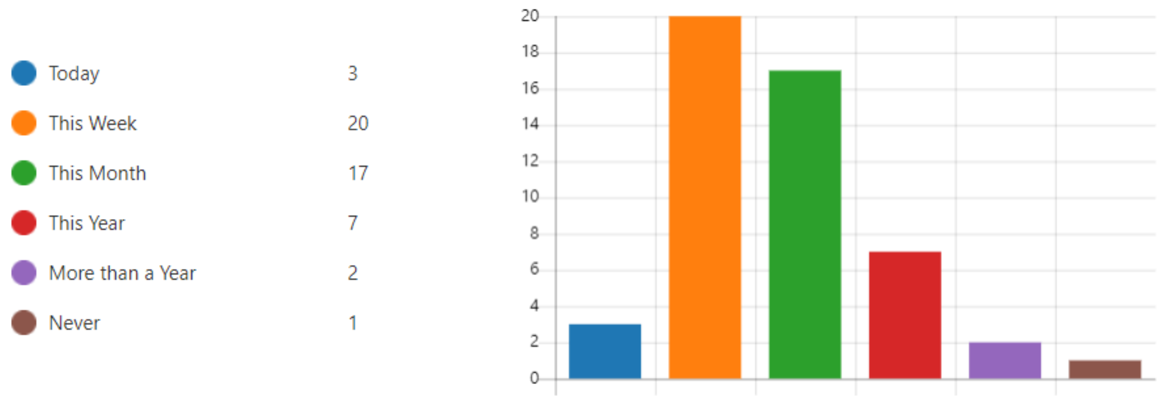


Figure 4.2: Histogram representing the amount times that the people heard about AI in the media (TV, newspaper, radio/podcasts, internet, magazines)

Another key aspect is to determine whether the people interact with these AI-based systems often, and based on the figure 4.3, up to 58% of the participants interact with these systems on a regular basis, and around 13% of the participants interact with these AI-based devices sometimes, and the rest of 16% of the participants never interacted with them.

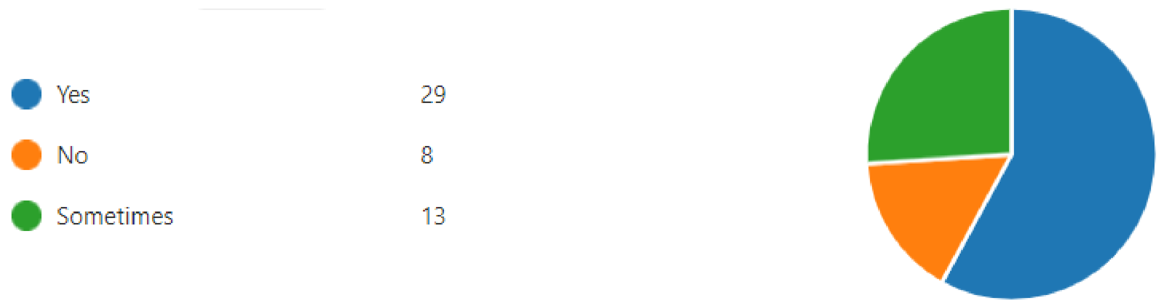


Figure 4.3: Pie chart representing the amount of participants that interact with AI-based devices

In addition to this, the participants were also asked to name an example of AI to identify what kind of AI-based devices that they encounter(ed), and the resulting raw responses are mapped into a word cloud and is shown in the figure 4.4 below.

Figure 4.4: Word cloud representing the raw responses for an AI example given by the participants

Since it is evident from the survey that the participants find affinity towards interacting with AI-based devices as almost 64% of the participants don't fear in the prominent application of AI in all devices, and more than 50% of three participants believe that the AI-based system could replicate human intelligence in machines in the future, and also with the possibility of AI replacing humans in the future.

4.2.2 User's opinion on interacting with the in-vehicle HMI

Before being asked about their perception towards interacting with AI-based HMI in general, the participants were initially asked about the different kinds of non-driving related activities that they tend to do while driving, and the resultant raw responses are converted into a word cloud which is shown in the figure 4.5 below.

Figure 4.5: Word cloud representing the participant's raw responses about their non-driving related activities

As expected, almost all of the participants responded as listening to music/radio/podcasts as the major non-driving related activity for them, followed by their interaction with the navigation system, and also reading books. The responses that were obtained from the survey, was also in-line with the research done by Pfleging et al., [67]. Getting to know what the participants wanted as the primary non-driving related activity is a key factor considering the user's needs and expectations.

Since the previous block tackled the opinions about the passengers about AI in general, the next part of this survey were designed to deal with some specific functionalities of the AI-based HMI. For this reason, the participants were initially, the type of interaction system that they prefer interacting with during driving. Up to 44% of the participants preferred interacting with the systems developed by the car manufacturer itself, and the rest of the participants preferred either Android Auto (34% of the participants) and Apple Carplay (16% of the participants) as their preferred HMI user interface.

In addition to this, it is also essential to identify the way that the participants preferred while interacting with the system. The survey conducted by Pfleging et al. [67], yielded that most of the respondent preferred to interact with the system using a touch screen. This was subsequently followed by speech input, and by operating the buttons on the steering wheel. The same questions was also asked to the participants in the survey who are instructed to rank their preferred input modality. In this case, most of the participants preferred to use the voice input the most, as shown in the figure 4.6, followed by touch input as the second choice and then the gesture input as their least preferred way.

Figure 4.6: Input Modalities ranked in the order preferred by the participants

Also for an effective interaction, most of the participants felt that the size of the central infotainment should have a significant high blueprint with an effective average size range varying within ***8 to 11 inches***, and it should have a size sufficient and large enough to enable the person on the other side of the seat to interact with the HMI without causing any discomfort between the occupants. This point is also one of the important factors, since almost 96% of the participants as shown in the figure 4.7, travel as a passenger when they are not currently driving.

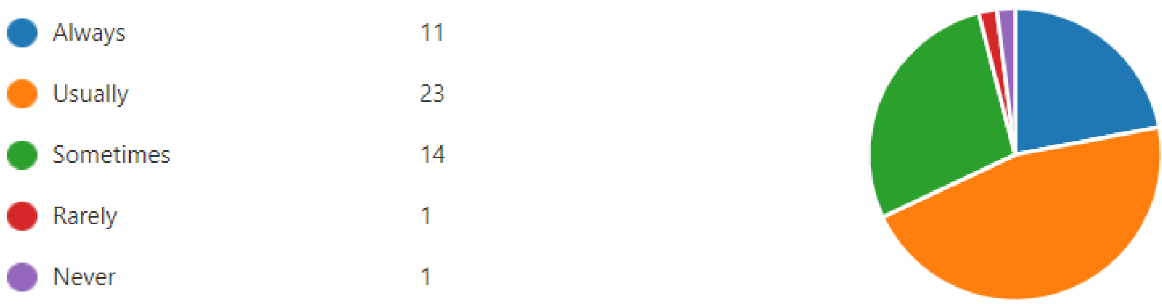


Figure 4.7: Pie chart representing people's preference to travel as a passenger

4.2.3 Users opinions about different HMI Functionalities

From the subsections 4.2.1 and 4.2.2, the user's general understanding about AI in general and also their opinions about interacting with an in-vehicle HMI is identified; but it is also essential to identify their perception towards the HMI regarding its various functionalities offered. Based on the participants raw responses for their preferred non-driving related activity (refer figure 4.5), along with considering the research done by Pfleging et al. [67], the following use cases for this thesis are selected. These use cases are identified to have the potential of improving the interaction between the vehicle's occupants and its AI-based HMI.

4.2.4 Use Case 1 - Destination Entry

Before delving deep into developing the functionality of this use case, first the preferred navigation systems/devices that the vehicle occupants constantly use while driving has to be determined. When this question is asked in the survey, 26 (52%) respondents preferred to use their smartphone's navigation features as their primary navigation device (refer figure 4.8), which is expected, and the rest of the participants either preferred to use the vehicle's on-board navigation system (13 respondents) or maps from third party applications like Android Auto or Apple Carplay (8 respondents).

Figure 4.8: Pie chart representing participant's preferred Navigation device

The main reason for people using their smartphones as their primary navigation devices is their familiarity and its ease of use; however usage of on-board navigation system poses a significant advantage over external devices, and one of the biggest advantage is the screen size. As mentioned in the subsection 4.2.2, most of the participants perceived that a bigger screen in the car will improve the interaction, and also the bigger the dimension of the screen, the fewer the glares will be, thereby reducing distractions. This also corresponds to the participant's responses as shown in the figure 4.9, where most of them preferred the navigation system to be displayed either only on the central display or on both of the central display and the instrument cluster.

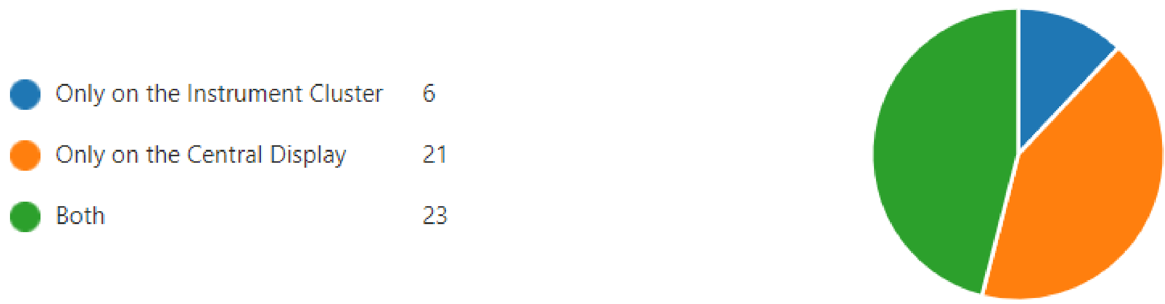


Figure 4.9: Pie chart representing participant's preferred way to display the navigation system

In addition to this, many people fear of losing the signal or internet connection while driving to a remote/unknown location. This can be easily mitigated since many of the present on-board navigation system uses GPS which has a better signal receiving quality. For this purpose, the user experience and their subsequent interaction with the on-board navigation system is measured in this thesis scope.

User's Preferences

Since the main objective of considering this use case is to improve the interaction between the vehicle's occupants and the HMI, even when the vehicle is stationary and also during driving. When destination entry is concerned, almost 43 respondents (86%) only enter the destination when the vehicle is stationary, and not during driving. This problem still persists even now, as the user takes up to ***33 keystrokes*** [48] which makes changing the address while driving almost troublesome. Therefore, the user interface must be designed in such a way that it encourages the user to enter the destination with as less keystrokes as possible. In order to achieve this, the way with which the user prefers to enter a destination has to be determined. So, the possible ways that the user might enter a destination are given below.

1. By entering the address of the street (number/name, city/town/village, state, etc.)
2. By entering the zip code of the location
3. By identifying and entering a nearby Point Of Interest (POI)
4. By selecting the previously entered destination points and POIs
5. By directly selecting the place by pinning the position on the map
6. By using a voice assistant

So, in order to find out the most and least preferred way for the user's to input a destination, the participants in the survey were asked to rank their preferences in the order of the most preferred option at the top and the bottom representing the least preferred option. The final rankings of the 6 possible ways is shown in the figure 4.10 below.

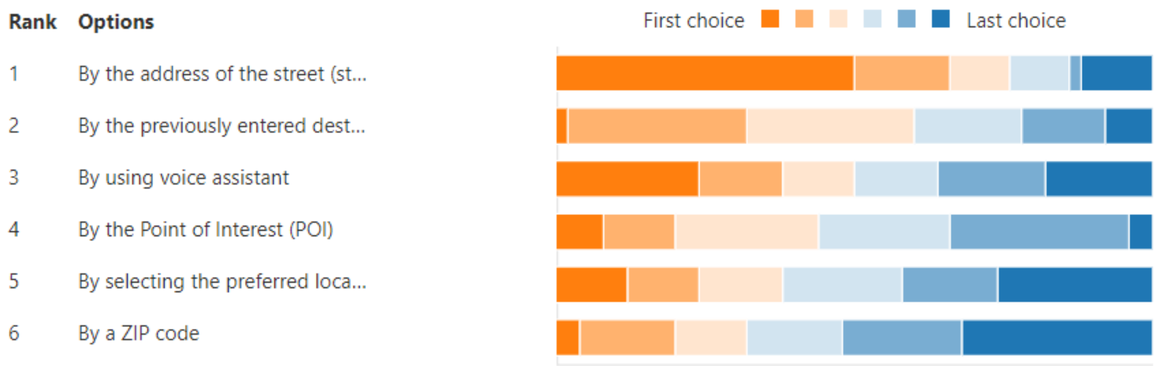


Figure 4.10: Participant's ranked responses indicating their preferred destination entry

From the rankings, it can be seen that entering the entire street address is most preferred way of entering the destination, and as mentioned in the previous paragraph, the problem of requiring a lot of key strokes to enter the destination is prominent in this case. Entering the destination through voice input was ranked second in overall cases, which is expected since the participants preferred the voice input modality the most while interacting with the display (refer figure 4.6). This is in fact due to to growing popularity of voice assistants due to them being constantly improved and upgraded every year. As per the market research done by ***Infopulse***[4], the user base for these voice assistants are expected to increase significantly, which is illustrated in the figure 4.11 retrieved from Infopulse below.

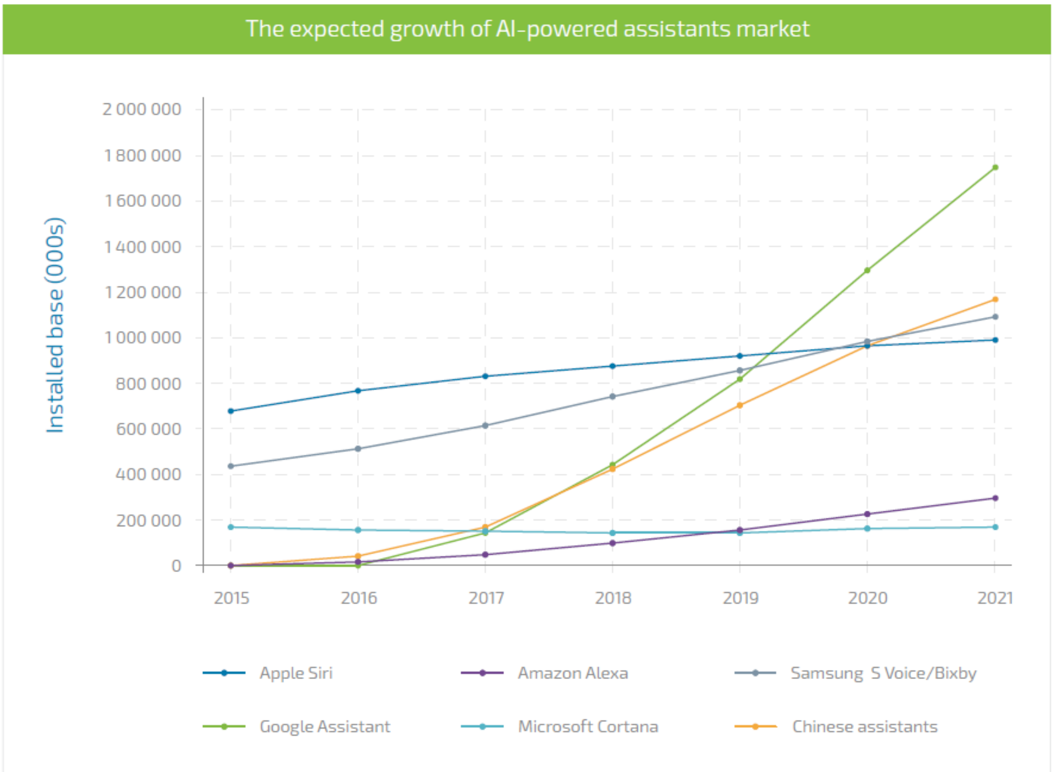


Figure 4.11: Expected growth of AI-powered assistants in the market

When speech input is taken into consideration, most of the research showed that, the speech interface in vehicles have a very low adoption rate due to some accuracy issues and errors perceived by the users. When these types of errors occur, the task completion time becomes very high which often leads to catastrophic conditions during driving [82]. In order to corroborate, the 14 HMI's that has been considered for the systematic comparison, is then checked for the occurrences of some of the commonly persisting speech recognition errors, which is elicited below.

1. Incorrect user input recognition
2. No match was found based on the user's input
3. Rejection of the user's input
4. Invalid command recognition (Spurious)
5. Frequent occurrences of "spoke too soon" prompt
6. Deletion of the user input
7. Insertion of the user input

[4] https://bit.ly/34gV7L4

Keeping the exact coding scheme as mentioned in the subsection 4.1.2, these 14 patents are checked for the above mentioned common speech recognition error occurrences using Screens Studio, and the resulting comparison is shown in the figure 4.12 below.

	SPEECH RECOGNITION ERROR OCCURANCES	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	
1	Incorrect recogntion	11	9	7	0	5	15	9	2	0	5	10	8	10	16	107
2	No match	6	4	6	7	2	3	1	1	0	7	6	8	3	13	67
3	Rejection	8	6	5	18	11	11	10	7	0	5	3	4	5	3	96
4	Timeout	1	0	1	0	0	0	2	2	0	0	8	7	2	2	25
5	Invalid command or other sounds recognized as a valid command	2	3	2	5	4	8	3	1	0	5	6	5	12	1	57
6	Spoke too soon	3	3	1	1	2	3	1	2	0	4	0	0	2	0	22
7	Deletion	1	1	0	0	0	2	0	0	0	0	0	0	0	0	4
8	Insertion	1	0	0	0	0	1	0	0	0	0	1	1	0	0	4
	Total	33	26	22	31	24	43	26	15	NA	26	34	33	34	35	382
	Total use cases analyzed in screens	99	60	70	52	50	104	82	41	NA	69	52	51	64	64	
	Error Occurrance Percentage	33%	43%	31%	60%	48%	41%	32%	37%		38%	65%	65%	53%	55%	
	SPEECH INPUT IMPLENTATION LEVEL	MED	MED	MED	LOW	MED	MED	MED	MED	HIGH	MED	LOW	LOW	LOW	LOW	
	GESTIC INPUT IMPLEMENTATION LEVEL	NA	NA	HIGH	NA	NA	NA	NA	NA	HIGH	NA	NA	NA	LOW	NA	

Figure 4.12: Existing HMIs speech recognition error occurrences

From the above figure, almost all of the HMIs had over 30% of these speech recognition error occurrences, and the levels of implementation of these speech recognition and its subsequent error correction is also low.

Summary

Thus, the major objective going forward is for the effective implementation of AI, to reduce the time taken by the users to input the data into the navigation system through continuous gathering of the user's location data and commonly entered destinations and POI's, and by facilitating interaction through various input modalities with ease. Incorporating this reduces the number of keystrokes required to input the destination, and also boost's confidence for the user's to enter the destination while driving.

4.2.5 Use Case 2 - Customizing the HMI settings

Even though most of the participants believed that the promise of AI application did not scare them, they did not want the system to take over all the controls, but they hoped that the system will give the control of their interaction to the user's without any invocation. This will allow the user to be in control of their interaction; however, some of the users does not really know what they want to customize, and sometimes basically showing a lack of interest in tweaking the settings.

Using settings to customize the user interface is in itself is extremely inconvenient for all the users, and based on the responses from the survey most of the participants ***rarely*** or even sometimes ***very rarely*** customize the features of the HMI's driving and system settings (refer figures 4.13 and 4.14 for scores distribution). Even though the participants used these settings rarely, the participants were also asked whether they have used every single possible feature provided to them, and as expected from the above statements, almost 86% participants felt they are either not sure whether they interacted with every settings or they didn't care to check out these settings. This might be due to the HMI's complicated menus though which the user has to navigate through multiple levels to access the desired functionality [26].

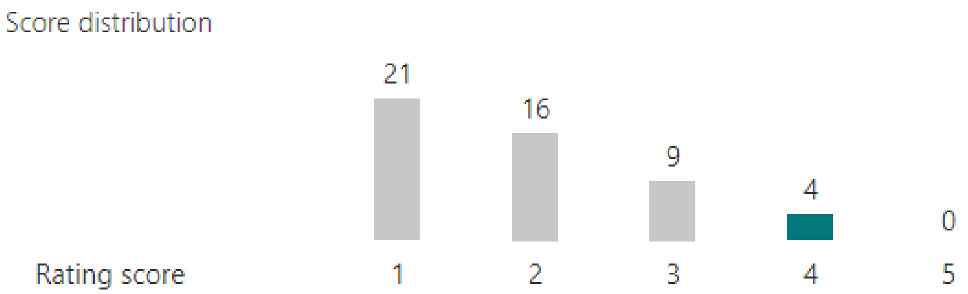


Figure 4.13: Participant's frequency-of-use of Driving Settings

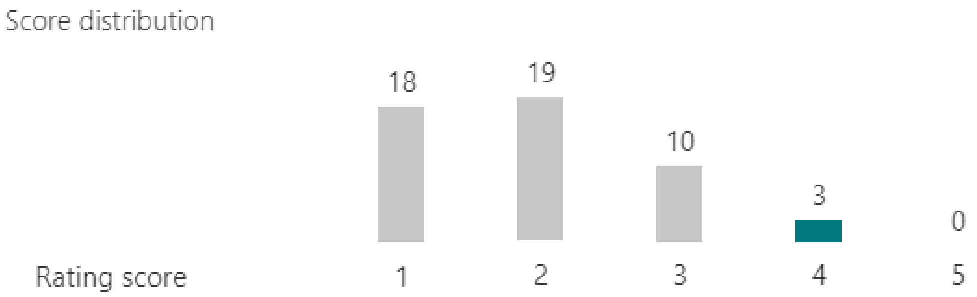


Figure 4.14: Participant's frequency-of-use of System Settings

Therefore, it is essential for the system to provide the settings and features that they tend to use more often without altering any of the primary driving related activities. Another necessity is the type of design and layout that the user's prefer while interacting with the in-vehicle HMI. For this purpose, the three types of design choices that are most commonly looked when customization is considered, is given in the table 4.4 below.

Table 4.4: Three common design choices for UI customization

Individualization	All the features that you tend to use more often are presented on the screen every time you start the car
Style	Offering a wide variety of styles to suit the mood and condition of the driving
Minimalism	Simple in design, and only essential features that are being used more often is provided directly. Some of the less and unwanted features are either removed or hidden

In order to identify the design from the user's point of view, the description of these 3 types of design choices are provided to the participants, and were asked to give their choice on their behalf. Based on their responses, half of the participants seem to incline towards the ***individualistic*** design and the other half of them are respectively leaning towards ***minimalistic*** design. Only a very small population of the respondents (20%) preferred ***style*** as the major design choice over other 2 considerations.

It is also important to consider how often the user's change the settings when they are willing to, and also whether they want all setting or only the specific setting that they tend to use more often to be displayed for them. When the participants were asked to provide their opinions about this, surprisingly most of participants wished that the system showed the entire settings cluster instead of showing only the settings that are related to the specific context. Thus, a perfect balance has to be maintained between the 3 types of design choices and also different kinds of settings that the system wishes to provide the user for changing during driving. Simply to get an understanding about the settings that the user's want to customize, the participants were asked to provide their responses, and the resulting raw responses are mapped as a word cloud and presented in the figure 4.15 below.

Figure 4.15: Word cloud representing the participant's raw responses about the settings they tend to change often

4.3 Remarks

The main objective behind the systematic comparison of HMIs of existing cars and patents, was to find several inconsistencies in the core guidelines that the interfaces are based on, and also to find sufficient trends in application of these guidelines. Once these inconsistencies are identified, a survey was distributed to gather the peoples opinions about AI in general, they way that they prefer to interact with the HMI, and their perception towards different functionalities offered by the HMI. Analyzing and summarizing these findings, inherently leads to development of various requirements that can be used as a guiding factor for UI development, which are explained briefly in the next chapter.

Chapter 5

User Interface Design and Development

The user's opinions and their preferences for their perceived interaction with the AI-based HMI, as mentioned in the previous chapter 4, forms the basis of the AI-based user interface development, thereby leading to elicitation various functional and non-functional requirements. As mentioned in the introduction of chapter 4, some inconsistencies are bound occur when drafting requirements for the system, since interaction design and software development require different skills, and are often done by different people. For this purpose, the user's opinions are identified first so that a clear understanding of the system's structure can be identified, which will lead to a usable system, thereby improving the experience of the user towards interacting with AI-based HMI. To measure this, an interactive prototype representing an AI-based HMI is designed by considering the guidelines, requirements, and mapped interactions, and is subsequently deployed for user testing for sole purpose of analyzing its impact on Human-HMI interaction. The development of the UI for this prototype is explained in the coming sections.

5.1 Functional and Non-Functional Requirements

During the requirements elicitation, it is necessary to categorize the requirements based on overall thesis goal, and this categorization is done based on the ***FURPS*** model proposed by HP [23], which stands for *Functionality, Usability, Reliability, Performance and Supportability.* These are done on the basis of supporting the ***traceability*** of these requirements with respect to other design components. The different categories of requirements that are elicited are drafted based on the following syntax [16].

$$[Condition]\ [Subject]\ [Action]\ [Object]\ [Constraint] \tag{5.1}$$

Considering the FURPS model mentioned above, 4 types of requirements were drafted using the syntax 5.1, which are inherently based on the systematic comparison HMIs, user opinions about AI and interaction with AI-based HMI. The requirements along with their subsequent ID's are listed below and in the coming subsections.

5.1.1 HMI Interaction Design Requirements

Table 5.1: HMI Interaction Design Requirements

ID	Requirement Specifications
D1	The HMI should ensure that the information presented through the vehicle's visual medium should minimize the driver's distraction from the road, subsequently maximizing the driver's performance
D2	The HMI should minimize verbosity by keeping all call-to-actions, menus and all text to a minimum
D3	The HMI should reduce the number of call-to-actions by minimizing the glance time, frequency and task completion time
D4	The HMI should incorporate lesser number of task types, menu layers and clickable elements
D5	The HMI should be designed considering the reachability of the screen by the vehicle's occupants and its subsequent readability
D6	The HMI should enable the user to clearly identify the primary and secondary actions of the function/task (i.e) the affordances of the HMI should be visible
D7	The HMI should not entirely depend upon the voice input and should lean towards incorporating all modalities
D8	The HMI should be designed in such a way that it reduces the amount of typing done by the user
D9	The HMI should consist a clear and minimalistic design without cluttering of all the available features
D10	The HMI shall include training modules of the necessary features and clickable items, only if the user needs it and/or log's in to the operating system for the first time

5.1.2 HMI Functional Requirements

Table 5.2: HMI Functional Requirements

ID	Requirement Specifications
F1	The HMI should store the recorded user data within the car's memory system and should be exported elsewhere
F2	The HMI shall access the stored data history of the user's which can be used further for facilitating interactions

F3	The HMI shall include a dark mode in addition to the normal mode since it has been a most requested feature in recent times
F4	The HMI should not have text with a font size less than 6mm and the text should be at least 34 to 38 pixels tall
F5	The HMI should lay more focus on media, calling and navigation when compared to other features
F6	The HMI shall prefer the functions to be represented in the form of icons over text labels, if the actions of the functions are clear and well-known
F7	The HMI should maintain the contrast of the display above the minimum ratio of 4.5:1
F8	The HMI should include the snapping effect when gestures are implemented

5.1.3 Navigational Design Requirements

Table 5.3: Navigational Design Requirements

ID	Requirement Specifications
N1	The navigational system shall convey the distance between the current location and the destination without the option of enabling the route guidance
N2	The navigational system should combine all the possible input modalities for enabling interaction with the system (preferably audio and touch - gestures are less common)
N3	While entering the destination, the system shall prompt the user to enter/say the street name first followed by the city name
N4	When voice input modality is considered, then the system should communicate/respond to the user in their native language thereby increasing it's adoption rate
N5	The system shall display the map both on the head unit and the instrument cluster (IC), where the IC is restricted to only the basic functions, and the remaining functions are placed on the head unit
N6	The system shall create a separate entry in the recent destinations card called the Address book to keep tabs of the recently visited contacts 'address separately from the frequently visited POI
N7	The system shall enable the user to select 3 diff types of routes based on the entered destination (faster time / shorter distance / less fuel)
N8	Design of the system should mimic the design on those on smartphones which in-turn provides a familiar experience to the users.

5.1.4 HMI Customization Requirements

Table 5.4: HMI Customization Requirements

ID	Requirement Specifications
C1	The system shall provide customization options in such a way that it requires the minimum keystrokes possible
C2	The design of the system shall be minimalistic in nature but has to provide the user with an easier access to their frequently used features
C3	The system shall enable the user to revert back to the original settings after some customization in addition to retaining the previous data
C4	The system shall display the most commonly used settings to user on the main card, keeping rest of the least used settings in the second menu level
C5	The system shall offer the customization settings to an extent that it does not have a significant influence of the stock design
C6	The parameters that determine the level of customization provided by the system should be made adjustable by the user

5.2 Mapping the User Interaction for the Use Cases

Since there are multiple available input modality (*touch, speech, gesture*) for the user to choose while interacting with the user interface, mapping different available interactions to the preferred input modality before beginning the design phase itself provides a clear idea about the cognitive model that could not only satisfy the requirements but also the technical AI-guidelines. The mapping of these interactions are done in such a way to map the preferred/alternate kinds of modality that the user performs the intended interaction, and also the way that the system reflects upon the interaction and conveys the intended information to the user through a preferred/alternate output modality. Drafting such a cognitive HMI model, enables the developers and researchers to determine when to use the appropriate input and output modalities.

Furthermore, it is also essential to identify the steps that the users take and the goals that the user's tend to form, while executing the tasks that the use cases has to offer. The user's attention might also shift constantly while executing these tasks. So, the frequent shift in the user's attention from the center to the periphery back to center is also measured [6, 59]. This attention shift executed by the users will also yield the most significant way of conveying information to the users during interaction. Thus, the mapped interaction of the use case along with subsequent Human-Action cycles are shown in the figures 5.1 and 5.2 below.

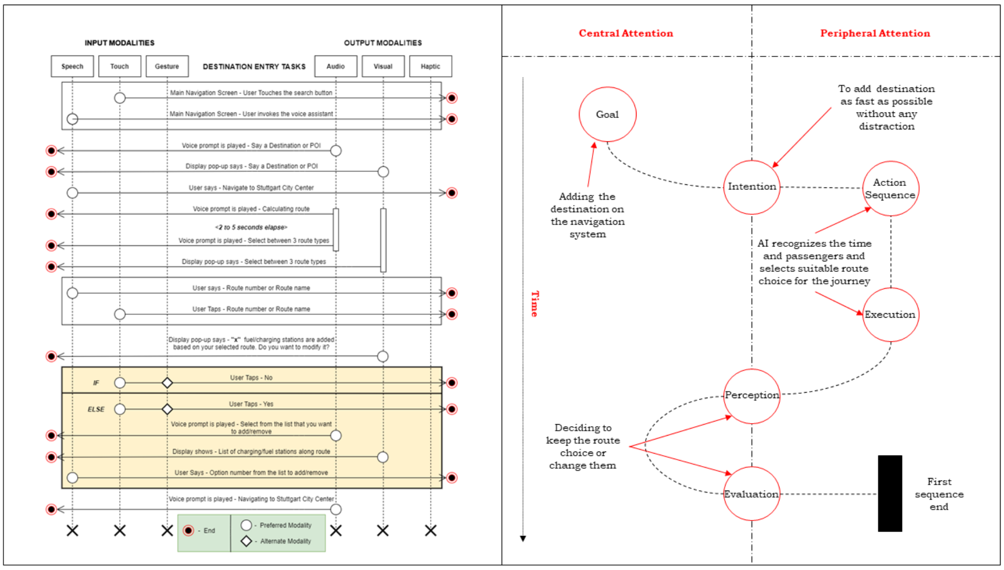


Figure 5.1: Interaction Mapping and Human-Action Cycle for Use Case 1

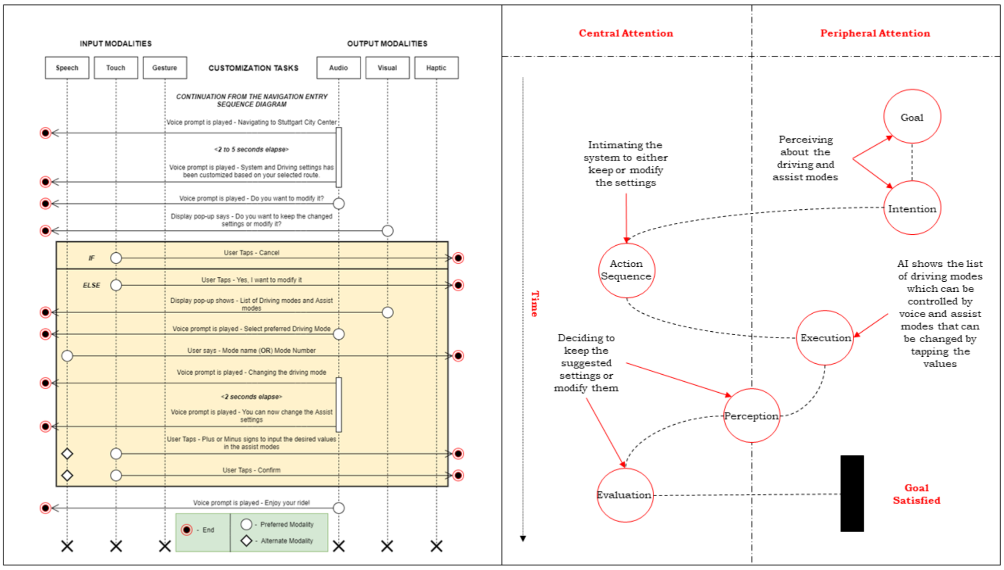


Figure 5.2: Interaction Mapping and Human-Action Cycle for Use Case 2

The left side of the above 2 figures represent the sequence diagram which represents the potential sequence of operations that happens when the user's interact with the system, and the right side of the figure shows the shift in the user's attention while performing the tasks with the help of the modified Human-Action cycle developed by Bakker et al. [6], based on the classic Action cycle developed by Don Norman [59]. One thing to note from the sequence diagram is that the gesture interaction has not yet been considered in the interaction mapping of both the use cases, since only the touch and voice interaction is prioritized in this thesis and the interaction through gestures is not currently out of this thesis' scope. The sequence of operations of both the use cases are explained in the sections below.

5.2.1 Use Case 1 - Interaction Sequence

In this section, the user's are instructed to enter a destination on the AI-based HMI under the assumption that the vehicle is driving. The interaction begins to start when the user either touches the search button on the navigation scree or by invoking the voice assistant by saying the preferred destination. An audio prompt is played along with displaying the visual prompt, which directs the user to say their preferred destination. Based on the same survey mentioned in the chapter 4, the participants were asked whether they prefer to enter the destination by entering the street name first followed by the city name or vice versa, and as shown in the figure 5.3, ***28*** respondents preferred to enter the street name first followed by the city name, and the rest of the ***22*** respondents preferred the opposite way.

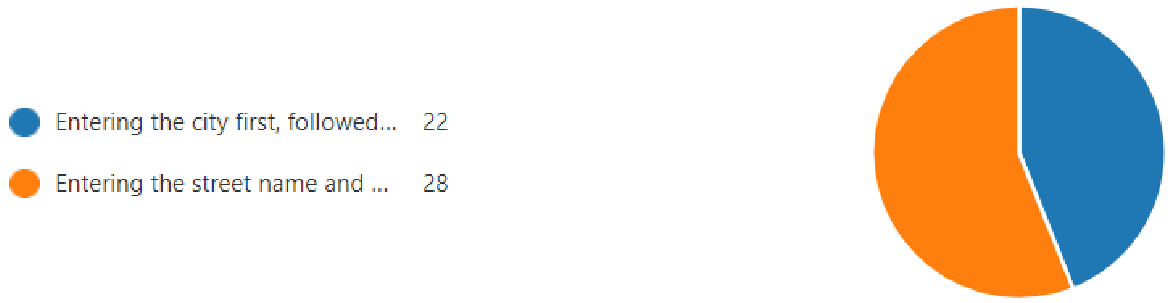


Figure 5.3: Pie chart representing the participant's preferred way of destination entry

Hence, for this use case, the users are then allowed to enter the destination by typing or saying the street name after the prior prompt given by the system. The system then calculates the route based on the user's input, and presents the entire route on the map. After a time elapse of 5 seconds, the system then provides the user via a display and an audio prompt an opportunity to select between three types of route modes varying between ***eco-route, route with shortest distance***, and ***route with faster time***. The user can then confirm the prompt either by tapping once on the screen or by saying the route name directly. In addition to this, the user can also ignore this prompt if they prefer to keep the stock settings. At the end of this interaction, a display prompt conveys the number of charging stations that has been added along the route based on the route mode selected, and along with the audio prompt provokes the user to either keep the added charging stations or to add/remove some of them. Once they users does their interaction, the system unlocks and enables the user to utilize the full functionality of the route guidance.

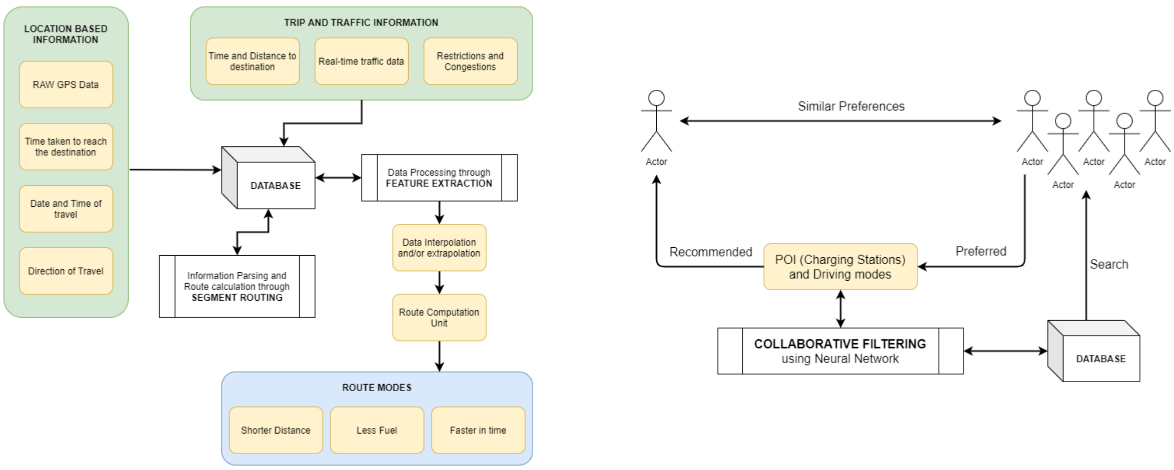
5.2.2 Use Case 2 - Interaction Sequence

This use case is a continuation of the previous destination entry use case, and the interaction begins when the user has added the necessary charging stations on the map. Here, the user is intimated via a voice prompt that the route guidance functions are unlocked and navigating to the destination said by the user. After a time elapse of 5 seconds, the system intimates the user in a low volume that the system and driving settings has been customized based on the route mode selected, and provides the user along with a display prompt that allows to user to either keep or customize these settings. If the user ignores/cancels the request, then the settings are not modified and are kept stock. However, there an option is provided so that the user can customize these setting at a later time if they want to.

On the other hand when the users present their need to modify the settings to the system (either by tapping the visual prompt on the screen or by confirming it through voice input), a display pop-up is displayed showing different driving modes available, along with the route modes as mentioned in the previous use case. Some of the driving modes displayed include ***comfort. sport, individual, off-road, eco-drive, normal*** modes, and these modes are designed in such a way that it gets automatically selected based on the type of route mode selected. However, the user are provided with the flexibility of changing these driving modes according to their needs, and can be done by the user either by tapping the desired driving mode on the screen or by saying the name of the desired route mode.

5.3 Reference Architecture for the AI-based HMI

One of the major problems in the automotive HMI development is due to design and provision of an HMI for every single individual system separately, and also questions arise concerning the integration of all the different function offered by the HMI into a whole functioning system considering the effect that it can have on interacting with the vehicle's occupants. In order to solve this, Amditis et al. [1] propose an overall integrated HMI architecture, where both the input and output are coordinated based on a central and shared in-vehicle controls. Based on this concept, the architecture of this AI-based HMI is constructed, and the architecture diagram for one the functionalities offered by the HMI is given in the figure 5.4 below.



(a) Destination Entry

(b) Point of Interest Recommendation

Figure 5.4: Architecture Diagram representing some of the HMI functionalities

From the above diagram, it can be seen that the system uses the ***Segment Routing***[1] to calculate the route based on the route and driving mode selected, and also adds the charging stations also the route using ***Collaborative Filtering***[2] method. Since the main goal of this thesis is to understand the interactions offered by the AI-based HMI, the more technical

[1] https://juni.pr/37pKuYl
[2] https://en.wikipedia.org/wiki/Collaborative_filtering

underpinnings of AI and it's functions are not considered and it out of scope. The main purpose of illustrating the architecture diagram is to get an idea about the data flow between the system, which in the end makes easy for the designers to map the interactions effectively.

5.4 Click-through Prototype of the AI-based HMI

The final step in the user interface development before conducting the user study is designing the user interface itself, and has to be designed based on the guidelines highlighted in the chapter 3, and by considering the user's opinions, needs and expectations as highlighted in the chapter 4, and has to ensure that the design was done in adherence to the requirements mentioned in the section 5.1 of this chapter. The prototype for the AI-based HMI realizing the use cases mentioned previously in the chapter 4, is realized as a click-through prototype, and this prototype was designed using ***InVision Studio***[3]. The prototypes that are developed using InVision are created as mock-ups for the different levels of functionalities offered by the AI-based HMI, and since it is not realized to its full extent, the user's will get a better impression of the interactions and features offered by the system. The home page of the prototype once the user's turns on the power is displayed in the figure 5.5 below.

Figure 5.5: Prototype Home Screen

[3]https://www.invisionapp.com/studio

Once the base template of the prototype as shown in the figure 5.5 is designed, it has to be incorporated for providing support to the user across the varying levels of interaction. As described by Lavie [44] and further elaborated by Walter [100], this prototype was designed to incorporate the 4 levels of user support, whose main objective is to examine the influence of various user characteristics and also the conditions/scenarios the tasks are performed. These 4 levels of user support are explained briefly in the sections below.

5.4.1 Level 1 - Manual Condition

The basic example of the manual condition is by enabling the user to enter their preferred input manually, where no prior support will be given to the user. As per the level 1, the figure 5.6 represents the traditional manual way of the user entering the destination and adding the point of interest on the map. No indications or prompts will be provided to the user at this level, and it is up to the the user to search and enter/add the destination and point of interest by themselves.

(a) Destination Entry (b) Point of Interest Addition

Figure 5.6: Entering/Adding the Destination and POI using the virtual keyboard

5.4.2 Level 2 - User Selection Condition

In this condition, the system presents the user with a set of options to select from, that has been curated automatically to the user needs based on the current scenario. From figure 5.7 it can be seen that the user is provided with the options of selecting their preferred route and/or driving mode when the route guidance is started. These options are either accessed automatically by the user with the help of recommendation provided by the system or can access by clicking the ***Driving Settings*** icon, which can be located on the left side of the screen immediately after the map area. The user is also guided automatically to this icon, in case if they missed the system's recommendation.

(a) Route Modes Selection

(b) Driving Modes Selection

Figure 5.7: Recommendation of preferred/alternative Route and Driving Modes

5.4.3 Level 3 - User Approval Condition

(a) Route Modes Modification Option

(b) POI Addition Option

Figure 5.8: Provision of recommendations to change/modify the settings and functions

This level is similar to that of level 2, except that the system provides the user with the functions/options that they tend to use more frequently and presents them with a choice of either to keep to current selected settings and functions or to modify them. Basically, the users are provided with a recommendation whether they accept or decline. As seen from the figure 5.8, the options of changing the route preference and the option of adding/removing the charging stations along the route is provided.

5.4.4 Level 4 - Highly Adaptive Condition

The highly adaptive condition increases the autonomy level of the vehicle to SAE level three and above, thereby providing the users with automatic responses based on the previously collected data. Instead of providing the user with a choice of selecting (*level 2*) or modifying (*level 3*), the settings and functions are automatically selected based on the current context and the system's knowledge base.

(a) Automatic Route Mapping (b) Error Correction Message (c) Alternate Responses

Figure 5.9: Provision of recommendations to the user automatically

Here, the figure 5.9(a) represents the automatic starting of the route guidance and also the addition of charging stations along the route, when the user just says the destination through voice input. For the figures 5.9(b), the case where the system provides an intimation to the user regarding incorrect recognition, and without letting the user to speak the desired address again, the system presents a list of potential choices as shown in the figure 5.9(c), where the user than then select the correct route.

5.5 Remarks

The main objective of designing the tablet-like user interface is to integrate every separate car functions into a single unit, thereby reducing the number of distractions and creating a holistic experience to the user, which are especially targeted towards the frequent users of smart devices. In order to keep up with the existing and evolving industry standard, the entire design, visual and voice aspects of the prototype were designed and developed by the application of taxonomy of guidelines by Normark et al.[61], along with the use of ***style guide*** from Microsoft[4].

[4] https://docs.microsoft.com/en-us/style-guide/welcome/

Chapter 6

Evaluation of Human-HMI Interactions and its Usability

Based on the guidelines drafted in the 3^{rd} chapter, and building up on the requirements elicited in the 5^{th} chapter, it is of absolute essential to design and develop the system which is good enough to satisfy the needs and expectations of the customers and stakeholders alike, and also allows them achieve their goals and objectives quickly and efficiently, thereby making the system ***usable***. Jakob Nielsen [57] call this type of UX design process as ***Usability***, which is a process that determines and examines how and why the user chooses and adopts the product, and the identifies the way that users seek to evaluate the product for achieving their goals and objectives. Based on the summary provided by Bevan et al. [11], the definition of Usability according the ***ISO 9241-11 Standard*** which is based on ISO 9241 - Ergonomics of Human Machine Interaction, is given below.

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

6.1 Evaluating Usability through Heuristic Evaluation

As mentioned in the chapter 2, the products to be designed based on the Usability has to be designed based on the 5 usability characteristics proposed by Whitney Quesenbery[1]. The process of designing the interface for its usability and user experience is an iterative process in nature, and some major usability problems will exist at the beginning phase of the product design. In order to identify the usability problems at the early stages of product development, ***Heuristic Evaluation*** method (refer section 2.7 of chapter 2), proposed by Nielsen and Mohlich [58], has to be carried out.

6.1.1 Experimental Design

The heuristic evaluation of the first iteration of the prototype was conducted virtually with each experts individually. For this session, a total of $N = 7$ experts from the field of automotive user experience and industrial design, gave their consent to participate in this session.

[1] https://www.interaction-design.org/literature/article/an-introduction-to-usability

The number of experts attended this study, is more than optimal number of evaluators as recommended by Nielsen[2]. Once the experts gave their informed consent, they are briefed about the Nielsen's 9 Heuristics (refer section 2.7 of chapter 2), and also about the UX assessment criteria that the expert has to perform after interacting with the prototype. For the UX assessment, the first step is where, the participants were asked to rate their interaction on a 7 point likert scale, which was developed based on the ***Dialogue Principles*** from the norm ***ISO 9241 - 110***. Finally, the expert concludes the UX assessment by completing a ***checklist***, which was derived based on the Nielsen's 10 Heuristics. Before the session begins, the every expert has been informed that their decisions will not be judged while they interact with the prototype, and they can execute the given tasks the way that seem fit to their mental model. Also the experts are requested to follow the ***Think Aloud protocol***[3]. The experts are also informed that the evaluation session will be carried out virtually, and they do not need to download/install any third party applications to take part in this session.

At the start of the session, the scenarios behind the development of the prototype along with the use case background were briefed to the experts, and they were instructed to assume the role of a car occupant who is interacting with the system, for the very first time. After the experts got familiarized with the prototype, a guided walkthrough of all the use cases and functionalities offered by the design was done, and subsequently some time were given to the experts to reflect on the tasks that they have been provided. Once the experts felt confident and comfortable about the tasks they have been provided, they were instructed to begin their prototype walkthrough, and during this they were encouraged to provide feedback, which were simultaneously recorded. Once the experts finish the UX assessment criteria for the prototype, the expert review session is then concluded.

6.1.2 Expert's Qualitative Feedback

Before proceeding with the assessment criteria, the qualitative feedback given by the participants during their interaction with the prototype were recorded, that yielded a very significant amount of feedback which are then later used to corroborate the ratings given by the participant. The display principles proposed by Wickens et al. [102], was used a measure of evaluating the expert's qualitative feedback. When looking into the qualitative responses by the experts, some of the most occurring and significant ones are mentioned here. Even though all of the experts agreed with the layout of the display and some design choices, all of the experts felt that the content and information provided through the display was ***overwhelming***, and they felt that there was a lot of ***visual clutter***. Since this comment was expected from the experts, their reasoning towards this statement was a bit unexpected, which was inherently due to the presence of some ***unmapped/unclickable icons*** in the layout. Another reason is that, the experts envisioned that the icons could be in a ***different space and size*** than what they saw, since they were not sure which menu/icon that they could focus on when performing the tasks.

The second unanimous response given by the experts while interacting the with the prototype, is to improve the ***quality of error messages*** and ***type of feedback*** received during interaction. Since the experts were not informed about the addition of voice feedback at the

[2] https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/
[3] https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/

later stage, they felt that it would be efficient to inform the occupants of the car that the AI is processing their request, and also make the system a bit more clear in the task that its going to perform based on the occupants' request. Most of the experts felt this could be improved because the given error messages and the explanations demanded a bit of time to be invested; furthermore, the nuance between some of the explanations are slightly small leading to confusion in perceiving the explanations, instead of being direct in nature. In the end, even though the experts felt the need to improve the quality of the explanations offered, all of them accepted the ***way it was conveyed to them***, which was through a pop-up recommendation message, except for one expert. In their case, they related to informations/explanations being conveyed in the form of a ***running text***, instead of being conveying the explanations through pop-ups every time.

6.1.3 UX Assessment

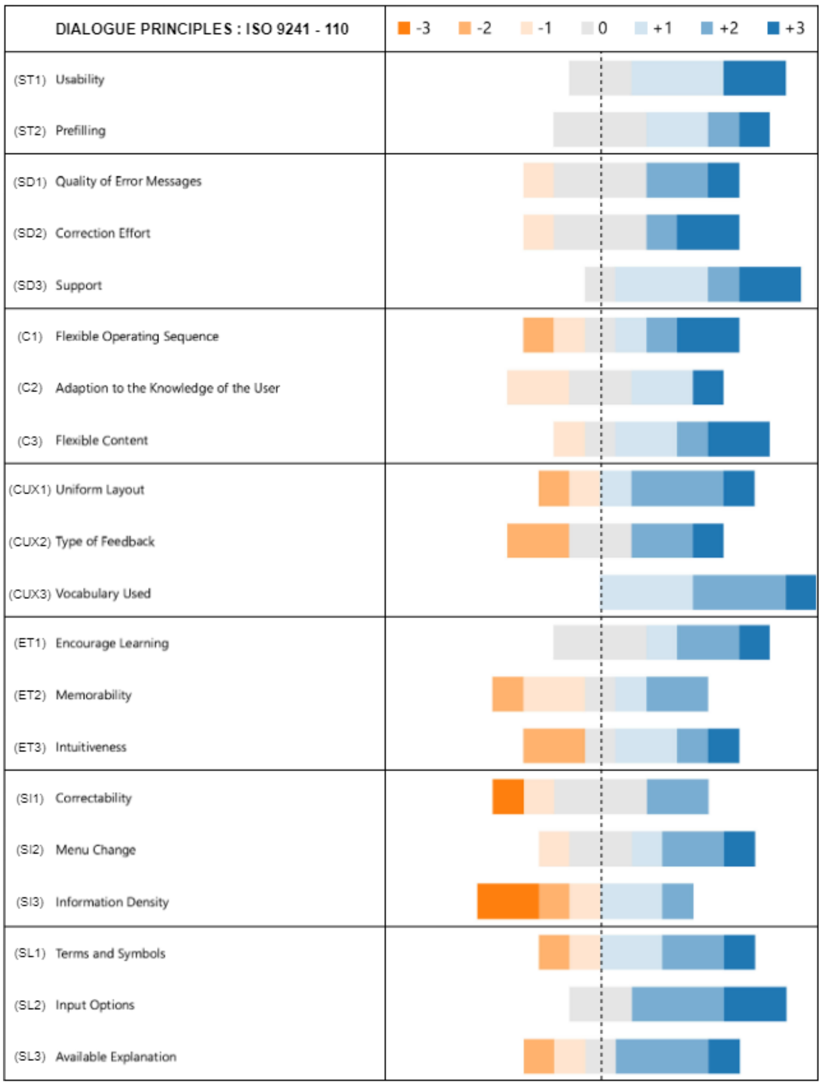


Figure 6.1: Expert's Responses for ISO 9241-110 Ratings

The next step in this expert review session, was for the experts to assess the interactions that they experienced with the prototype, by rating these interactions on a scale, followed by completing a custom checklist, which was drafted based on the Nielsen's Heuristics under consideration. The expert's responses for their experienced interactions are recorded based on the constructs from the standard ISO 9241-110, which is shown in the figure 6.1. Here, the experts are asked to rate their interaction with the prototype on a 7 point likert scale, and their responses are weighted with a rating of ***-3*** representing non-compliance to ***+3*** indicating a full-compliance to the norm.

The expert's responses from this standard was in-line with the qualitative feedback provided by the experts, which can be confirmed by observing the negative scores on some constructs like ***Information Density, Type of Feedback, Memorability***, and ***Correctability*** [11]. Setting these constructs aside, a fairly positive response ratings has been observed for other constructs, which in turn lays a strong emphasis and priority on addressing and adjusting these issues on the final design. The constructs from the standard ISO 9241-110, as seen in the figure 6.1, are elaborated in the table 6.1 below.

Table 6.1: ISO 9241-110 Constructs

ST	Suitability for the Task
SD	Self-Descriptiveness
C	Controllability
CUX	Conformity with User Experience
ET	Error Tolerance
SI	Suitability for Individualisation
SL	Suitability for Learning

Finally, the usability of the prototype is checked using Nielsen's Heuristics for Usability testing [58], and the experts were asked to rate the severity of the problems that they encountered on a severity scale. For measuring severity, Nielsen[4] suggests to use a 0 to 4 rating scale, based on which the severity can be assessed on the ***frequency, impact*** and ***persistence*** of the problem within the interface. The severity rating for the 9 selected heuristics is shown in the figure 6.2, and as expected, the heuristic ***Flexibility and Efficiency of Use*** recorded a higher severity rating. This in turn corroborated the experts qualitative feedback and also their ratings based on the standard mentioned above. In addition to this, the experts responses on the heuristic ***User Control and Freedom*** spanned across all the ratings, which suggests that some of the experts believed the interaction to be in sequence, thereby lacking in support of undo/redo. This response was obvious, since the prototype is in early stages of development, and this support will be considered in the final state of the prototype. Finally, some usability gaps that experts missed to address during their assessment, were then identified briefly by providing them with a ***checklist***, that was drafted based on the heuristics under considerations, and the expert's responses are listed in the appendix D.

[4] `https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/`

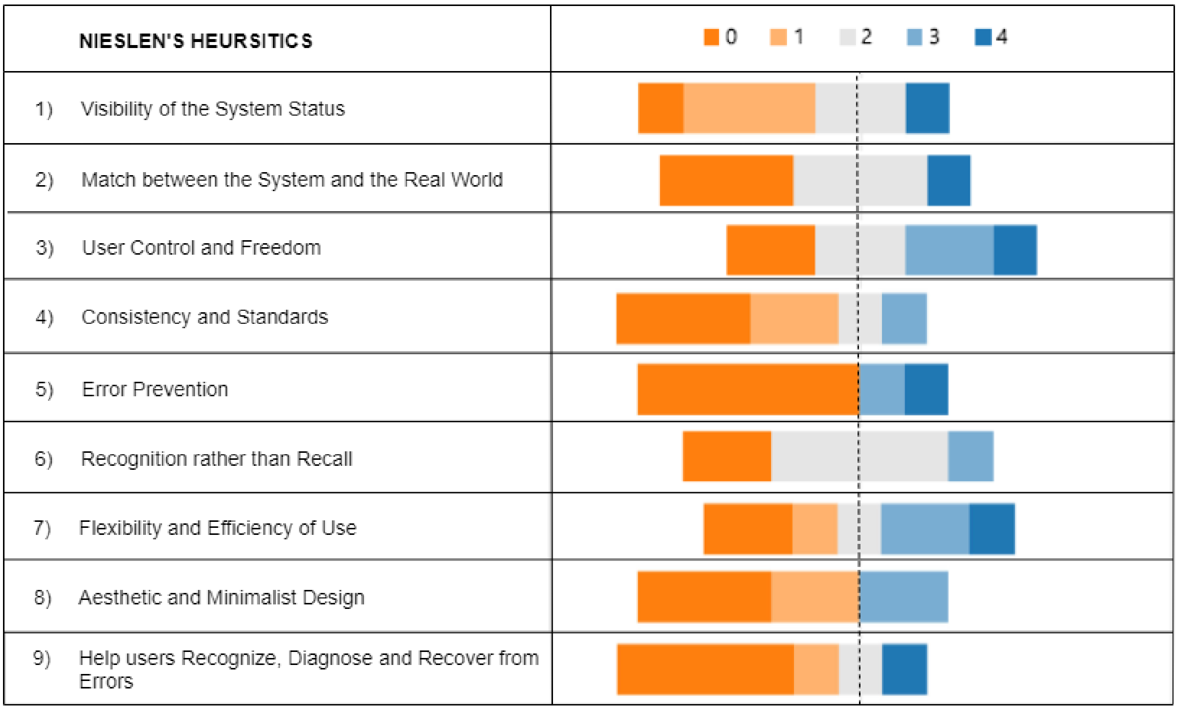


Figure 6.2: Severity Ratings for Nielsen's Heuristics

6.2 Final Prototype

Based on the qualitative feedback and the UX assessment of the prototype, the final prototype addressing the found usability has to be designed, especially giving a higher priority for ***reducing the visual inconsistencies and clutter***, and also for ***improving the quality of error messages and feedback provided***. Thus, the main goal is provide the user with a interface designed based on the drafted guidelines (refer chapter 3) and requirements (refer chapter 5), in addition to addressing the usability issues that was found during expert review. The issues that were addressed and the changes that were made during the final design iteration are given below. Some of the minor design changes are not mentioned here, but explained briefly in the coming sections.

1. ***Route Mode specific layout*** has been added as shown in the figure 6.3, where whenever a route mode is selected, the layout of the screen changes to adapt/add route specific view and its correspondent settings and features to the screen.

2. Some of the unused icons on the screen, such as the ones on the status bar, are completely ***removed*** from the final prototype (refer figure 6.3), and the other icons that are not in current use, are ***dimmed off*** to enable the user to select the more prominent icons (refer 6.4(b)). If necessary, the status bar can accessed in a way similar to notification tray found in mobile UI's.

3. Since, the route mode settings and the driving mode settings are dependent upon each other, the section of route and driving modes (in addition to being accessed trough the

settings icon on the status bar), are made to be accessed on a single window itself as shown in the figure 6.4(b), instead of displaying it on a different task bars (6.4(a)).

4. In addition to conveying information through a display pop-up every time, ***voice feedback function*** is also added for improving the quality of messages delivered. Due to limited scope, haptic feedback has been implemented in this prototype version.

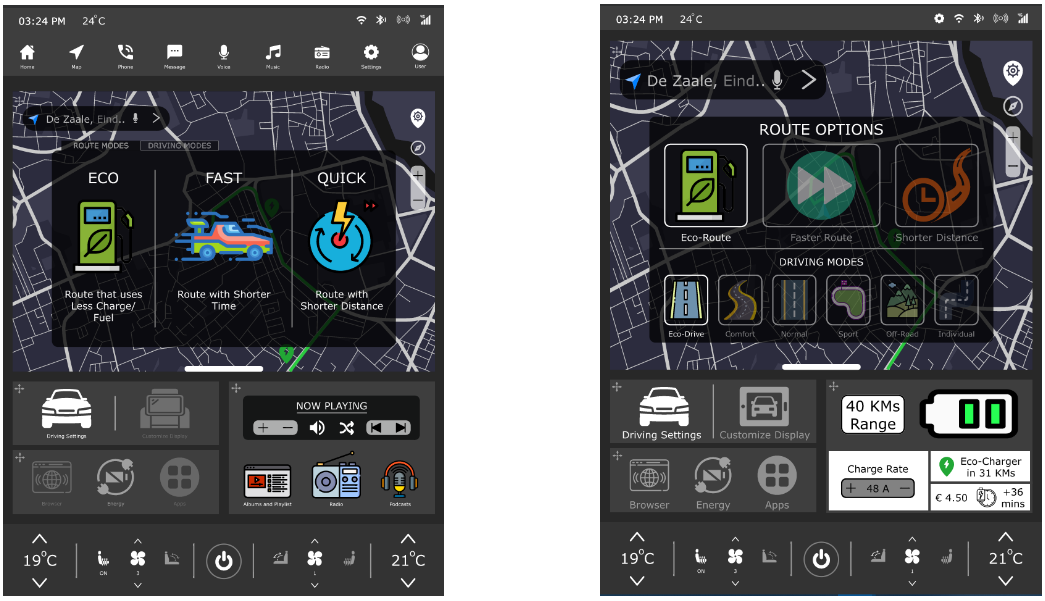
5. In addition to the automatic addition of charging stations based on the range, the functionality of ***automatic addition of some point of interests*** along the charging stations are also added, as shown in the figure 6.5. These point of interests are added based on the selected route mode, and the time taken for the vehicle to charge at the selected charging station.

(a) Eco-Route　(b) Route with Shortest Time　(c) Route with Shortest Distance

Figure 6.3: Layout change based on Route Settings

Figure 6.3 represents the different types of layout and settings change that occurs whenever a route mode is selected. As seen in the figure, the layout changes based on the types of route mode selected, for example, the energy settings has been made more prominent in the Eco-route mode as seen in the figure 6.3(a), other than the Fast mode (figure 6.3(b)) and Short mode (figure 6.3(c)) since it is obvious and necessary that the user pays more attention towards energy management. Keeping the layout change aside, the user's are also provided with an option of viewing the entire energy settings as whole, which can be accessed using the ***Energy*** icon, which are present on the services tray of the layout and can be easily accessible. In case for the layout based on shortest time (6.3(b)), the entire energy settings can be accessed by long pressing the energy tray.

Figure 6.4 highlights the differences in mode selection before (6.4(a)) and after (6.4(b)) the expert reviews. Since both the route modes and driving modes are dependent upon each other, it makes displaying both of these modes on a single screen more viable and usable. As seen in figure 6.4(b), ***Eco-Drive*** driving mode has been set as default since ***Eco-Route*** is the current route mode. For allowing more flexibility in mode selection, any driving mode option can be selected by the user irrespective of the default route mode selected, and the reverse of this case is also possible. In order for the user to quickly identify and change these settings, except for the selected route and driving modes, the rest of the buttons are dimmed off, which gives the user a visual cue regarding which buttons are and are not selected.



(a) Mode Selection - OLD (b) Mode Selection - NEW

Figure 6.4: Mode Selection in Previous and Final Iteration

(a) POI Addition 1 (b) POI Addition 2

Figure 6.5: Point of Interest Addition along the charging stations

As explained earlier, the figure 6.5 represents multiple point of interests that are added along the route. One thing to note from the above figure is, the amount of POIs added will be larger for eco-charging stations, and will be smaller in number along the fast charging stations. By doing this, the user's will significantly understand the decisions taken by the system, since it was in-line with their mental model (one instance being less POIs around fast charging stations since the charging time is low). Furthermore, ***no voice feedback*** will be provided to the user intimating the addition of POIs when the selected route mode is based on the route with *shortest time*, instead of the Eco-route mode where a voice feedback will be provided. In this case, the user can automatically add these POIs by just tapping the map icon on the screen, once a ***red bubble***, similar to notification bubble found in ***Android 10***[5], appears on the map, One thing to note that this feature is possible only in the case of fast charging stations. This red bubble notification can be found in the figure 6.4(b).

6.3 Remarks

Even though Heuristic Evaluation is considered as a type of ***discounted usability engineering method***[6], it is the most widely successful methods for identifying the major usability problems, before proceeding towards the final user study. One thing to note that, performing heuristic evaluation does not guarantee perfect results every time, but it aims to identify the underlying usability problem, based on which redesigning can be done. Hence, by performing these expert reviews, the usability problems associated with the prototype under consideration are identified, which ultimately leads to the development of a prototype that addresses its usability issues.

[5] https://zd.net/34v5n28
[6] https://www.nngroup.com/articles/guerrilla-hci/

Chapter 7

Evaluation of User Experience and Acceptance

The final step in exploring the impact of AI-based HMI on Human-HMI interaction is the process of ***usability testing***, where the final prototype designed by addressing some its major usability issues (refer section 6.2), is tested with user's in general to identify how the implementation of AI-based HMI affects the user experience and acceptance. Studying the user acceptance regarding the usage of AI-based HMIs is a key factor for this research based on which, the potential benefits and the corresponding caveats of implementing AI in Automotive HMI can be found and communicated.

7.1 Goals and Objectives of the Study

The goal of this usability testing is twofold. The **first goal**, is to compare AI-based and non AI-based HMI on the basis of a given task (in this case, a destination entry task using the navigational system of the prototype), for determining the cognitive workload taken by the user to perform these tasks. The **next goal** is to determine and analyze the mental model that the user develops while interacting with the system [41], and how accurate the developed mental model is, by analyzing with the elicited requirements, which can be found in the section 5.1. Thus, the main focus of the study is to evaluate the user experience and acceptance of the AI-based HMI, and this shall be achieved by satisfying the following research questions, which can also been in chapter 1.

Table 7.1: Research Questions

Code	Questions
RQ1	How does an AI-based HMI of an autonomous vehicle which allows the occupants to interact and customize its interface result in an improved user experience and acceptance of the system?
RQ2	How well do the occupants of an autonomous vehicle understand and anticipate the behavior of the system based on the context that was presented to them?

7.1.1 Study Description

The experiment is conducted in the form of ***Remote Moderated Usability Testing***[1], where the experimental session will be carried out virtually via a video sharing and conferencing platform, such as Microsoft Teams (preferred) or through Zoom. In this session, the link for the HMI prototype will be shared to the participants as a click-through prototype, where they can interact with it. As mentioned in the chapter 5, the prototype is designed using a tool called ***InVision Studio***[2], and the participants do not need any account to access the prototype, as the prototype will be hosted from InVision studio's servers. Since this is a virtual study, there is no need for observing the people's physical behavior. There is no deceit involved in this study and only the way the participants interact with the prototype will be observed. This will be done, only after acquiring a proper signed consent from the participants before starting the study. Thus for this study, participants who are aged above 18, fluent in Dutch, German and/or English, and who are/expected to use means of transportation will be considered. Participants under the age of 18, and who pose physical/mental impairments limiting their participation, will not be considered as a viable research population for this study; however, these group might be considered for future research.

7.1.2 Study Assumptions

For evaluating the user experience and acceptance, some assumptions has been made before the start of the study, especially related to functionalities that AI has to offer and provides to the occupants of the car. These assumptions were made in such a way that the final results will not compromise the elicited guidelines and requirements. Some of the assumptions (related to prototype design and experiment) is given below.

1. The car under consideration is assumed to be fully electric and consists of level 4 or higher autonomous vehicle, which is based on the Society of Automotive Engineers' (SAE) 5 levels of autonomous vehicles

2. The user has given permissions to the system to collect and store location based data and has authorized the system to access the user's calendar

3. The system has already collected and stored a substantial amount of above-mentioned data, so that it can provide and aid the user during navigation and customizing the interface

4. Due to the assumption being the vehicle is a higher level autonomous vehicle, it also assumed that both the driver and the occupants of the vehicle will have the exact same interaction with the vehicle's HMI

5. It is also assumed that in this instance, no gesture input will be present within the car, and only touch and voice input modality is taken under consideration

6. As this study is to be conducted entirely online, it is assumed that the destination entry task is performed when the car is at idle, and the customization task is done when the vehicle is in motion

[1] `https://www.nngroup.com/articles/moderated-remote-usability-test-why/`
[2] `https://www.invisionapp.com/studio`

7.2 User Research Methods

As per the main goal of this study, evaluating the user acceptance and experience of the AI-based HMI is made possible by providing the study participants to perform some simple tasks on both AI and non-AI based prototype, to determine their amount of workload during interaction, and also has to answer some questions based on the levels of customization offered, to comprehend and analyze their mental model of the system. Thus, the tasks to be performed by the participants in this study is given below.

Table 7.2: Tasks to be performed by the participants

Code	**Task Description**
T1	Destination and Point of Interest entry through the touch input using the keyboard provided, and starting the route guidance
T2	Destination entry using the AI-enabled speech input, and thereby starting the route guidance
T3	Develop a mental model of the system by watching a video inferring the levels of customization offered by the system, along with interacting with the prototype showcased in the video

Before delving deep into the tasks itself, it is necessary to identify the sample size of the participants to be invited for the study. In order to get a more statistically significant rating, Nielsen recommends a minimum of 5 participants to be significantly enough, if the user testing is iterative in nature[3]. Thus, a sample size of minimum ***12*** potential participants was fixed initially for this remote moderated usability study. Since this is qualitative study, the participants are planned to be recruited through ***convenience sampling***[4]. The advantage with convenience sampling is that it enables us to collect the data more quickly, it is economical and enables us to choose from a readily available sample with a minimum bias.

The participants will be invited to take part in this experimental study by sending a link for a video conference using Microsoft Teams prior to the start of the experiment. The cumulative duration of the study (45 $\sim$ 60 minutes) will also be notified to the participants in the study invitation. The participants will be invited either through their E-mail or by using social media platforms such as Facebook or LinkedIn.

Once the participants are recruited, the usability study will be performed in ***2 iterations***, where the 1^{st} iteration will be considered as a pre-study, and will be performed initially with 4 participants. Conducting the pre-study is essential at the early stages of user testing where doing this will lead to some problems and mistakes with the experiment design, which can rectified easily. If no changes are to be made to the design and the experiment itself based on the results from the pre-study, then the study data collected will be carried over to the final results once the 2^{nd} iteration ends. The experiment for the main study is similar to what performed in the pre-study, and the steps in these study are explained below. One thing to

[3] https://www.nngroup.com/articles/how-many-test-users/
[4] https://en.wikipedia.org/wiki/Convenience_sampling

note that, if some changes are observed in the pre-study, then the results obtained from the pre-study ***will not be carried over to final data tally***, and the main user study will be carried out with 12 participants, instead of 8.

7.2.1 Comparison of AI and non-AI based HMI

In lieu with the first goal, the all of the participants were asked to interact with both the AI and non-AI based HMI prototype by performing the tasks T1 and T2, where the workload of the participants when interacting with the prototype is measured. Since all of the participants are going to interact with all aspects of the AI and non-AI based prototype, the participants are split equally into control and experimental groups, and are requested to perform these tasks on both the prototypes. While evaluating with the prototype the issue of ***order*** and ***practice effects***[5] are bound to occur, and in order to balance these errors, the control group is made to interact with the non-AI based prototype first followed by AI-based one, and the reverse is done with participants placed in the experimental groups. The most common and widely used scale for measuring the participants perceived subjective workload is the ***NASA TLX (Task Load Index)*** scale, which as the name suggests, was developed by NASA [32] in the 1980s.

7.2.2 Analysis of User's Mental Model

As mentioned in the section 3.3 of 3^{rd} chapter, determining how the user's perceive and anticipate the behavior of the system is important to evaluate the traceability of the system, which in turn is possible by analyzing the ***mental model*** developed by the user [13, 41]. The final prototype is designed in such a way that it can enable the user to develop the mental model of the system easily, and if not, then conveys informations and/or explanations in such a way it helps the user to build and develop sufficient mental model to understand the system. Considering the taxonomy of user's frequent doubts while interacting with the system [86], the questions that the user's might ask while interacting with this system, and the expected answers that the participant might give after interacting with the system is provided in the table below.

Table 7.3: Mental Model Questions and Expected Answers

ID	Questions	Expected Answers
Q1	Why was the system customized in that specific way?	The system customized the layout based on learning the recurrent behavior of the user
Q2	Based on what context, the customization was done?	Based on the route mode and driving mode selected by the user. Time of the commute and the date also can be taken into consideration

[5] `https://bit.ly/3aG8gRJ`

Q3	Based on the customization tasks showcased in the video, how certain are you that the AI adheres to customization context?	75 – 100% Expected 50 -75% Partial Match $\leq$ 50% No Match
Q4	Are there any other possibilities of performing this customization in a different way?	Changing the layout, Providing personalized shortcuts and recommendations
Q5	What if the system does not offer this kind of customization and why?	System will not be usable / AI has not been started or invoked yet

In addition to the above mentioned questions, the participants were also asked to describe a situation in general, where these type of customization showed will be useful. For analyzing this, a ***Wizard of Oz*** style experiment will be carried out, where a video showing a wizard operating the prototype will be shown to the participants, and once the video is finished playing, they are provided with the prototype again and made them to experience the same levels of customization's as shown in the video. At the end, the participants are asked the questions mentioned in the table 7.3, and their responses are checked and analyzed.

7.2.3 Usability Measures and Questionnaires

For the evaluation of user experience and acceptance, a couple of standardized questionnaires has been employed in this study. A brief description of the questionnaires are given below.

1. As explained in the subsection 7.2.1, ***NASA TLX Scale*** [32] has been used to measure the workload of the participants performing the tasks on the prototype. This scale is used to assess the perceived workload of the participants on a scale of 0 (very low) to 100 (very high).

2. The user acceptance of the model is analyzed using the scale developed by Osswald et al. [63], called the ***Car Technological Acceptance Model*** (CTAM), which is derived from the standardized acceptance models such as TAM, TAM2, and UTAUT, these of which are explained in the section 1.6 of chapter 1. Using this scale, the behaviour of the participants while interacting the in-vehicle infotainment system of the car can be predicted, and also is used to support the design process of the infotainment system.

3. Finally, the pragmatic and hedonic quality of the HMI design is analyzed using another standardized scale, such as ***Attrakdiff***[6], developed by Marc Hassenzahl et al.

In addition to the above 3 scales, some questions about addressing the user's mental model (mentioned in the table 7.3) is asked qualitatively, and the users responses will be coded as a ***match, partial-match*** or ***no-match*** to the expected answers of the mental model.

[6] `http://attrakdiff.de/index-en.html`

7.3 Experimental Design

The steps in this usability study is split up over 7 phases, and the procedure involved in each phases of the study in described below. One thing to note that, the first half of the study (Phase 3 and Phase 4) is a task based study, being that the participant uses the prototype, and the second half is video based study (Phase 5).

Phase 0 - Participant Invitation

The participants are invited to take part in this experimental study by sending a link for a video conference using Microsoft Teams prior to the start of the experiment. The cumulative duration of the study (45 minutes $\sim$ 60 minutes) will also be notified to the participants in the study invitation. The participants will be invited either through mail or by using social media platforms.

Phase 1 - Formal Introduction and Consent

The goal of the study and the subsequent use cases are introduced to the participants at the beginning of the video conference. In addition to this, the participant needs to sign a declaration of consent to participate in this study, followed by filling out a short demographic questionnaire.

Phase 2 - Experiment Briefing

In the second phase, the participants will be notified regarding the experimental steps, and tasks they are required to perform. The participants are also notified again that they do not need to install any specific apps or opening an account to work around the system. Once briefing is done, along with getting the consent of the participants again, the moderator will start the recording of this session.

Phase 3 - Manual Interaction Task

After briefing session, the link of the prototype will be shared to the participants and they are instructed to perform a manual destination entry task and point of interest addition task on the prototype by inputting the given address and following the necessary instructions. Once these tasks are done, the participants are supplied with the NASA TLX questionnaire.

Phase 4 - AI based Interaction Task

This phase will be split up into 2 parts. Before beginning this phase, another link for the prototype will be shared with participants, and the participants are instructed to say the address that was mentioned in the task sheet. This action by participants has to be completed twice in this phase, and the 2 interaction instances is briefed below.

1. Entering the destination and enabling the route guidance by entering the destination through voice input
2. Choosing/Not choosing to add the interim destination based on a message notification received after enabling the route guidance

Once the participants are done with these tasks, they are again asked to complete the questionnaire, which is the NASA TLX scale.

Phase 5 - Assessing the User's Mental Model for AI-based Customization

This phase is different than other phases. Here, a video will be shown to the participants of the Wizard interacting with the prototype, and performing customization tasks. The participants have to look at it and after this, they are encouraged to answer some questions mentioned in the table 7.3, which are subsequently used to determine the user's mental model.

Phase 6 - Final Questionnaire

After all the tasks are done, the participants are requested to complete 2 questionnaires which are brief in the subsection 7.2.3], which are used for recording the user's acceptance scores and also scores corresponding to user experience of the system

Phase 7 - Study Conclusion

Once the questionnaire are finished, the participants are sometimes probed for more explanations based on their particular set of responses. After this, the participants are told that this is the end of the user study, and the session will be concluded.

7.4 Study Hypotheses and Expected Outcomes

In lieu with the research questions stated in the table 7.1, the following research hypothesis along with their expected outcomes for this experimental study are outlined in the table 7.4 given below.

Table 7.4: Research Hypotheses and Expected Outcomes

Hypothesis 1	
H1	Users rate the performance and acceptance of AI-based HMI significantly higher than the performance and usability of the non AI-based HMI
NH1	The performance and acceptance rating of the AI-based HMI provided by the users are equal or significantly lower to the ratings recorded for traditional non AI-based HMI
Expected Outcomes	
• The rating from the NASA TLX scale shows the perceived workload of the AI-based destination entry task is significantly lower than the perceived workload of the manual destination entry task • The determinants from the CTAM scale are rated positive for the interactions offered by the AI-based HMI, and also has secured significantly higher ratings when compared to the traditional non AI-based HMI	

Hypothesis 2	
H2	Users develop a mental model based on the customization context provided to them, and significantly determine what the system does at that moment and what corresponding actions to be taken
NH2	The mental model developed by the user does not help them to determine and understand the context of customization provided by AI-based HMI
Expected Outcomes • The responses given by the users matches with the potential expected answers to the questions to be asked, for determining the user's mental model • The users rated the AI-based HMI in the context of customization, with a significantly high acceptance scores	
Hypothesis 3	
H3	The interactive AI based HMI prototype presents a significant increase in usability and user experience
NH3	There was no significant increase found in the usability and user experience measures of AI based HMI prototype
Expected Outcomes • Responses provided by the user indicate a significantly positive scores for usability and user experience • The prototype has secured a significantly higher scores measuring its pragmatic and hedonic qualities	

7.5 Pre-Study Results and Observations

For this pre-study, a sample size of $N = 4$ participants were recruited, as per the recruitment criteria mentioned in the section 7.2, and these participants were chosen only from Europe. One thing to note that, the main purpose of conducting this study is to identify and rectify some changes in the experiment and design itself, and also to identify the initial trend of the results gathered. The statistical analysis of these results will be done later based on the observed results. As mentioned in the section 7.2.1, before starting this study, the recruited participants were allotted a randomly generated anonymized number and were placed in their subsequent control/experimental groups.

Once the participants finished filling the demographics questionnaire, they were asked to perform some tasks on the prototypes presented to them in blocks 1 and 2 (phase 3 and 4), were upon completion has to rate their amount of perceived workload on evaluating the prototype using the NASA TLX scale. For this instance, only the rating scale from the TLX was used, and the subsequent word pairs were left out [31]]. The total average workload calculated for both the blocks of prototype for the control and experimental group is shown in the table 7.5 below. The individual workload of each NASA TLX constructs is provided in the appendix E for further reference.

Table 7.5: Pre-Study : Workload Average Values

	Block 1	**Block 2**
Control Group	42	28
Experimental Group	28	38

From the above table it can be seen that, the participants from the control group rated interacting with the manual prototype with a higher average workload value of ***42*** than the average workload value of ***28*** for the AI-based prototype. This seems to be in-line with the probable expected outcomes of hypothesis H1, described in the table 7.4. When looking into the average workload values rated by the experimental group, AI-based prototype had a higher average workload value of ***38*** when compared with the average workload value of ***28*** for the manual prototype. For this instance, there might be a case of the participants ***uncertainty/unfamiliarity*** with the AI based prototype. There might also be another instance where they might have rated the manual prototype with a lower value due to them getting accustomed to using the former one (manual prototype) on a day-to-day basis. It is difficult now to corroborate this trend with the hypothesis and still more data need to be collected and analyzed to confirm this trend.

Similar to H1, for corroborating the hypotheses H2, the user's mental model while interacting with the AI-based prototype are then analyzed based on the method described in the ***phase 5*** of the experimental design, and the number of participant's responses that either fully match (FM) or partially match (PM) or in no match (NM) with the expected user's behaviors and their understanding of the system (refer expected answers of the table 7.3) are given in the table 7.6 below.

Table 7.6: Matching the pre-study's participants responses with the expected answers

ID	**Questions**	**FM**	**PM**	**NM**
Q1	Why was the system customized in that specific way?	3	1	-
Q2	Based on what context, the customization was done?	3	1	-
Q3	Based on the customization tasks showcased in the video, how certain are you that the AI adheres to customization context?	2	1	1

Q4	Are there any other possibilities of performing this customization in a different way?	3	1	-
Q5	What if the system does not offer this kind of customization and why?	2	1	1

By analyzing the mental model of the pre-study participants using the table 7.3 and 7.6, it can be seen that the responses provided by 3 of the 4 participants was in-line with the mental model requirement for the prototype. Only 1 participant's mental model does not meet with that of requirements of the prototype. Their major statement was this system does not resonated with them and they prefer the manual model over the AI-based HMI. This one seems to corroborate with the average workload scores for the experimental group as seen in the table 7.5, since the workload score was higher for the AI-based prototype than the manual one.

Even though most of the yielded results seems to corroborate with the requirements of the mental model, the responses provided by one of the participants, seems to be in contradiction with the expected answers. The necessity to conclude the mental model only with this data seems insignificant right now, as there may be many other participants who has the same thought of that particular participant, or this result might be an outlier. So, more data needs to be collected to check this fact. As mentioned earlier, participants were also asked to describe a scenario in which they expect the system to customize the settings and features for them. Some of the responses provided by the 4 participants that significant to analyze the mental model of the user are given below.

1. Music recommendations, different layout based on profiles, to use more social media applications

2. Route settings other than the city scenario which for example in mountains, plains and many more

3. Removing the route settings that are not suitable for the route; for example, there is no purpose of keeping off-road setting when driving in the city

4. General recommendations and personalized services

5. Route selection based on the occupants mood and also the type of travel

Once all the blocks are finished, the participants are then asked to fill out the questionnaire based on the CTAM scale [63], and the resulting average scores graph are provided in the appendix E for reference. Looking into the average scores, most of the acceptance scores rated by the participants are mostly positive based on the CTAM constructs. Negative values are expected for ***Anxiety***. Some of the constructs that are having a more neutral average score, are the constructs that are insignificant and out of scope of this study. This might be due to the participants being not to know what to rate for those constructs, and also being confused due to its irrelevancy to this context. For this reason, they were instructed to provided a neutral score or to provide a score by thinking of the entire use case in general.

Finally, the pragmatic and hedonic qualities of the AI-based prototype is analyzed using the Attrakdiff survey. From the semantic differential graph of the word pairs (refer appendix E) and its corresponding average values, it can be seen that most of the average rating given by the participants are on the positive side, which is in-line with the expected outcomes. However, all of the participants believed that interacting with the AI prototype were a bit ***too technical*** for them, even for the 2 participants, who perceived a less average workload by interacting with the AI-based prototype (refer table 7.5). Also based on the scores, the hedonic quality of the prototype is found to be higher than its pragmatic quality, which corroborates the fact that the participants rated the prototype to be more technical in the word pairs. For now, the rating for the prototype tend to incline between ***self-oriented and neutral***. This represents that even though the prototype has a higher hedonic quality score, it is still sell-oriented in use. This can be seen in the figure 7.1 below.

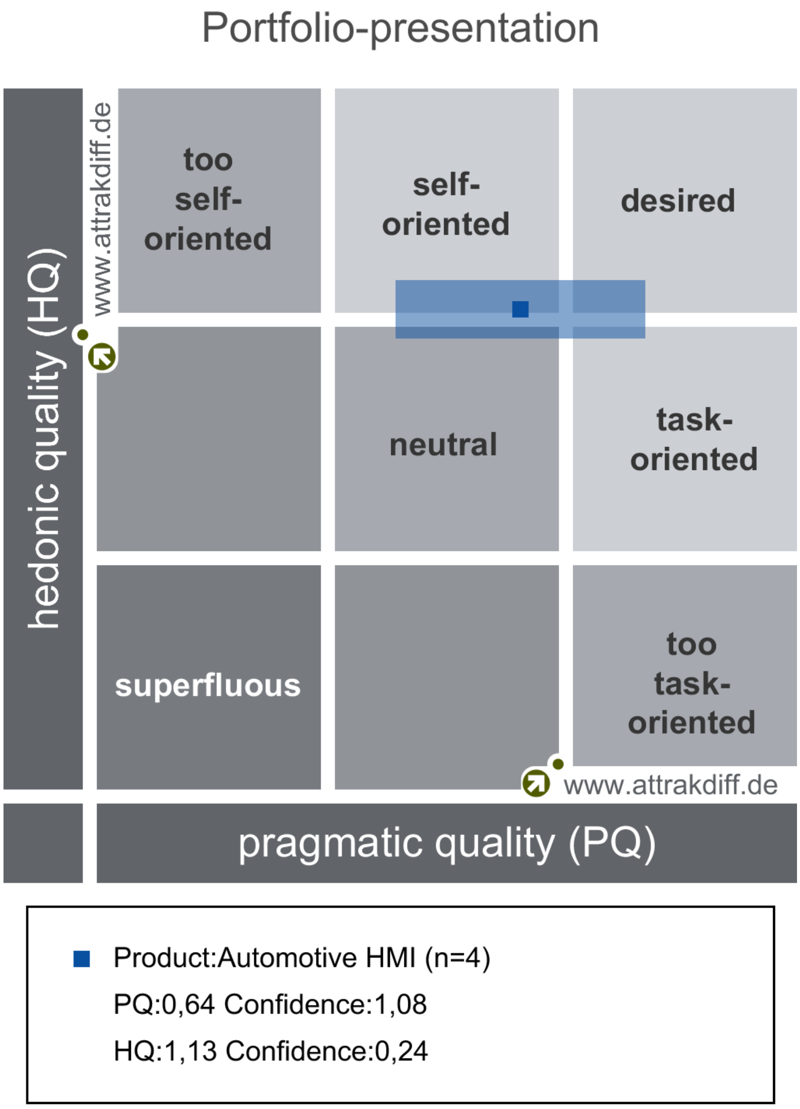


Figure 7.1: Portfolio of the results for the Pre-Study

Hence, based on the observations of the responses recorded during the pre-study, the data indicates that even though some trends are to be seen, still more data are needed to be collected in order to confirm these findings with the research hypotheses. Even though the AI-based interface seems to be less reliable right now due to the big confidence rectangles as seen in the figure 7.1, the rating of the word pairs seems to be in-line with the expected outcomes of the hypothesis H3. Thus, looking into the observations from the pre-study, there arises no need for any drastic changes within the experiment design and the design itself, and it is was deemed safe to add these data to the final data tally of the collected responses from the main usability study.

7.6 Results of the Usability Study

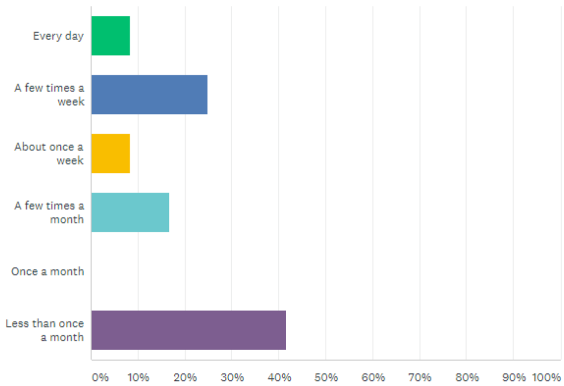
In spite of some surprising trends revealed by the pre-study, the final usability study was carried out as an extension of the pre-study, since no changes in the data were found in the results. Similar to the pre-study, the entire session of the usability was carried out virtually using a video conferencing platform. The results obtained from this usability study is presented in the coming sections. Due to the quantitative nature of the study, the results obtained from the questionnaires were analyzed using statistical descriptives which was extracted from IBM SPSS Statistics[7], and the corresponding verbal feedback of the participants were also recorded, which were used further to affirm and corroborate the rated scores in general. For statistical calculations, the SPSS survival manual from Julie Pallant [64] and the tutorials provided by Laerd Statistics[8] were used as reference.

7.6.1 Participants

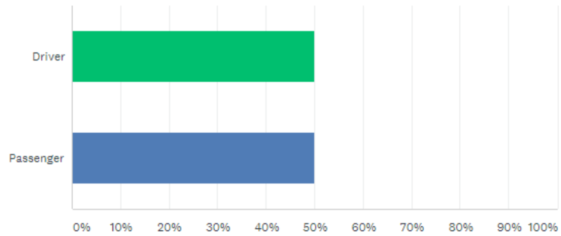
Similar to target group recruited during the survey mentioned in the chapter 4, the target group recruited for this study was within the age range of 18 to 40. So, a total of ***12*** participants, including 4 participants from the pre-study (N = 12 ; 9 Male ; 3 Female), aged between 18 to 34, were recruited for this usability study through ***convenience sampling*** method. From the recruited population, 10 participants were from Europe (5 each from Netherlands and Germany), and the other 2 participants were from India. All of the participants had a significant technical background relating to self-driving cars in general, and also significantly interact with the in-vehicle infotainment systems. When asked about driving in-general, a wide distribution of responses are recorded, and an equal split can be seen in their driving preference, as 6 participants each said they only like to travel as drivers and the rest of them as a passenger. The responses of the participants for the above details are illustrated in the figure 7.2 below.

Finally, the participants are allotted with a randomly generated anonymized number, and are placed in their respective control and experimental groups. Stating once again that, the participants from the pre-study are considered for this usability study, and no shifting/shuffling of these participants were done within the groups. The group that were placed in the pre-study remained the same. One thing to mention that the participants were allowed to participate in this usability study, only if they gave their consent to participate.

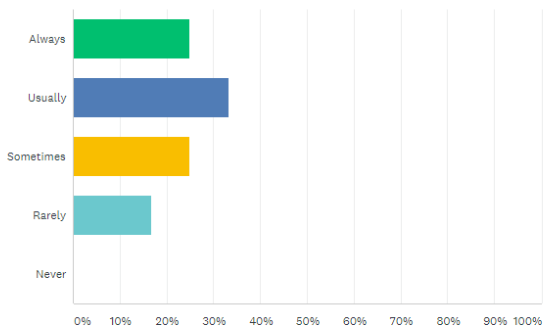
[7] https://www.ibm.com/products/spss-statistics
[8] https://statistics.laerd.com/



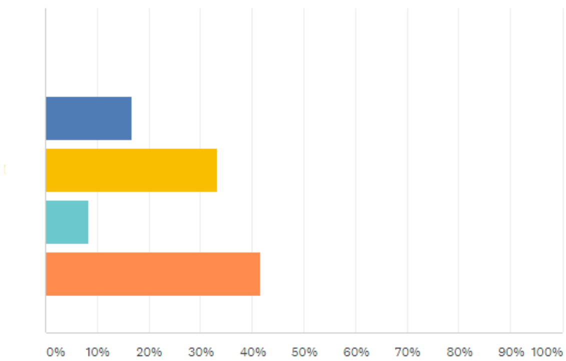
(a) Amount of Time Driven



(b) Driving/Riding Preference



(c) HMI Interaction Frequency



(d) Technical Knowledge w.r.t Self-Driving Cars

Figure 7.2: Participant's responses for driving-related demographic questions

Once the participants finished the demographics questionnaire, they are provided with the web-link, clicking on which they can get full access to perform the interactions on the clickable prototype. Before proceeding with the tasks, the participants are reminded once again that their decision will not be judged and are free to express their thoughts about the HMI, and also they were encouraged to think out loud while performing the given tasks. The tasks and their subsequent results are presented in the coming sections below. The total amount of time taken by the participants to finish this usability study was roughly between 45 to 60 minutes.

7.6.2 Comparison of Perceived Workload during Interaction

The main objective of this part of this usability study is to compare and estimate the total amount of workload that the user's perceive while performing the first use case, which is the destination entry task as mentioned in the section 4.2.4. This block constitutes to the phase 3 and 4 of the experimental design, and the participants rated their perceived workload on a NASA TLX scale. The workload ratings provided by the participants for both the manual prototype and AI-based prototype, are then compared and analyzed for their significance. For this, the an ***independent-samples t-test*** were carried out for the main objective of identifying whether there is a significant difference in the average workload scores between both the control and experimental group [64]. The results from the independent-samples t-test are provided in the table 7.7 below.

Table 7.7: Independent-samples t-test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
B1_AVG	EVA	.131	.725	-.309	10	.764	-3.83167	12.39921	-31.45882	23.79549
	EVNA			-.309	9.991	.764	-3.83167	12.39921	-31.46217	23.79883
B2-AVG	EVA	.939	.355	-.216	10	.834	-2.49833	11.59017	-28.32284	23.32618
	EVNA			-.216	7.256	.835	-2.49833	11.59017	-29.71011	24.71344

It can be seen that, SPSS provides results for 2 conditions, one being equal variances assumed (EVA) and the other being equal variance not assumed (EVNA). So in order to check the assumptions of the t-test, the significance value yielded from ***Levene's test for equality of variances*** are taken into account. Since the significance value of both B1_AVG (Sig. = 0.725) and B2_AVG (Sig. = 0.355) are way ahead of the standard value of 0.05 [64], the t-test scores are taken with respect to EVA (Equal Variance Assumed), and the difference between the two groups can be assessed. To achieve this, the ***Sig. (2-tailed)*** column from the ***t-test for Equality of Means*** are taken into consideration. By looking into this, it can be observed that, no significant difference in the scores between the control (B1_AVG (M = 31.3350 ; SD = 21.15274) ; B2_AVG (M = 32.8350; SD = 21.79458)) and experimental (B1_AVG (M = 35.1667; SD = 21.79458) ; B2_AVG (M = 35.3333; SD = 25.51148)) groups were found both for block 1 (t(6) = -.309 ; p = .764 ; two-tailed) and block 2 (t(6) = -.216 ; p = .834 ; two-tailed). Both the blocks yielded very high p-value, which is way above the designated standard value of 0.05.

Calculation of Effect Size Statistics

In addition to identifying whether the difference could have occurred by chance, the magnitude of the differences between the two groups should also be identified, which can done by considering the ***effect size statistics***. According to Julie Pallant [64], the most commonly methods for this is by using ***eta squared*** method, which represents the proportion of variance in the dependent variable that is explained by the grouping variable. The other method is by calculating ***Cohen's d***, which presents the difference between the two groups a measure of standard deviation. Since variances are under consideration, ***eta squared method*** is used for calculating the effect size statistics, which can be done using the formula given below.

$$\eta^2 = \frac{t^2}{(t^2 + (N1 + N2 - 2))} \tag{7.1}$$

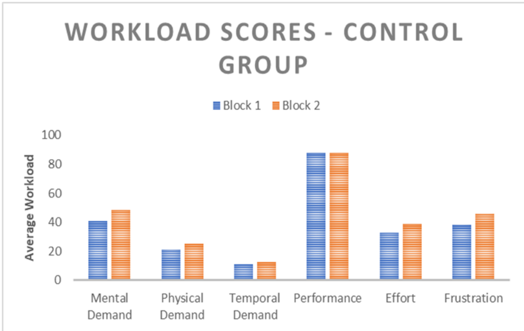
By substituting the t scores, and mean values of both the groups, which can be seen in the table 7.8, the value of eta squared can be calculated. Before analyzing the results, it is essential to identify the guidelines that these values are to be interpreted upon. This is made possible by J Cohen [14], where he posited that the effect size is small for values equalling 0.1, moderate for values in the range of 0.6, and larger for values equal to 0.14. By calculating the effect size of these samples, it was found that the magnitude of the differences in means (B1_AVG (mean difference = -3.83167 ; 95% Confidence Interval = -31.45882 to 23.79549) ; B2_AVG (mean difference = -2.49833 ; 95% Confidence Interval = -28.32284 to 23.32618)) was very small for both the blocks 1 (eta squared = 0.0001468) and 2 (eta squared = 0.0000705). Thus, it can be inferred that, there is very little statistical difference in values between the control and group, both through chance and magnitude on the whole.

Table 7.8: Group Statistics

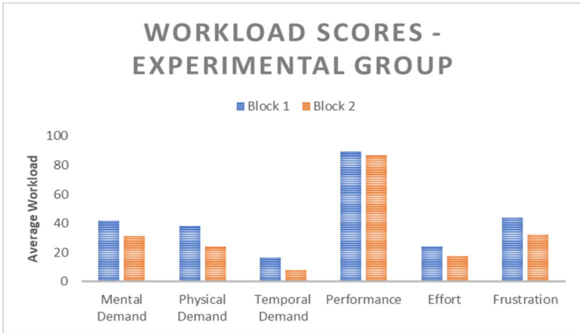
	GROUP	N	Mean	Std. Deviation	Std. Error Mean
B1_AVG	1 Control	6	31.3350	21.15274	8.63557
	2 Experimental	6	35.1667	21.79458	8.89760
B2_AVG	1 Control	6	32.8350	12.45619	5.08522
	2 Experimental	6	35.3333	25.51148	10.41502

Individual Workload Constructs Comparison

However, when individual constructs of the workload from the NASA TLX scale are considered, a significant difference between the average values between the control and experimental groups for both the blocks can be seen. Before asserting the results based on the data, it is necessary to consider that during the usability study, the participants in the control group are provided with the AI-based prototype first and then the manual prototype. The reverse is true for the participants from the experimental group where they interacted with the manual prototype first followed by the AI-based prototype.



(a) Workload Values - Control Group



(b) Workload Values - Experimental Group

Figure 7.3: Workload Values for Individual NASA TLX Constructs

For the participants placed in the control group, the perceived that interacting with the manual prototype posed a slight increase in mental (M = 48.32) and physical demand (M = 25.16) than interacting with the AI-based prototype (mental demand : M - 40.77 ; physical demand : M = 21.3). The same trend can also be seen from the ratings given by experimental group, where in this case, a much more significantly larger difference in perceived workload between the manual (mental demand : M = 42.04 ; physical demand : M = 38.5), and the AI-based prototype (mental demand : M = 31.3 ; physical demand : M = 24.125). Even though almost all of the participants, both in the control and experimental group were successful in completing the given tasks (***performance*** scores are same in both the figure 7.3), they felt that interacting with the manual prototype required a slightly higher effort when compared to accomplishing tasks with the AI-based prototype. These can be seen in the above figure, where for the control group, the effort ratings for the manual prototype (M = 38.68) were found to be slightly higher than that of the AI-based prototype (M = 32.57). The same trend can also be observed in the ratings provided by the participants in the experimental group (manual : M = 24.125 ; AI-based : M = 17.65). Furthermore, the participants, both in the control and experimental group felt a little bit more frustrated with the interaction choices offered by the manual prototype over the interactions offered by the AI-based one.

7.6.3 Analyzing the user's mental model of the system

The participants both from the control and experimental group perceived a less amount of workload when interacting with the AI-based prototype than the manual one, which are explained in detail in the previous sections. Now, in order to make the AI-based prototype resonate with the user, it is essential to identify the mental model that the user develops while interacting with the AI-based prototype. So, based on the predicted outcomes that the user's might arrive, which is elaborated in the table 7.3, tracing of the user's mental model is done, based on the process mentioned in the phase 5 of the experimental design (refer section 7.3). Based on the data gathered from the participants' responses to the questions, similar to the analysis done during the pre-study, the number of answers that either fully match (FM), partially match (PM), or does not match (NM) with the mental model requirement is coded, and illustrated in the table 7.9 below.

Table 7.9: Matching the participants responses with the expected answers

ID	Questions	FM	PM	NM
Q1	Why was the system customized in that specific way?	9	3	-
Q2	Based on what context the customization was done?	11	1	-
Q3	Based on the customization tasks showcased in the video, how certain are you that the AI adheres to customization context?	8	3	1
Q4	Are there any other possibilities of performing this customization in a different way?	7	2	3
Q5	What if the system does not offer this kind of customization and why?	9	2	1

From the above table it can be seen that most of the participants are able develop significant understanding of the system and are able to trace back to interactions offered while performing the tasks. By delving deep into this, the responses provided by the participants these questions are summarized below, and are then referred to the user's mental model.

Q1 : Why was the system customized in that specific way?

Responses given by the participants for this questions are mostly in-line or partially in-line with the expected answers, which means that the participant's predicted that the current actions taken by the system, is due to learning the user's recurring behavior while interacting with the system, thereby satisfying the mental model of this requirement. When looking into the 3 partially matching answers, the participants understood the intentions of the system at that particular context, but they didn't recognize the reason behind those actions; however, they accurately anticipated the system's behavior, and hence their mental model was also satisfied to an extent.

Q2 : Based on what context the customization was done?

For this questions, all of the participants except one, exhibited the correct mental model by accurately interpreting and understanding the context/use-case based on which the system provided specific features and/or setting for customization. The one participant who exhibited a partial mental model towards this requirement was the same one from the pre-study, whose response can be seen in the table 7.6.

Q3 : Based on the customization tasks showcased in the video, how certain are you that the AI adheres to the customization?

The mental model exhibited by 8 participants totally resonated with their expectations of certainty of the customization based on the showcased context. For the other participants, one participant had a significant gap in their mental model regarding the workings of the system, and the rest of the 3 participants also had a small gap in their mental model; however, their ratings were significantly closer to that of the expected answer. The reasons according them were, they were able to reason behind the settings/features offered by the system but the system still needs a bit more flexibility in allowing them to customize it significantly.

Q4 : Are there any other possibilities of performing this customization in a different way?

Even though most of the participants (7 of them) agreed with the actions taken by the system for customizing the settings and/or features for them, a significant gap in mental model is observed for 5 other participants. When looking into the rest of the participants, 2 of them partially resonated with the mental model exhibiting their interest in changing the layout frequently, even where the customization features were not offered by the system. When these 2 users were prompted to add any other responses in addition to this, no other ideas came to them at that time, and for this reason their responses were partially in-line with the mental model requirement. This does not mean that they did not develop the required mental model of the system; instead, was very slow in developing them. Similar to the previous case, out of the other 3 participants. one of the participant did not even provide a response to this question, as no ideas came to their mind when they came across this question. The rest of the participants suggested some improvements (not mentioned in the expected answers) to the existing methods of customization.

Q5 : What if the system does not offer this kind of customization and why?

Most of the participants (partially)agreed that for the occupants who are used to these types of system, and if the system does not offer these kind of customization features, they will be frustrated and have difficulty in interacting with the system. Some of the participants also provided the perfect answer of whether there might be a problem in invoking the AI assistant, where if the occupants does not know about this, their behavior will not be influenced. Out of all the participants, only one user said that it wont affect their behavior when these features are not provided to them.

In addition to the above 5 questions, the participants are also asked to provided a scenario, different than they encountered in this usability study, where they would like the system to customize the settings and/or features for them, and the responses of the participants are provided in the appendix F. Looking into the responses provided by the participants, almost all of the participants either fully or partially satisfied the mental model requirement, which in-line with the expected outcome of the hypothesis H3, mentioned in the table 7.4, thereby satisfying it. This marks the end pf phase 5 of the experimental design, and the users were then asked to rate the AI-based prototype based on user experience and acceptance. The results obtained from these questionnaires are explained in the sections below.

7.6.4 Results obtained from User Experience evaluation

Once the evaluation of the perceived workload and the mental model of the user's are done, the next step towards evaluating the pragmatic and hedonic qualities of the prototype is done by using the Attrakdiff survey. The ratings given by the participants for the word pairs presented to them during the survey is analyzed and the resulting semantic differential graph of average values, and the corresponding portfolio-presentation of the prototype is analyzed. Before analyzing these results, first the reliability of the attrakdiff constructs are to be analyzed, which are done by calculating the Cronbach's Alpha on each constructs. Based on the results, the scores corresponding to attractiveness (Alpha = 0.892) and pragmatic quality (Alpha = 0.815) posed very high reliability, since their alpha values are way ahead of the standardized score of 0.7 [17].

However, both the HQ-I (Alpha = 0.608) and HQ-S (Alpha = 0.431) constructs of hedonic quality posed a significantly lesser alpha value than the standard 0.7 value mentioned by DeVellis [17], which is due to the lesser number scales within the constructs. For such short scales, Briggs et al. [12] recommend an optimal range from 0.2 to 0.4 Alpha values, which makes the above two constructs, reliable. Furthermore, a unified hedonic quality construct for this prototype cannot be made, as the alpha scores, even though being reliable, are very less compared to the standardized score of 0.7; but a unified hedonic quality construct can be considered for analyzing the portfolio of the prototype, which can be seen in the figure 7.4.

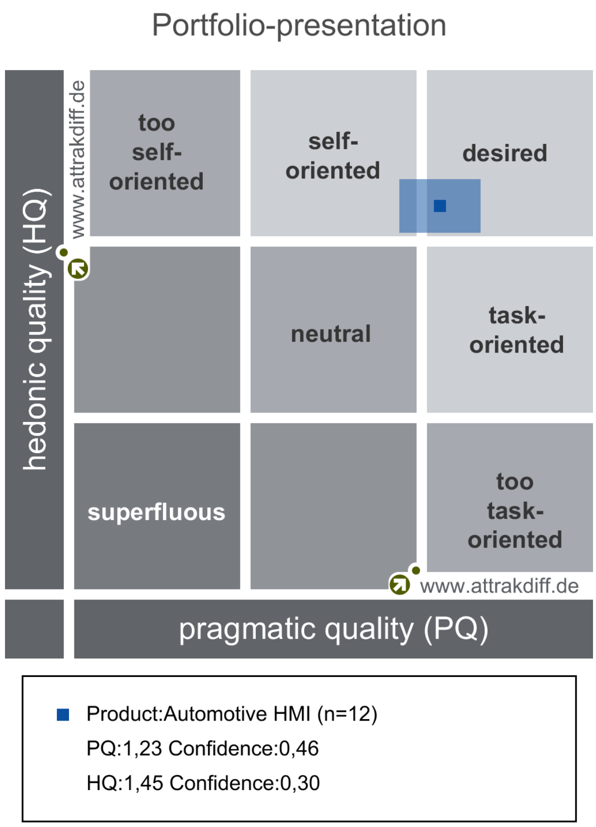


Figure 7.4: Portfolio of the results for the AI-based HMI

From the above diagram it can be seen that, the AI-based HMI prototype was rated very highly both in pragmatic and hedonic quality, which is represented by a small dark blue dot inside the confidence rectangle. By looking into the confidence rectangle of the prototype, it some area of it extends from the desired region towards the self-oriented region. Since, only a small part of the area spans towards that region, it can be said that almost all of the participants felt that the product was desirable. Even though the prototype is declared as desirable, the reason for the area extension towards the self-oriented region should also be examined. This can be done by analyzing the individual word-pairs provided within each of these four constructs. The mean and standard deviation ratings of these 4 qualities are provided in the table 7.10 below.

Table 7.10: Rating provided by the participant's for Attrakdiff qualities

Quality	**Word Pairs**	**Mean**	**Standard Deviation**
PQ	technical - human	3.17	1.030
	complicated - simple	5.33	1.435
	impractical - practical	5.83	1.115
	cumbersome - straightforward	5.25	1.357
	unpredictable - predictable	5.42	1.084
	confusing - clearly structured	5.50	1.314
	unruly - manageable	6.08	.669
HQ-I	isolating - connective	5.25	1.138
	unprofessional - professional	5.92	.900
	tacky - stylish	5.83	.937
	cheap - premium	5.50	1.087
	alienating - integrating	5.42	.996
	separates me - brings me closer	4.33	1.371
	unpresentable - presentable	5.92	.515
HQ-S	conventional - inventive	6.00	1.128
	unimaginative - creative	6.08	.669
	cautious - bold	4.58	1.443
	conservative - innovative	5.50	1.567
	dull - captivating	5.58	.996
	undemanding - challenging	5.08	1.505
	ordinary - novel	5.25	1.288

ATT	unpleasant - pleasant	5.33	1.614
	ugly - attractive	5.75	1.138
	disagreeable - likeable	5.92	1.165
	rejecting - inviting	5.50	1.168
	bad - good	6.33	.888
	repelling - appealing	5.75	.965
	discouraging - motivating	5.58	1.564

From the table it can be seen that, all of the word-pairs in their corresponding qualities recorded a significantly higher average value, which was calculated based on the ratings provided as in a 1 to 7 likert scale, attractiveness in particular, recorded the highest mean average value among the other 3 qualities, as shown in the figure 7.5. The higher mean average values can also be seen in the other three constructs, with notable exceptions. Similar to the result obtained from the pre-study, the participants rated the pragmatic quality scales very highly except for one where they felt that interacting with the AI-based prototype felt a bit too technical for them (PQ ; technical - human ; M = 3.17 ; SD = 1.030). Out of all the scales rated by the participants, only this word pair gained negative ratings. When hedonic qualities of the prototype is taken into consideration, in addition to the significantly higher average ratings, the participants thought the prototype was bold enough (HQ-S ; cautious - bold ; M = 4.58 ; SD = 1.443) to achieve its intended outcomes; however, the participants were significantly divided whether interacting with it separates them or will bring closer to people (HQ-I ; separates me - brings me closer ; M = 4.33 ; SD = 1.371).

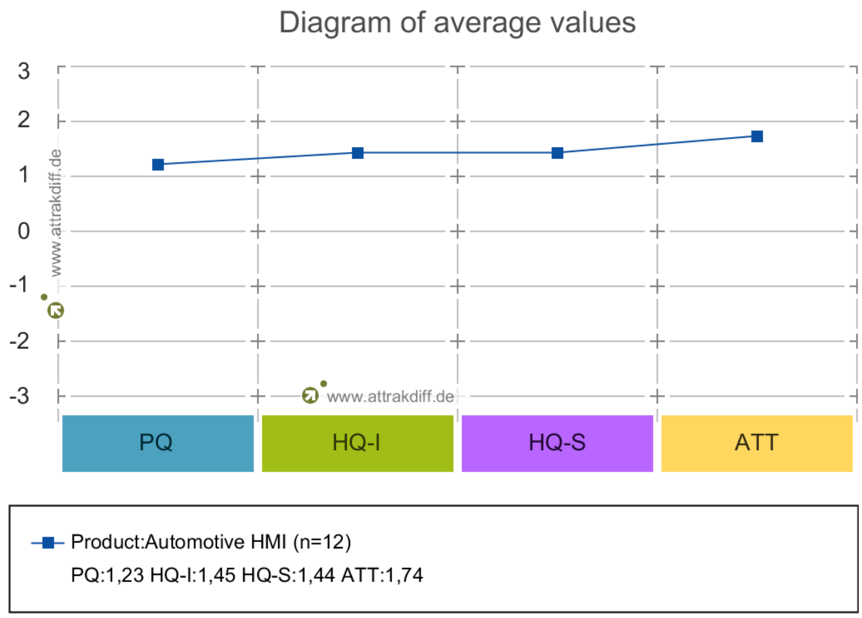


Figure 7.5: Diagram of average values for AI-based HMI

One thing to note from the above diagram, is that it was generated from the attrakdiff website based on the likert scale of -3 to +3, but for analysis through SPSS, the results are reverse coded to reflect the likert scale of 1 to 7. The entire semantic differential diagram representing the average values of the word-pairs for all the qualities as mentioned in the table 7.10 are mentioned in the appendix F.

7.6.5 Analyzing the acceptance scores

The user acceptance of the AI-based prototype is then measured by asking to participants to rate their opinion on a scale based on the CTAM [63], and the rating provided by the participants for the CTAM constructs are given in the table 7.11 and figure 7.6 below.

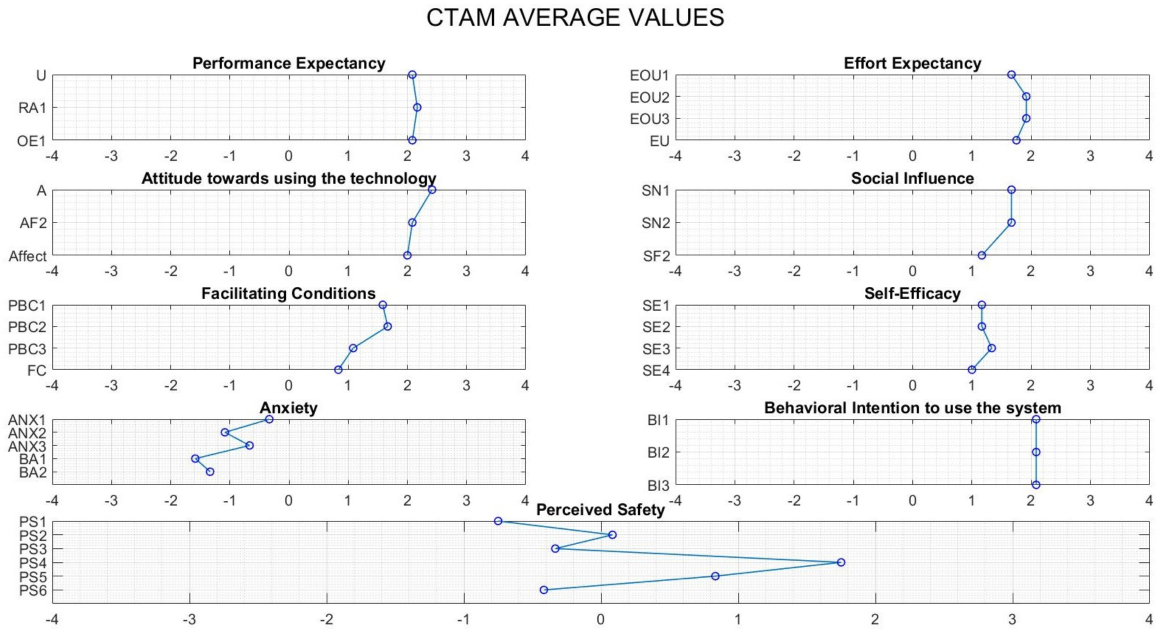


Figure 7.6: Average acceptance values plot for AI-based HMI

Table 7.11: Descriptive Statistics for CTAM scale ratings

		Min	Max	Mean	SD	Variance
Performance Expectancy	U	-1	3	2.08	1.240	1.538
	RA1	-1	3	2.17	1.115	1.242
	OE1	-1	3	2.08	1.505	2.265
Effort Expectancy	EOU1	-1	3	1.67	1.231	1.515
	EOU2	-2	3	1.92	1.443	2.083
	EOU3	-1	3	1.92	1.505	2.265
	EU	-2	3	1.75	1.815	3.295

Attitude towards using the technology	A	0	3	2.42	.900	.811
	AF2	0	3	2.08	1.165	1.356
	Affect	-1	3	2.00	1.414	2.000
Social Influence	SN1	0	3	1.67	1.303	1.697
	SN2	-1	3	1.67	1.155	1.333
	SF2	-1	3	1.17	1.193	1.424
Facilitating conditions	PBC1	-1	3	1.58	1.240	1.538
	PBC2	-2	3	1.67	1.497	2.242
	PBC3	-1	3	1.08	1.379	1.902
	FC	-1	3	.83	1.642	2.697
Self-Efficacy	SE1	-2	3	1.17	1.586	2.515
	SE2	-1	3	1.17	1.467	2.152
	SE3	-2	3	1.33	1.826	3.333
	SE4	-2	3	1.00	1.595	2.545
Anxiety	ANX1	-3	3	-.33	1.923	3.697
	ANX2	-3	0	-1.08	.996	.992
	ANX3	-3	3	-.67	1.875	3.515
	BA1	-3	2	-1.58	1.832	3.356
	BA2	-3	2	-1.33	1.875	3.515
Behavioral intention to use	BI1	-1	3	2.08	1.165	1.356
	BI2	-1	3	2.08	1.165	1.356
	BI3	-1	3	2.08	1.240	1.538
Perceived Safety	PS1	-3	2	-.75	1.658	2.750
	PS2	-3	2	.08	1.881	3.538
	PS3	-3	3	-.33	1.723	2.970
	PS4	0	3	1.75	1.215	1.477
	PS5	-2	3	.83	1.801	3.242
	PS6	-2	3	-.42	1.564	2.447

The CTAM constructs, the variables behind the constructs, and it's corresponding items of these constructs are provided in the appendix A for reference.

Looking into the CTAM constructs, the first construct is the performance expectancy, and all the participants rated this construct with a high positive score. This means that the participants thought that interacting with the HMI would be useful them while commuting (U : M = 2.08 ; SD = 1.240), and also helps them to significantly develop their mental model faster to achieve their goals quickly (RA1 : M = 2.17 ; SD = 1.115). Compared to their ability in achieving their goals quickly, the participant's felt that the interacting with the HMI is easier (EOU3 : M = 1.92 ; SD = 1.505) and the effort required to perform the interactions is also significantly low ((EOU1 : M = 1.67 ; SD = 1.231) and (EOU2 : M = 1.92 ; SD = 1.443), but not as effective as the performance expectancy variables (EU : M = 1.75 ; SD = 1.815). Due to the system enabling the participants to achieve their goals quickly, and being easy to use and understand, interacting with it enabled the user to have a positive belief (A : M = 2.42 ; SD = 0.9), thereby making interacting with the HMI fun (AF2 : M = 2.08 ; SD = 1.165) and likeable (Affect : M = 2 ; SD = 1.414).

One thing to note that while interacting with the prototype is that, the participants felt that the HMI was in the early stages of development, and as a result of this, the ratings for the social influence construct, even though positive, was a bit lower than the performance expectancy and the attitude construct. Even though the participants felt that they would feel very proud in showing this system to their social contacts (SN1 : M = 1.67 ; SD = 1.303), they felt that they would like to encourage others to use the system, albeit not right now (SF2 : M = 1.17 ; SD = 1.193). This belief has also affected the participant's self-efficacy scores too. In spite of being doubtful, the participants felt that they could complete the task/activity provided by the HMI under various conditions (refer the self-efficacy construct in the table 7.11); however, the participants exhibited a significant level of concern about asking someone for assistance in case of any difficulties faced (FC : M = 0.83 ; SD = 1.642). In addition to this, the participants also exhibited some concerns regarding the compatibility of the interface across all platforms, but they believed that this issue will be solved in the later iterations (PBC3 : M = 1.08 ; SD = 1.379).

In spite of this, the participant's felt that they can have a significant amount of control in maintaining their driving behavior (PBC1 : M = 1.58 ; SD = 1.240) and have the necessary knowledge to use the system (PBC2 : M = 1.67 ; SD = 1.497), since the scores of the performance and effort expectancy constructs are significantly positive and higher. On the whole, the participants strongly asserted their intention to use the HMI (BI1 : M = 2.08 ; SD = 1.165) for the next several months (BI3 : M = 2.08 ; SD = 1.240), and also did neither have any concerns (ANX1 : M = -0.33 ; SD = 1.923) nor had any fear while interacting with it (ANX3 : M = -0.67 ; SD = 1.875). As mentioned in the starting of this section, the prototype is in the early stages of development, and it is difficult to analyze the real world implications of it. The same issue the participant's faced while rating the perceived safety construct of this scale. So the participants were asked to think of a situation of interacting with this AI-based HMI in general and are asked to rate based on that. This seems to have a huge impact on the perceived safety scores, since all of the rating except one, raked a near neutral score. One positive aspect from this construct is the participants felt that they feel safe while using the system (PS4 : M = 1.75 ; SD = 1.215).

7.7 Discussion

The main goal of this usability study is to analyze the impact implementation of AI in an exiting car HMI has, in improving the interaction between the human and the HMI. For this, the usability and the user experience of the AI-based HMI prototype were analyzed and the results were produced in the previous sections. As mentioned earlier, since the goal of this usability study is twofold, the first aspect of this study is to analyze the amount of workload that the user's perceive while interacting with AI-based HMI in comparison with the conventional HMI. Even though the results from the independent-samples t-test showed that there is no overall significant difference between these two HMIs, the individual workload constructs of the NASA TLX scale proved to be otherwise. Even though the participants from both the control and experimental group were quite successful in performing the tasks given to them, they exhibited a higher physical and mental demand when interacting with the manual prototype, instead of the AI-based one.

In addition this, the participants also exhibited some frustrations when interacting with the manual prototype, especially the participants from the experimental group felt more frustrated with the interaction choices offered by the manual version. This might be due to them being interacting with the manual prototype first followed by interacting with the AI-based prototype. Comparing the results from the pre-study, where the participants from the experimental group perceived a higher average workload while interacting with the AI-based HMI, the final results, although having a slight difference in the overall perceived workload, matches with the expected outcomes and in-turn satisfies the hypothesis H1.

The second aspect of this usability study is to address the traceability of the model by assessing the mental model that the user develops while interacting with the system. Thus, in-order to make the user to recognize AI-based systems, it is essential to trace the exact scenario of how the user's develop significant knowledge about the system's behavior, and for this, the users are primed regarding some aspects of the system. By matching the participants responses with the expected mental model requirement as shown in the table 7.3, it is safe to assume that mental model of almost all the users are satisfied, except for a few aspects. Even though the participants were able to determine the why and how behind the customization, most of them thought that there were other means based on which these customization's can be done. The above is reasons is in fact a matter of personal preference, since similar responses can also be found in the user opinions survey presented in the chapter 4.

Another important data recorded from the participant's responses is that, even though some of the participant's response fully matched with the expected requirement, the participants are still not 100% certain that the system adhered to the customization context. The biggest reason behind these responses might be the product being in the early stages of development, and more functionalities are needed to be added. In addition to this, there might be another possibility concerning the nature of the user testing, which was carried out virtually. If the same study were to be performed in a physical simulator environment, then there might be a slight change in the participant's responses. Thus, even though a small number of participant's responses did bot match with the expect mental model requirement, it is safe to say that the gap present in the user's mental model were significantly lower, and thus confirming the satisfaction of hypothesis H2.

Finally the pragmatic and hedonic qualities of the AI-based HMI prototype is analyzed along with the measuring its user acceptance. Based on the scores, it can be seen that both the pragmatic and hedonic qualities of the prototype were significantly higher, which can be referred from the prototype presentation as shown in the figure 7.4. In addition to this, the average values of all the attrakdiff qualities also raked a significantly higher positive score, making the prototype desirable across all platforms. However, one exception can be found in the ratings, were the participants unanimously agreed that interacting with the AI-based prototype was a bit too technical for them. The same trend for this specific word-pair can also been from the participant's responses from the pre-study. Thus, the participants felt that even though the AI-based prototype was usable, it still leans a bit towards the technical side; however, they also expressed that by frequent learning, both by the system and by the user, the ratings might shift towards the human side.

The trend similar to that seen in the user experience scores, can be seen in the user acceptance scores, where the ratings provided by the participants for the CTAM constructs, are mostly in-line with the expected outcomes of the research questions. As mentioned in the section 7.6.5, interacting with the prototype increased their beliefs (attitude towards using the technology) regarding achieving their goals (performance expectancy), and also its subsequent ease-of-use (effort expectancy). In addition to these, the participant's felt that they had a significant amount of control (facilitating conditions) while interacting with the HMI, which in turn increased their intention to use the prototype all times (behavioral intention to use) and also will enable them to recommend this HMI to their social contacts (social influence). Even though minimal, they neither did exhibit any frustrations nor any fears when interacting with these AI-based HMI (anxiety). Even though the participants provided some neutral scores in some of the out-of-context variables, in the end, they felt safe when using/interacting with the system (perceived safety). In the end, the ratings observed from the user experience and the acceptance scale reflects a significant improvement in the user experience and acceptance when presented with an AI-based HMI, which in turn satisfies the hypothesis H3.

7.8 Summary

In this chapter, the final prototype designed by addressing the usability problems identified in the heuristic evaluation, is subjected to usability testing based on a determined experimental design. Before proceeding with the user study, a pre-study was conducted with a part of the research population with a main objective of finding some final problems within the experimental design. Once no changes are identifies, the final usability study was carried out with a main objective of measuring the user experience and acceptance towards using the AI-based HMI, and the subsequent mental model that the user develops while interacting with the prototype is analyzed and the results are discussed in detail.

Chapter 8

Conclusion

The impact the incorporation of artificial intelligence has on the interactions offered by the human machine interfaces of the vehicle, and its significant influence on the user experience and acceptance, has been explored in detail in this thesis. From the inception of this thesis to its end, the entire research contributions and findings from the usability study, are summarized in this chapter. From these contributions, the potential research that has to be carried out in the future of analyzing and implementing these AI-based human machine interfaces are listed out, followed by the concluding remarks.

8.1 Research Summary

When looking into to the thesis scope mentioned in the chapter 1, the main factors that are necessary to make the product usable and to motivate the users to accept and use the product in the future, is first by identifying the amount of workload that the user's perceive while interacting with the system, followed by analyzing the mental model that the user develops while interacting with the system. In order to measure and analyze this, the research that has been done in entirety in thesis is summarized below.

8.1.1 Interaction Guidelines to support development of AI-based HMI

As mentioned in chapters 1 and 2, most of the automakers are scaling the implementation of AI based systems into their workflows to suit their individual needs, and if these AI systems are to be incorporated into other workflows, then modifications are to be required at a large extent, and there arises the problem of inconsistencies. Even though many guidelines and requirements are in place for designing these AI-based system, consistency is still a problem, and hence to address this, as seen in the chapter 3, a thorough systematic review of all the guidelines has to been carried out to propose a set of interaction guidelines that can be consistent when considered for implementation and also to support the development of these AI-based interfaces. Since these systems are continuously developing, in order to keep the guidelines fresh and relevant, significant efforts has been made in such a way that majority of the materials that are used as a reference for drafting these guidelines are based on journals, blogs and patents that are/has been published in the last decade. Some exceptions has been made for older guidelines also, since these are considered as standards.

8.1.2 Identifying the inconsistencies in the existing HMIs

Before progressing further into the design of the interface using these guidelines, it would be essential first to use this guidelines as a reference to find the inconsistencies present in the HMIs that are available in the market and on the patents that are either pending till now or approved. As seen in the chapter 4, in order to identify this, over HMIs of 14 cars that are available in the market and 8 patents that are filed in the topic of AI and HMI, are sorted out for this systematic comparison, where the functionalities that these HMIs offer at every levels of interaction are checked for match by using the guidelines as reference. As expected, the results from this systematic comparison yielded several inconsistencies in the overall build of the HMI, and the inconsistencies found were used a building block for elicitating the requirements that the prototype has to be designed upon.

8.1.3 Understanding the User's Needs and Expectations

Apart from identifying the inconsistencies from the existing HMIs, in order to truly design a product that the user can accept, it is of atmost importance to gather their opinions about AI in general followed by identifying their needs and expectation when interacting with an AI-based HMI. In order to gather a more comprehensible responses, the use cases in which the incorporation of AI will pose a significant impact on the interaction between the user and the system, has to be considered, which based on another systematic review, yielded navigation and customization as the viable ones.

Analyzing the responses from 50 participants, provided significant insights which are further carried over along with the results from the systematic comparison towards the elicitation of functional and non-functional requirements as seen in the chapter 5, and based on these along with the technical guidelines, the HMI prototype was designed for usability testing.

8.1.4 Evaluating the User Experience and Acceptance of AI-based HMI

Once the usability problems has been identified in the designed prototype based on some expert reviews as seen in the chapter 6, the HMI prototype is then subjected to usability study conducted by remote moderated usability testing method virtually as per the experimental design mentioned in the chapter 7.

After acquiring proper consent from the participants, the usability study is carried out and during the course of the study, the amount of workload that the participants perceive while interacting with the prototype, and the mental model that they develop while interacting with the prototype are analyzed. Finally, the pragmatic and hedonic qualities of the prototype, along with their motivation to accept the interactions offered by the prototype is analyzed via Attrakdiff and CTAM questionnaire.

8.2 Answers to the Research Questions

Following the usability study, the responses from the participants are collected and the quantitative data were then subjected to descriptive statistics calculation. Based on the test results discussed in the chapter 7, the answers to the research questions raised in the 1^{st} chapter of this document are provided below.

How does an AI based HMI of an autonomous vehicle, allows the occupants to interact and customize with its interface, resulting in an improved user experience and acceptance of the system?

After completing the provided interactions, the AI-based prototype scored a significantly higher scores on both the user experience and acceptance scales. Even though most of the participants thought interacting with the AI-based HMI was a bit too technical for them, they were motivated enough to accept the levels of interactions and customizations offered by the HMI, and significantly concluding the AI-based HMI was desirable. Due to restrictions of not interacting with the prototype physically and in-person, the ratings for the perceived safety in the acceptance scale did not reflect with the expected outcomes but in the end, the participant's thought that the AI-based prototype was safe to use.

How well does the occupants of the autonomous vehicle, understand and anticipate the behavior of the system, based on the context that was presented to them?

In order to measure the traceability of the HMI, the participants are then primed about some questions about the levels of customization offered by the AI-based HMI, and their responses are then matched with the expected answers. Even though all of the participants were not 100% certain about the level so f customization offered to them, out of 12 participants, the mental model of 11 participants resonated well with the prototype as they were able to predict and anticipate the behavior and reason behind the customizations. This evidently concludes that the users were able to their mental model with the underpinnings of the AI-based HMI.

8.3 Research Boundaries and Future Work

Owing to the COVID 19 pandemic that struck in 2020, it was decided to conduct this study virtually through the method of remote moderated usability testing. Even though this method has some advantages, the data collected from these usability tests can sometimes be often less detailed than the one conducted from in-person tests. These kind of drawbacks are even more significant in the case of physical products. In this study context, even though the results obtained from analyzing the data satisfied with the research hypothesis, the quality of the obtained would have been much higher if a realistic simulated environment was present during testing. Furthermore, analyzing the physical behavior exhibited by the participant while interacting with the prototype was also not possible to the online nature of the testing. Even though being an additional answer, analyzing the physical behavior of the participants will also provide more assertion to the research hypothesis, and which significantly improve the mental model of the participants even further. Also, there might be a case where a significant difference in the participants' perceived workload can be observed if the testing was carried out in-person instead of being completely online. Although causing an impact on the study, these limitations provides a potential scope of improvement in the future.

Also the prototype, even though usability tested to find some potential problems, the development of it is still far ahead and still new functionalities has to be added in order to motivate the user to accept and interact with it. While analyzing the mental model of the user, the participant's were asked to provide a situation in a context different than the use cases considered, and it is suffice to say from their responses is that, even though they accept the technologies used in the system, they will set high expectations while interacting it. This case not only represented for the two use cases being considered in this project, but also for the entire functionalities and features that these systems offer. As the HMIs being designed today are based on considering specific methodological choices ranging from the technical specifications to the customer base in order to increase the confidence associated with the provided result making the future assumptions explicit, it is of absolute essential to design these interfaces that understands and takes into consideration the mindsets of the users, each and every time during post launch monitoring. Since the user's understanding of the systems keeps on changing based on the products life cycle, learning from the past and understanding the interface in the plausible future will significantly aid towards improvement of these AI-based user interfaces, thereby making them accepted by the consumers.

Bibliography

[1] Amditis, A., Andreone, L., Pagle, K., Markkula, G., Deregibus, E., Rue, M. R., Bellotti, F., Engelsberg, A., Brouwer, R., Peters, B., et al. Towards the automotive hmi of the future: overview of the aide-integrated project results. *IEEE Transactions on Intelligent Transportation Systems 11*, 3 (2010), 567–578. 57

[2] Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems* (2019), pp. 1–13. 27, 29, 30, 31

[3] Antifakos, S., Kern, N., Schiele, B., and Schwaninger, A. Towards improving trust in context-aware systems by displaying system confidence. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services* (2005), pp. 9–14. 31

[4] Arnold, V., Clark, N., Collier, P. A., Leech, S. A., and Sutton, S. G. Explanation provision and use in an intelligent decision aid. *Intelligent Systems in Accounting, Finance & Management: International Journal 12*, 1 (2004), 5–27. 29, 30, 31

[5] Bader, R., Siegmund, O., and Woerndl, W. A study on user acceptance of proactive in-vehicle recommender systems. In *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2011), pp. 47–54. 29, 30, 31

[6] Bakker, S., van den Hoven, E., and Eggen, B. Peripheral interaction: characteristics and considerations. *Personal and Ubiquitous Computing 19*, 1 (2015), 239–254. 54, 55

[7] Barthelmess, P., and Kaiser, E. Workshop on tagging, mining and retrieval of human related activity information. In *Proceedings of the 9th international conference on Multimodal interfaces* (2007), pp. 383–384. 26

[8] Becker, J., Joachim, K., Koldewey, C., Reinhold, J., and Dumitrescu, R. Scaling digital business models: A case from the automotive industry. In *ISPIM Conference Proceedings* (2020), The International Society for Professional Innovation Management (ISPIM), pp. 1–15. 36, 39

[9] Beckles, B., Welch, V., and Basney, J. Mechanisms for increasing the usability of grid security. *International Journal of Human-Computer Studies 63*, 1-2 (2005), 74–101. 29

[10] Benčić, F., Nemet, S., Barać, M., and Kukolj, D. Trends in ai based automotive industry using patent analysis. In *26th TELFOR 2018* (2018). 1, 2, 3

[11] Bevan, N., Carter, J., and Harker, S. Iso 9241-11 revised: What have we learnt about usability since 1998? In *International Conference on Human-Computer Interaction* (2015), Springer, pp. 143–151. 62, 65

[12] Briggs, S. R., and Cheek, J. M. The role of factor analysis in the development and evaluation of personality scales. *Journal of personality 54*, 1 (1986), 106–148. 88

[13] Carroll, J. M., and Olson, J. R. Mental models in human-computer interaction. In *Handbook of human-computer interaction*. Elsevier, 1988, pp. 45–65. 32, 73

[14] Cohen, J. *Statistical power analysis for the behavioral sciences*. Academic press, 2013. 84

[15] Davis, F. D. *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. PhD thesis, Massachusetts Institute of Technology, 1985. 8

[16] Delligatti, L. *SysML distilled: A brief guide to the systems modeling language*. Addison-Wesley, 2013. 51

[17] DeVellis, R. F. *Scale development: Theory and applications*, vol. 26. Sage publications, 2016. 88

[18] Díaz-Oreiro, I., López, G., Quesada, L., and Guerrero, L. A. Standardized questionnaires for user experience evaluation: A systematic literature review. In *Multidisciplinary Digital Publishing Institute Proceedings* (2019), vol. 31, p. 14. 8

[19] Dickmanns, E. D. *Dynamic vision for perception and control of motion*. Springer Science & Business Media, 2007. 12

[20] Doshi-Velez, F., and Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608* (2017). 31

[21] Dove, G., Halskov, K., Forlizzi, J., and Zimmerman, J. Ux design innovation: Challenges for working with machine learning as a design material. In *Proceedings of the 2017 chi conference on human factors in computing systems* (2017), pp. 278–288. 26, 29, 30, 31, 32

[22] Driving, A. Levels of driving automation are defined in new sae international standard j3016: 2014. *SAE International: Warrendale, PA, USA* (2014). 1

[23] Eeles, P. Capturing Architectural Requirements What Is an Architectural Requirement? *IBM Rational developer works* (2005). 51

[24] Eiband, M., Buschek, D., and Hussmann, H. How to support users in understanding intelligent systems? structuring the discussion. *arXiv preprint arXiv:2001.08301* (2020). xii, 26, 32, 33

[25] Fenton, R. E., and Olson, K. W. The electronic highway. *IEEE spectrum 6*, 7 (1969), 60–66. 11, 20

[26] Garzon, S. R. Intelligent in-car-infotainment systems: A contextual personalized approach. In *2012 Eighth International Conference on Intelligent Environments* (2012), IEEE, pp. 315–318. 48

[27] Gasser, T. M., and Westhoff, D. Bast-study: Definitions of automation and legal issues in germany. In *Proceedings of the 2012 road vehicle automation workshop* (2012), Automation Workshop. 16

[28] Geddes, N. B. *Magic motorways*. Read Books Ltd, 2011. 10

[29] Glass, A., McGuinness, D. L., and Wolverton, M. Toward establishing trust in adaptive agents. In *Proceedings of the 13th international conference on Intelligent user interfaces* (2008), pp. 227–236. 26, 29, 30, 31, 32

[30] Hamilton, K., Karahalios, K., Sandvig, C., and Eslami, M. A path to understanding the effects of algorithm awareness. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 2014, pp. 631–642. 31

[31] Hart, S. G. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (2006), vol. 50, Sage publications Sage CA: Los Angeles, CA, pp. 904–908. 78

[32] Hart, S. G., and Staveland, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52. Elsevier, 1988, pp. 139–183. 73, 74

[33] Hengstler, M., Enkel, E., and Duelli, S. Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change 105* (2016), 105–120. 2

[34] Höök, K. Steps to take before intelligent user interfaces become real. *Interacting with computers 12*, 4 (2000), 409–426. 3, 26

[35] Horvitz, E. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (1999), pp. 159–166. 26, 29, 30, 31

[36] International, S. Automated driving levels of driving automation are defined in new sae international standard j3016, 2014. 16

[37] Jobin, A., Ienca, M., and Vayena, E. The global landscape of ai ethics guidelines. *Nature Machine Intelligence 1*, 9 (2019), 389–399. 27, 30, 31, 32

[38] Kayacik, C., Chen, S., Noerly, S., Holbrook, J., Roberts, A., and Eck, D. Identifying the intersections: User experience+ research scientist collaboration in a generative machine learning interface. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), pp. 1–8. 31, 32

[39] Kitchenham, B. Procedures for performing systematic reviews. *Keele, UK, Keele University 33*, 2004 (2004), 1–26. 5

[40] Kraut, R. E., and Resnick, P. *Building successful online communities: Evidence-based social design*. Mit Press, 2012. 6

[41] Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., and Wong, W.-K. Too much, too little, or just right? ways explanations impact end users' mental models. In *2013 IEEE Symposium on Visual Languages and Human Centric Computing* (2013), IEEE, pp. 3–10. 32, 33, 70, 73

[42] Lallemand, C., Koenig, V., Gronier, G., and Martin, R. Création et validation d'une version française du questionnaire attrakdiff pour l'évaluation de l'expérience utilisateur des systèmes interactifs. *European Review of Applied Psychology 65*, 5 (2015), 239–252. 8

[43] Laugwitz, B., Held, T., and Schrepp, M. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and usability engineering group* (2008), Springer, pp. 63–76. 7

[44] Lavie, T. Adaptive user interfaces for in-vehicle devices. In *International Conference on User Modeling* (2005), Springer, pp. 509–511. 59

[45] Liao, Q. V., Gruen, D., and Miller, S. Questioning the ai: Informing design practices for explainable ai user experiences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (2020), pp. 1–15. 29, 30, 31

[46] Lieberman, H. User interface goals, ai opportunities. *AI Magazine 30*, 4 (2009), 16–16. 3

[47] Lim, B. Y., Dey, A. K., and Avrahami, D. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), pp. 2119–2128. 29, 30, 31, 32

[48] Llaneras, R. E., and Singer, J. P. In-vehicle navigation systems: Interface characteristics and industry trends. 46

[49] Luger, E., and Sellen, A. " like having a really bad pa" the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (2016), pp. 5286–5297. 26

[50] Lungaro, P., Tollmar, K., and Beelen, T. Human-to-ai interfaces for enabling future on board experiences. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct* (2017), pp. 94–98. 3

[51] Marnau, N. Comments on the "draft ethics guidelines for trustworthy ai" by the high-level expert group on artificial intelligence. 27, 30, 31

[52] McNee, S. M., Riedl, J., and Konstan, J. A. Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems* (2006), pp. 1097–1101. 30

[53] Meixner, G., Orfgen, M., and Kuemmerling, M. automotive hmi - model-based in-vehicle infotainment description as a stepping stone for the integration of car and web. 4

[54] Monroe, J. J., Denby, R. M., Zhang, L., Maunders, R. S., and Piche, M. E. Vehicle information/entertainment management system, May 15 2018. US Patent 9,970,768. 32

[55] Mori, M., MacDorman, K. F., and Kageki, N. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine 19*, 2 (2012), 98–100. 21

[56] Ng, A. Ai transformation playbook: how to lead your company into the ai era. *Landing AI* (2018). 13

[57] Nielsen, J. *Usability engineering*. Morgan Kaufmann, 1994. 62

[58] Nielsen, J., and Molich, R. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1990), pp. 249–256. 4, 6, 23, 62, 65

[59] Norman, D. *The design of everyday things: Revised and expanded edition*. Basic books, 2013. 54, 55

[60] Norman, D. A. How might people interact with agents. *Communications of the ACM 37*, 7 (1994), 68–71. 26

[61] Normark, C. J., and Gärling, A. Assessment of automotive visual display guidelines and principles: A literature review. *The Design Journal 14*, 4 (2011), 446–474. 23, 61

[62] Ohashi, K., Katayama, A., Hasegawa, N., Kurihara, H., Yamamoto, R., Doerr, J., and Magin, D. P. Focusing requirements elicitation by using a ux measurement method. In *2018 IEEE 26th International Requirements Engineering Conference (RE)* (2018), IEEE, pp. 347–357. 36

[63] Osswald, S., Wurhofer, D., Trösterer, S., Beck, E., and Tscheligi, M. Predicting information technology usage in the car: towards a car technology acceptance model. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2012), pp. 51–58. xii, 8, 74, 79, 91

[64] Pallant, J. *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge, 2020. 81, 83, 84

[65] Parasuraman, R., and Miller, C. A. Trust and etiquette in high-criticality automated systems. *Communications of the ACM 47*, 4 (2004), 51–55. 30, 31, 32

[66] Patrick, E. Intelligent user interfaces: Introduction and survey. 4

[67] Pfleging, B., Rang, M., and Broy, N. Investigating user needs for non-driving-related activities during automated driving. In *Proceedings of the 15th international conference on mobile and ubiquitous multimedia* (2016), pp. 91–99. 20, 43, 44

[68] Pfleging, B., Schneegass, S., and Schmidt, A. Multimodal interaction in the car: combining speech and gestures on the steering wheel. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2012), pp. 155–162. 21

[69] Pieters, W. Explanation and trust: what to tell the user in security and ai? *Ethics and information technology 13*, 1 (2011), 53–64. 29, 31

[70] Pu, P., and Chen, L. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems 20*, 6 (2007), 542–556. 30, 31

[71] Raghunathan, S. K., and Jaswal, A. Methods and systems for providing personalized and context-aware suggestions, Apr. 17 2014. US Patent App. 13/747,509. 29

[72] Raman, M. Method and system for providing driving quality feedback and automotive support, Oct. 30 2014. US Patent App. 13/871,585. 29

[73] Ribeiro, M. T., Singh, S., and Guestrin, C. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (2016), pp. 1135–1144. 30

[74] Ricci, C. P. Proactive machine learning in a vehicular environment, Oct. 16 2014. US Patent App. 14/253,727. 29, 30, 31

[75] Riegelsberger, J., and Sasse, M. A. Ignore these at your peril: Ten principles for trust design. 30

[76] Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies 62*, 3 (2005), 381–422. 30, 31

[77] Riener, A., Boll, S., and Kun, A. L. Automotive user interfaces in the age of automation (dagstuhl seminar 16262). In *Dagstuhl reports* (2016), vol. 6, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 20

[78] Ross, T., Midtland, K., Fuchs, M., Pauzié, A., Engert, A., Duncan, B., Vaughan, G., Vernet, M., Peters, H., Burnett, G., et al. Hardie design guidelines handbook. *Human Factors Guidelines for Information Presentation by ATT Systems 530* (1996). 23

[79] Russell, S., and Norvig, P. Artificial intelligence: a modern approach. 2

[80] Rzepka, C., and Berger, B. User interaction with ai-enabled systems: a systematic review of is research. 21

[81] Sambasivan, N., and Holbrook, J. Toward responsible ai for the next billion users. *interactions 26*, 1 (2018), 68–71. 31

[82] Schalk, T. B., Faoro, P., He, Y., and Hirschenberger, F. Multimodal user interface design, Sept. 18 2014. US Patent App. 14/195,242. 47

[83] Schelenz, L., Segal, A., and Gal, K. Best practices for transparency in machine generated personalization. *arXiv preprint arXiv:2004.00935* (2020). 3, 29, 30, 31, 32

[84] Sherman, D. K., and Cohen, G. L. Accepting threatening information: Self–affirmation and the reduction of defensive biases. *Current directions in psychological science 11*, 4 (2002), 119–123. 30

[85] Shneiderman, B., and Plaisant, C. *Designing the user interface: strategies for effective human-computer interaction.* Pearson Education India, 2010. 23

[86] Silveira, M. S., Barbosa, S. D., and Souza, C. S. *Getting Users to Understand Designers via Help Systems.* PUC, 2000. 73

[87] Soegaard, M. The basics of user experience design. *Interaction Design Foundation, ed* (2018). 5

[88] Söllner, M., Hoffmann, A., Hoffmann, H., and Leimeister, J. M. Towards a theory of explanation and prediction for the formation of trust in it artifacts. 29, 31

[89] Stevens, A., Quimby, A., Board, A., Kersloot, T., and Burns, P. Design guidelines for safety of in-vehicle information systems. *TRL Limited* (2002). 23

[90] Stumpf, S., Rajaram, V., Li, L., Burnett, M., Dietterich, T., Sullivan, E., Drummond, R., and Herlocker, J. Toward harnessing user feedback for machine learning. In *Proceedings of the 12th international conference on Intelligent user interfaces* (2007), pp. 82–91. 29, 30, 31

[91] Stumpf, S., Rajaram, V., Li, L., Wong, W.-K., Burnett, M., Dietterich, T., Sullivan, E., and Herlocker, J. Interacting meaningfully with machine learning systems: Three experiments. *International Journal of Human-Computer Studies 67*, 8 (2009), 639–662. 30

[92] Tintarev, N., and Masthoff, J. Effective explanations of recommendations: user-centered design. In *Proceedings of the 2007 ACM conference on Recommender systems* (2007), pp. 153–156. 29, 30

[93] Tintarev, N., and Masthoff, J. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction 22*, 4-5 (2012), 399–439. 30, 31

[94] Trimble, T. E., Bishop, R., Morgan, J. F., Blanco, M., et al. Human factors evaluation of level 2 and level 3 automated driving concepts: Past research, state of automation technology, and emerging system concepts. 16

[95] Tzirkel-Hancock, E., Malka, I., Winter, U., Custer, S. D., and Pop, D. P. System for intelligent passenger-vehicle interactions, Aug. 3 2017. US Patent App. 15/159,347. 31

[96] Uga, B. Towards trustworthy ai: A proposed set of design guidelines for understandable, trustworthy and actionable ai, 2019. 3, 31, 32

[97] Venkatesh, V. Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. *Information systems research 11*, 4 (2000), 342–365. 8

[98] Venkatesh, V., and Bala, H. Technology acceptance model 3 and a research agenda on interventions. *Decision sciences 39*, 2 (2008), 273–315. 8

[99] Vöst, S., and Wagner, S. Towards continuous integration and continuous delivery in the automotive industry. *arXiv preprint arXiv:1612.04139* (2016). 15

[100] Walter, N. *Personalization and context-sensitive user interaction of in-vehicle infotainment systems.* PhD thesis, Technische Universität München, 2019. 32, 59

[101] West, D. M. Moving forward: self-driving vehicles in china, europe, japan, korea, and the united states. *Center for Technology Innovation at Brookings: Washington, DC, USA* (2016). 1

[102] Wickens, C. D., Gordon, S. E., Liu, Y., et al. An introduction to human factors engineering. xii, 22, 23, 63

[103] Yang, Q., Yin, J., Ling, C., and Pan, R. Extracting actionable knowledge from decision trees. *IEEE Transactions on Knowledge and data Engineering 19*, 1 (2006), 43–56. 29

[104] Yuan, J., Di Guardo, M. C., Sterling, R., and Alonso, R. Exploring end-user acceptance and exploitation routes of industrial use cases related to ai and hmi. Tech. rep., EasyChair, 2020. 3

[105] Złotowski, J., Proudfoot, D., Yogeeswaran, K., and Bartneck, C. Anthropomorphism: opportunities and challenges in human–robot interaction. *International journal of social robotics 7*, 3 (2015), 347–360. 31

Appendix A

CTAM Constructs

Performance Expectancy

U	The system would be useful while driving.
RA1	Using the system enables me to accomplish my goals more quickly.
RA2	Using the system increases my driving performance.
OE	If I would use the system I will reach my destination safely.

Effort Expectancy

EOU1	My interaction with the system would be clear and understandable.
EOU2	It would be easy for me to become skillful at using the system.
EOU3	I find the system easy to use.
EU	Learning how to operate the system is easy for me.

Attitude towards using technology

A	Using the system is a good idea.
AF1	The system makes driving more interesting.
AF2	Interacting with the system would be fun.
Affect	I would like interacting with the system.

Social Influence

SN1	I would be proud to show the system to people who are close to me.
SN2	People whose opinions are important to me would like the system too.
SF1	My passenger(s) would be helpful when using the system.
SF2	In general, people who I like would encourage me to use the system.

Facilitating Conditions

PBC1	While using the system I can maintain an safely driving behavior.
PBC2	I have the knowledge necessary to use the system.
PBC3	The system is compatible with other systems I use.
FC	There would be somebody I can ask for assistance with system difficulties.

Self-efficacy : I could complete a task or activity using the system..

SE1	.. if there was no one around to tell me what to do.
SE2	.. if I could call someone for help if I got stuck.
SE3	.. if I had a lot of time.
SE4	.. if I had just the built-in help facility for assistance.

Anxiety

ANX1	I have concerns about using the system.
ANX2	I think I could have an accident because of using the system.
ANX3	The system is somewhat frightening to me.
BA1	I fear that I do not reach my destination because of the system.
BA2	I am afraid that I do not understand the system.
BA3	I am confident that the system does not affect my driving.

Behavioral Intention to use the system

BI1	Assuming I had access to the system, I intend to use it.
BI2	Given that I had access to the system, I predict that I would use.
BI3	If the system is available I plan to use the system in the next months.

Perceived Safety

PS1	I believe that using the system is dangerous.
PS2	Using the system requires increased attention.
PS3	The system distracts me from driving.
PS4	I feel save while using the system.
PS5	Using the system decreases the accident risk.
PS6	I can use the system without looking at it.

U = perceived usefulness; RA = relative advantage; OE = outcome expectations; EOU = perceived ease of use; EU = ease of use; A = attitude; AF = affect towards use; Affect = affect; SN = subjective norm; SF =social factors; PBC = perceived behavioral control; FC = facilitating conditions; SE = self efficacy; ANX = anxiety; BA = behavioral anxiety ; BI = behavioral intention to use technology; PS = perceived safety

Appendix B

Systematic Comparison Results

	AI INTERACTION GUIDELINES	Audi - MMI		BMW - iDrive	Honda	Hyundai - SmartSense	MINI	Mercedes Benz - MBUX			Porsche - PMC	Tesla		VW	Volvo
		e-tron 2019	Q8 2018	X7 2019	Honda e 2020	IONIQ 2019	Cooper SE 2020	EQC 2019	A-Class 2018	S-Class 2021	Taycan 2019	Model Y 2020	Model 3 2019	Golf V III 2020	s60 2019
1	Explain the user with respect to the capabilites that the AI framework has	MA	MA	MA	MA	MA	NA	MA	MA	SA	MA	SA	SA	MA	MA
2	Identify, educate and mitigate the potential errors	MA	NA	NA	MA	MA	NA	MA	MA	MA	MA	MA	MA	NA	MA
3	Ai should be transparent in what data it has of the user and how it can be accessed by the user	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA
4	Enable the user to provide feedback indicating their preferences during regular interaction with the AI system	SA	MA	SA	MA	SA	MA	SA	SA	SA	SA	MA	MA	MA	MA
5	Maintaining the working memory of recent interactions by learning continuously from user behavior	SA	MA	SA	SA	SA	MA	SA	SA	SA	SA	SA	SA	NA	MA
6	Provide personalized explanations	SA	MA	SA	SA	SA	MA	MA	MA	SA	MA	MA	MA	NA	MA
7	Develop and show significant, value-added and contextually relevant information	MA	MA	SA	MA	SA	NA	MA	MA	MA	SA	MA	MA	MA	MA
8	Update and adapt to different users cautiously	MA	MA	MA	MA	MA	NA	MA	MA	MA	MA	MA	MA	NA	MA
9	Match relevant social norms	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA
10	Mitigate social biases	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA
11	Support efficient AI invocation, dismissal and	MA	MA	SA	MA	MA	MA	SA	SA	SA	MA	MA	MA	NA	MA
12	Scoping precision of service to match uncertainty, variation in goals	SA	MA	SA	MA	MA	MA	SA	SA	SA	SA	SA	SA	NA	MA
13	Be accessible	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA
14	Let the user govern the AI	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA	MA
15	Dont interrupt services and time them based on context of the interaction - Speed is critical	MA	MA	MA	MA	SA	NA	SA	SA	SA	SA	SA	SA	NA	MA
16	Make AI delightful	SA	MA	MA	SA	SA	MA	SA	SA	SA	SA	SA	SA	MA	MA
17	Security and Privacy Matters	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA	SA
SA	Strongly applied	41.2%	11.8%	47.1%	29.4%	47.1%	11.8%	47.1%	47.1%	58.8%	47.1%	41.2%	35.3%	11.8%	11.8%
MA	Moderately applied	58.8%	82.4%	47.1%	70.6%	52.9%	58.8%	52.9%	52.9%	41.2%	52.9%	58.8%	64.7%	47.1%	88.2%
NA	Not applied	0.0%	5.9%	5.9%	0.0%	0.0%	29.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	41.2%	0.0%

Systematic Comparison of Existing HMIs

	AI INTERACTION GUIDELINES	ML in Navigational systems	Continental patent - Tailoring HMI	FCA - Info management system	Ford - Intelligent music systems	GM - Vehicle comfort settings	GM - Sys for Intelligent VI	GM - Proactive infotainment in AV	Symbol technologies
1	Explain the user with respect to the capabilites that the AI framework has	MA	NA	MA	MA	NA	SA	SA	MA
2	Identify, educate and mitigate the potential errors and unexpected edge cases that might occur during user interaction	SA	SA	SA	NA	MA	MA	SA	SA
3	Ai should be transparent in what data it has of the user and how it can be accessed by the user	NA	NA	NA	MA	MA	NA	MA	NA
4	Enable the user to provide feedback indicating their preferences during regular interaction with the AI system	SA	SA	SA	SA	SA	SA	SA	SA
5	Maintaining the working memory of recent interactions by learning continuously from user behavior	SA	SA	SA	SA	SA	SA	SA	SA
6	Provide personalized explanations	SA	SA	SA	SA	SA	SA	SA	SA
7	Develop and show significant, value-added and contextually relevant information	MA	SA	MA	SA	SA	SA	SA	SA
8	Update and adapt to different users cautiously	MA	MA	SA	SA	SA	SA	SA	SA
9	Match relevant social norms	NA	NA	MA	MA	NA	NA	NA	NA
10	Mitigate social biases	MA	MA	NA	MA	NA	NA	NA	SA
11	Support efficient AI invocation, dismissal and correction	MA	SA	NA	SA	SA	SA	SA	SA
12	Scoping precision of service to match uncertainty, variation in goals	SA	SA	SA	SA	SA	MA	SA	MA
13	Be accessible	MA	MA	MA	MA	NA	MA	MA	NA
14	Let the user govern the AI	MA	NA	NA	MA	NA	NA	NA	NA
15	Dont interrupt services and time them based on context of the interaction - Speed is critical	MA	NA	MA	SA	MA	MA	MA	NA
16	Make AI delightful	NA	SA	SA	SA	SA	SA	MA	SA
17	Security and Privacy Matters	SA	MA	SA	SA	SA	SA	SA	SA
SA	Strongly applied	35.3%	47.1%	47.1%	58.8%	52.9%	52.9%	58.8%	58.8%
MA	Moderately applied	47.1%	23.5%	29.4%	35.3%	17.6%	23.5%	23.5%	11.8%
NA	Not applied	17.6%	29.4%	23.5%	5.9%	29.4%	23.5%	17.6%	29.4%

Systematic Comparison of HMI patents

Appendix C

Survey Questions

Opinions regarding AI and interacting with in-car HMI

This survey is being conducted as a part of my master thesis project done in collaboration with Eindhoven University of Technology (TU/e) and Fraunhofer IAO, Stuttgart, Germany. This research is based on analyzing the user experience and acceptance on implementing Artificial Intelligence (AI) on a Human Machine Interface (HMI) of a fully autonomous vehicle.

As this survey is being conducted through prolific, you will be only identified through your prolific ID. The information regarding the types of data collected and confidentiality, can be found in the consent form which appears in the next section.

If you should have any questions about this research study, please feel free to contact the researcher and also the supervisor. The respective email address is provided below:

Main Researcher – Abhishek Suresh, a.suresh2@student.tue.nl

Supervisor – Dr. rer. nat. Bastian Pfleging, b.pfleging@tue.nl

Section 1

Informed Consent

The following information is provided to inform you about the research project and your participation in it. Please read the form carefully and feel free to ask any questions you may have about this study. The information regarding whom to contact, is provided at the end of this consent form. You are not obligated to participate in this research or to answer the questions, and you are also free to withdraw from the survey at any time. In the event of any new information becomes available that may affect your participation in the study, you will be notified so that you can make an informed decision of whether to (not) participate in this study.

1. Purpose of the survey

In this survey, we would like to gain some insights about the user's opinion regarding the implementation of AI within the car, especially for facilitating the interaction with the vehicle's HMI (referring to vehicle's Central Infotainment System). Subsequently, we are looking for your opinion on using the vehicle's navigational system, and also using the central display for customizing the driving and system settings. These insights will be suitable for evaluating how the AI can be implemented in facilitating the interaction between you and the vehicle's HMI. There are some questions which you may not be familiar with. In that case, just answer it by assuming yourself in the question's context.

2. Description of data to be collected and approximate duration of the study

As a part of this research, we will be recording the following data in the survey. No personal data such as your name, email address and phone numbers will be collected. This survey is expected to take around 20 minutes of your time, and can be completed using desktops and smart phones.

- Your Prolific ID, age, gender, country of residence, level of education and income level
- Certain demographic data regarding driving and commuting
- Some general opinions about technology and artificial intelligence (AI)
- Your opinions regarding the use cases formulated for this research

3. Description of discomforts, risks and/or inconveniences that can be reasonable expected as a result of participating in this study

There are no foreseeable risks, discomforts, or inconveniences associated with participating in this survey. Questionnaire data includes only low risk information with results being only presented in aggregated form. Participants will not be exploited, and none of their data would place participants at risk of criminal or civil liability.

4. Data confidentiality

All efforts, within reason, will be made to keep the participant's personal information private, without compromising the privacy. The collected data will be coded by assigning a randomized number to each data set provided by the participants. The coded data will be kept on a password protected institutional repository at (or recommended by) the Eindhoven University of Technology. All the personal data collected during the study will be processed confidentially and test subjects will never be recognizable in publications, academic material or any other mean unless explicit consent (e.g. for pictures) is given. As per the Netherlands code of conduct for research integrity, the data will be stored for a maximum period of 10 years. Participants can ask the researchers for an electronic copy of the data that he/she has provided or that have been measured directly. If they are dissatisfied with how data privacy is handled, they can submit a complaint to the Chief Information & Security Officer, the Privacy & Security Officer and/or the Data Protection Officer of the Eindhoven University of Technology via privacy@tue.nl or contact the Dutch Data Protection Authority.

5. Withdrawing consent

You can withdraw your consent to the use of your data at any time. This applies to this study and also to storage and use for future research. The study data collected until the moment you withdraw your consent will still be used in the study.

1. I have read this informed consent documents and the material contained in it has been explained. I understand that in each part of the document, my questions have been answered, and I freely and voluntarily choose to participate in this survey.

- Yes, I consent
- No, I do not consent

Section 2

Participant's Demographics

2. Enter your age

Enter your answer

3. What gender do you identify as?

- Woman
- Man
- Prefer not to say

4. Specify your current country of residence

Enter your answer

5. What is your highest degree / level of education?

- High School

- Bachelor
- Master
- PhD
- Post Doc

6. What is your current employment status?

- Employed full-time
- Employed part-time
- Seeking opportunities
- Student
- Retired from work
- Prefer not to say

7. What is your annual household income?

- Less than $25,000
- $25,000 - $50,000
- $50,000 - $100,000
- $100,000 - $200,000
- More than $200,000

Section 3

Driving related questions

8. How often do you drive?

- Every day
- Few times a week
- Once a week
- Few times a month
- Once a month
- Less than once a month
- Never

9. How long you have been driving for? Enter the answer in years (Enter 0 if you have never driven)

Enter your answer

10. If you don't drive, then how often you travel as a passenger?

- Always
- Usually
- Sometimes
- Rarely
- Never

11. Do you have a valid driver's license or a learner's permit (if you are currently learning to drive)?

- Yes
- No

12. What mode of transportation do you use to commute daily?

Enter your answer

13. What are the non-driving related activities that you would like to do while you are in the vehicle? Name some instances briefly.

Enter your answer

Section 4

Opinion regarding AI and it's implementation

In this block, we would like to know your opinion in general regarding artificial intelligence, it’s implementation and its impact in the future

14. Kindly share your existing technical knowledge with self-driving cars in general. Rate your answer on a scale of 1 to 5, with 1 being “no technical background” and 5 being “a strong technical background”

1 2 3 4 5

15. Kindly share your existing technical knowledge regarding artificial intelligence in general. Rate your answer on a scale of 1 to 5, with 1 being “no technical background” and 5 being “a strong technical background”

1 2 3 4 5

16. Did you take any classes or some courses regarding AI?

- Yes
- No

17. Name an example of Artificial Intelligence.

Enter your answer

18. When was the last time you heard about AI in the media? (Media being TV, newspaper, radio/podcasts, internet, magazine)

- Today
- This Week
- This Month
- This Year
- More than a Year
- Never

19. Do you interact with AI-infused systems often? (Some of the AI-infused systems being Smart appliances, voice assistants in smartphones, in-car assistant systems, computer software, robotic products.)

- Yes
- No
- Sometimes

20. Bringing cars into context, what kind of in-car assistant systems you prefer interacting with?

- Manufacturer's Own
- Android Auto

- Apple Carplay
- Other

21. What kind of data that you are willing to share it with the in-car assistant systems, expecting it to deliver a personalized experience to you?

Enter your answer

22. Do you think AI can replicate human intelligence in machines?

Rate your opinion in the scale below ranging from highly unlikely to highly likely.

	Highly Unlikely	Unlikely	Neutral	Likely	Highly Likely
Answer					

23. Do you believe that AI-based systems will be prominent in the future by replacing humans? For example conventional cars are replaced by a fleet of self-driving cars with no drivers.

Rate your opinion in the scale below ranging from highly unlikely to highly likely.

	Highly Unlikely	Unlikely	Neutral	Likely	Highly Likely
Answer					

24. What do you think that AI might bring in the future?

Enter your answer

25. Does the promise and the increase in application of AI scare you?

- Yes
- No

Section 5

Questions regarding interacting with the in-car Navigational System

This block is used to gather opinions regarding the common interaction with the navigation system of the car. One of the potential use cases that we are going to investigate in this research is how the infusion of AI in the navigation system facilitates interaction between the system and the occupants of the vehicle. In this case, we would like to know some general types of interactions that you do while using the navigational system, followed by a specific use case of destination entry in the navigation system.

26. Click "I understand" to proceed further in this section

- I understand

27. What kind of navigation systems and/or devices that you use or prefer to use often while driving?

- On-board Navigation System
- Maps on Smartphones
- Maps from third party applications (like Android Auto, Apple Carplay))
- External GPS Navigation Device
- Other

28. Could elaborate briefly regarding why did you chose the option from the question above?

Enter your answer

29. Rank the order in which kind of input modality you prefer to interact with the navigational system.

Touch input

Voice input

Gesture input

30. Explain in brief the reason behind your ranking of the input modalities mentioned in question 29.

Enter your answer

31. If you prefer the touch input, then do you prefer to type using keyboard input or use the touch pad (if the car has) for entering the data on the navigation system?

- Keyboard Input
- Touchpad
- Other

32. Let’s assume that you do not have or want to use your personal navigation devices at this time, and you rely on using the on-board navigation system in the car.

Where would you prefer the maps to be displayed?

- Only on the Instrument Cluster
- Only on the Central Display
- Both

33. Explain briefly on why you chose the particular option in the previous question.

Enter your answer

34. Now, how big do you think the screen size shall be so that you can interact with it effectively. You can also consider the possibility of your car not having an instrument cluster (like in Tesla model 3).

Please enter your desired number (in inches) below.

Enter your answer

Section 6

Destination Entry

In this section, let us assume that you are going to plan the route using the car’s navigation system. By step-by-step destination entry, you can set one destination point or multiple points for the long trip, entering different patterns, and also you can add different points of interest and landmarks. For this, we would like to know in general how would you approach and perform the destination entry.

35. There are potentially 6 possible ways for entering a destination in the navigational system. Please rank in the order of which would you prefer the most to the least.

By the Point of Interest (POI)

By a ZIP code

By using voice assistant

By the previously entered destination points

By selecting the preferred location by tapping the map

By the address of the street (street number/street name/city/town/village, state, etc.)

36. From the ranking in the previous question, kindly suggest the reason behind the selection of most and least preferred way of entering the destination.

Please try to explain it in 1 to 2 lines.

Enter your answer

37. Do you prefer to use the destination entry feature while the vehicle is stationary or during driving?

- When the vehicle is stationary
- During driving

38. Now, let us consider entering an address using the entry feature of the car's navigational system. The address of your potential destination is address of TU Eindhoven, which is given below.

    De Zaale, 5612AZ
    Eindhoven, The Netherlands

    The first line represents the street and postal code, followed by the city and country name on the next line. While entering the above address, which one do you prefer doing?

    - Entering the city first, followed by street name and postal code
    - Entering the street name and postal code, followed by the city name

39. For selecting a particular city from the map, which one of the following you prefer the most?

    - Typing the name of the city in the search bar
    - Pinning the city location by directly tapping on the map

40. Kindly elaborate the reason behind why you chose that particular option in the previous question.

    Enter your answer

41. Name 2 potential point of interest (POI) that you might add as an interim destination after setting the destination.

    Enter your answer

42. How often do you add point of interests in the navigational system while driving?

Rate your opinion in the scale below with 1 being "very rarely" to 5 being "too often"

1 2 3 4 5

Section 7

Questions about the Customization aspect of the HMI

In this section, we wanted to know your opinions regarding customizing the instrument cluster and the central infotainment screen. Similar to the navigation system, we are aiming to improve the customization aspect of the display which will be assisted by AI. One thing to note that, we are focusing only on the customization aspect of the display, not the personalization aspect.

43. The main aim while considering this use case is not the system providing you content, but how the system might assist you in reshaping the content according to your needs. More information regarding Personalization and Customization can be found in this link (https://www.nngroup.com/articles/customization-personalization/).

Click I understand to proceed further in this section.

- I understand

44. Before answering the next question, I would like to quickly elaborate to you, the 3 types of design styles that everyone prefer during customization. If your car does not provide you the functionality to customize the display, then you can compare this assumption as similar to customizing your mobile phone or laptop displays. Please read the 3 categories carefully.

Individualization - All the features that you tend to use more often are presented on the screen
every time you start the car

Style - Offering a wide variety of styles to suit the mood and condition of the driving

Minimalism - Simple in design, and only essential features that are being used more often
is provided directly. Some of the less and unwanted features are either
removed or hidden.

From the above 3 design features, what type of customization do you prefer the most?

- Individualization
- Style
- Minimalism

45. Kindly explain the reason behind your choice in brief.

Enter your answer

46. How often do you customize the instrument cluster and head unit of the display?

Rate your opinion in the scale below with 1 being "very rarely" to 5 being "too often"

1 2 3 4 5

47. Do you prefer the stock design style of the instrument cluster and the head unit or would like to change the style of these displays often?

- I would prefer the stock design
- I would prefer changing the design style

48. List 3 features (in no order ----- and excluding the features from the navigation system) that you would be very much likely to be seen using it often, and kindly elaborate on why did you chose those features.

Enter your answer

49. List 1 instance of the feature that you miss the most or the feature you might think you need it the most when driving.

Explain in 1 to 2 lines why did you chose that feature.

Enter your answer

50. Have you ever used every single possible feature provided to you in the central infotainment display?

- Yes
- No
- Not so sure

Section 8

Some extra questions regarding the driving and system settings.

In this section, we will ask you some questions that help us to gain further insights regarding changing of driving settings and the system settings using the main central infotainment display. Please read and answer the following questions.

51. What type of driving settings you prefer to set often and why? Example including sport mode, comfort mode and many more.

    Explain your reasoning in 1 to 2 lines.

    Enter your answer

52. How often do you change the driving settings?

    Rate your opinion in the scale below with 1 being "very rarely" to 5 being "too often"

    1 2 3 4 5

53. How often do you change the system settings?

    Rate your opinion in the scale below with 1 being "very rarely" to 5 being "too often"

    1 2 3 4 5

54. Do you know how to change or moderate the entire system settings?

    - Yes
    - No

55. Would you rather prefer the system only shows the settings that you change often or you want all of the settings to displayed?

    Please share your opinion in 1 to 2 lines

Enter your answer

Section 9

THANK YOU VERY MUCH FOR CONTRIBUTING TO THIS SURVEY

I know there are a lot of questions in this survey, and it's a taken a long time to finish it, but I am really grateful for taking your time to contribute to my research. Thank you for all your time and effort.

Best Regards

Abhishek Suresh

Appendix D

Expert's Responses to Heuristic checklist

Heuristic 1 - Visibility of the System Status

Table D.1: Participants responses to the checklist for H1

Q.No	Checklist	Yes	No
1.	Is there a consistent icon design scheme and stylistic treatment across the system?	6	1
2.	Is a single, selected icon clearly visible when surrounded by unselected icons?	4	3
3.	Do menu instructions, prompts, and error messages appear in the same place(s) on each menu?	5	2
4.	Is there some form of system feedback for every operator action?	5	2
5.	After the user completes an action, does the feedback indicate that the next group of actions can be started?	3	4
6.	Is there a feedback in menus or dialog boxes about which choices are selectable?	4	3
7.	Are response times appropriate to the task?	2	5
8.	Are response times appropriate to the user's cognitive processing?	2	5
9.	Does the system provide visibility (i.e.) by looking, can the user tell the state of the system and alternatives for action?	3	4
10.	Does the GUI menus makes obvious which items can be selected/deselected?	4	3

Heuristic 2 - Match Between the system and the real world

Table D.2: Participants responses to the checklist for H2

Q.No	Checklist	Yes	No
1.	Are icons concrete and familiar?	6	1
2.	Are the menu choices ordered in the most logical way, given the user, the item names and the task variables?	4	3
3.	If there is a natural sequence to menu choices, has it been used?	5	2
4.	If the shape is the virtual cue, does it match the cultural conventions?	6	1
5.	Do the selected colors correspond to the common expectations about color codes?	6	1
6.	When prompts imply a necessary action, are the words in the message consistent with the actions?	6	1
7.	Do menu choices fit logically into categories that have readily understood meanings?	6	1
8.	Has the interface been designed so that the keys with similar names do not perform opposite (also potentially dangerous) actions?	6	1

Heuristic 3 - User control and freedom

Table D.3: Participants responses to the checklist for H3

Q.No	Checklist	Yes	No
1.	If setting up the interface is a low frequency task, is it fairly easy to remember?	5	2
2.	When a user's task is complete, does the system wait for a signal from the user before processing?	3	4
3.	Are users prompted to confirm commands that have drastic, destructive consequences?	3	4
4.	Is there any error correction option at any stages of the interface?	3	4
5.	Can users cancel out the operation that are currently in-progress?	3	4

6.	Do the users have an option of initiating the sequence through multiple input modalities (i.e.) by using touch/gesture/voice input?	6	1
7.	Can users effectively navigate between various fields or dialog box options?	5	2
8.	Can user's effectively and easily reverse their actions, and is there any retracting input to allow for multiple redo's?	1	6
9.	Can users set their own system, session, file, and screen defaults?	3	4

Heuristic 4 - Consistency and Standards

Table D.4: Participants responses to the checklist for H4

Q.No	Checklist	Yes	No
1.	Have industry or company standards, that are publicly available, applied consistently on all menu screens in the system?	5	2
2.	Is the most important information placed at the beginning of the prompt?	6	1
3.	Are the user actions named consistently across all prompts in the system?	6	1
4.	Are the menu choice names consistent, both within each menu, and across the system, in grammatical style and terminology?	7	0

Heuristic 5 - Error Prevention

Table D.5: Participants responses to the checklist for H5

Q.No	Checklist	Yes	No
1.	Does the system prevent the users from making errors whenever possible?	4	3
2.	Does the system warn the users if they are about to make a potentially serious error?	4	3
3.	Does the system intelligently interpret variations in user commands?	6	1

Heuristic 6 - Recognition rather than recall

Table D.6: Participants responses to the checklist for H6

Q.No	Checklist	Yes	No
1.	For questions and answer interfaces, are visual cues and white space used to distinguish questions, prompts, instructions, and user input?	7	0
2.	Are prompts, cues and messages placed where the eye is likely to be looking on the screen?	6	1
3.	Have prompts been formatted using white space, justification, and visual cues for easy scanning?	5	2
4.	Is there an obvious visual distinction made between "choose one" menu and "choose many" menus?	4	3
5.	Have items been grouped into logical zones, and have headings been used to distinguish between zones?	6	1
6.	Have zones been separated by spaces, lines, color, letters, rules, lines, or shaded areas?	6	1
7.	Does the system provide "mapping" (i.e.) are the relationships between controls and actions apparent to the user?	7	0
8.	Does the user interface make obvious when selection is possible?	3	4

Heuristic 7 - Flexibility and efficiency of use

Table D.7: Participants responses to the checklist for H7

Q.No	Checklist	Yes	No
1.	If the system supports both novice and expert users, are multiple levels of error message detail available?	2	5
2.	Does the system allow novice users to enter the simplest, most common form of each command, and allow expert user's to add parameters?	1	6
3.	Can expert users bypass nested dialog boxes with either type-ahead, user-defined macros, or keyboard shortcuts?	1	6

Heuristic 8 - Aesthetic and Minimalist design

Table D.8: Participants responses to the checklist for H8

Q.No	Checklist	Yes	No
1.	If the system uses a standard GUI interface where menu sequence has already been specified, do menus adhere to the specification whenever possible?	7	0
2.	Are meaningful groups of items separated by space?	6	1
3.	Are menu titles brief, yet long enough to communicate?	6	1

Heuristic 9 – Help Users Recognize, Diagnose, And Recover From Errors

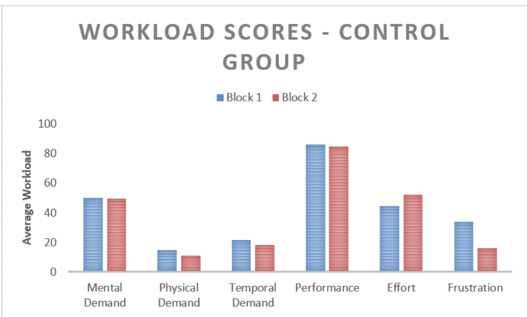
Table D.9: Participants responses to the checklist for H9

Q.No	Checklist	Yes	No
1.	Is there any type cues to indicate an error?	3	4
2.	Are prompts stated constructively, without over or implied criticism of the user?	5	2
3.	Do prompts imply that the user is in control?	6	1
4.	Are error messages worded so that the system, not the user, takes the blame?	6	1
5.	Do error messages avoid the usage of violent or hostile words?	5	2
6.	Do messages place users in control of the system?	5	2
7.	Do error messages inform the user of the error's severity?	4	3
8.	Do error messages suggest the cause of the problem?	5	2
9.	Do error messages indicate what action the user needs to take to correct the error?	3	4

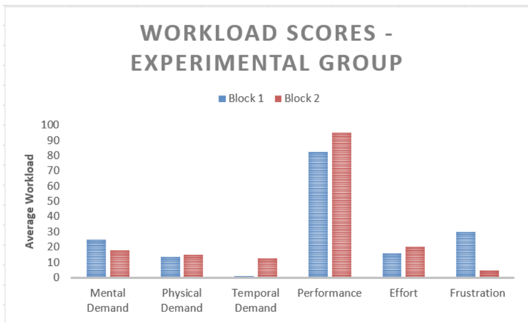
Appendix E

Pre-Study Results

E.1 Workload Results



(a) Workload Values - Control Group



(b) Workload Values - Experimental Group

Figure E.1: Workload Values for Individual NASA TLX Constructs in Pre-study

E.2 CTAM Average Scores

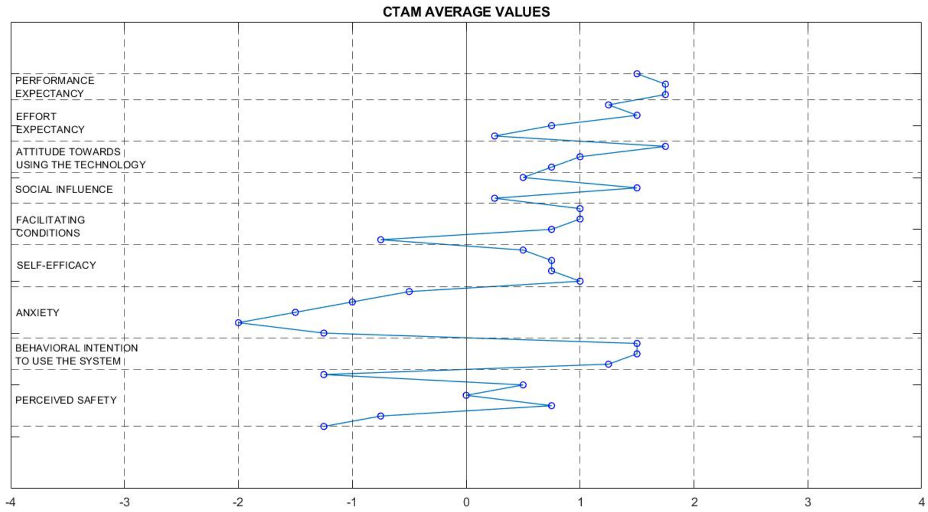


Figure E.2: Average values plot for the CTAM constructs in the Pre-study

E.3 Attrakdiff Scores

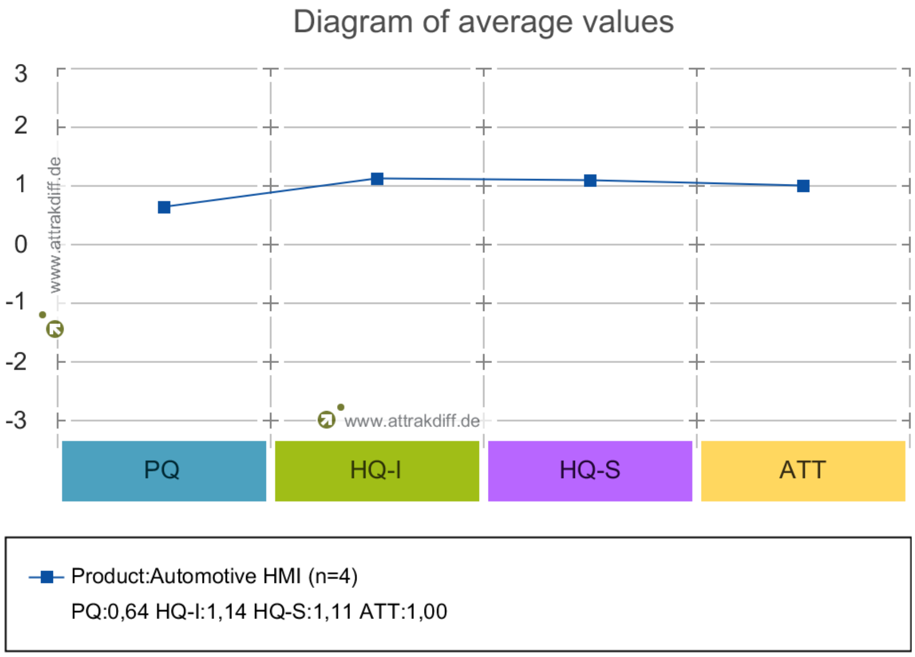


Figure E.3: Averages Values Plot for the Attrakdiff Qualities

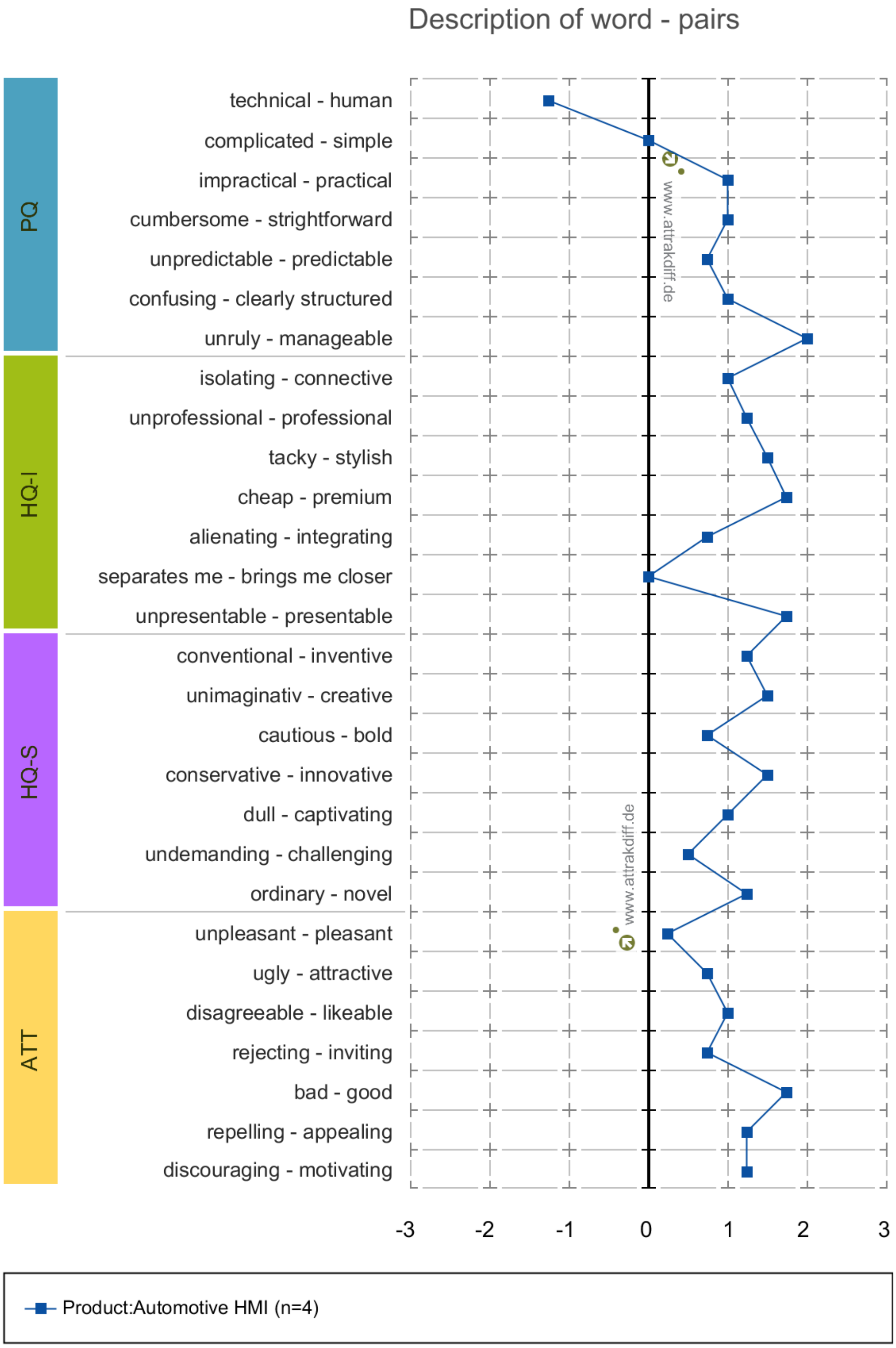


Figure E.4: Attrakdiff Pre-study Scores : Semantic Differential

Appendix F

Usability Study Results

F.1 Mental model responses

In addition to the 5 questions mentioned in the table 7.3, the participants are also asked to provide a situation different than the use cases where the system can customize the settings and features for them, and the responses provided by the participants for this question is given below.

1. Music recommendations, different layout based on profiles
2. Route settings other than the city scenario which for example in mountains, plains, etc.
3. Removing the route settings that are not suitable for the route; for example, there is no purpose of keeping off-road setting when driving in the city
4. General recommendations and personalized services
5. Route selection based on the occupants mood and also the type of travel
6. Mood based routes
7. Making the system to have a conversation proactively with the occupants when/during travel
8. Tap=or-say situation
9. Specific music and playlists for long distance travel
10. If the destination address is new to the user, then the car can provide address with the country name first, followed by the city and street name in that order
11. Detects the temperature outside and adjusts the climate settings accordingly
12. For the point of interest list on the map, list down the complete name of the place and distance from the charging station
13. In addition to the route mode displayed, emergency routes can also be provided as an option
14. More details need to be provided in the context of battery and energy management

F.2 Semantic Differential Graph for the Usability Study Word Pairs

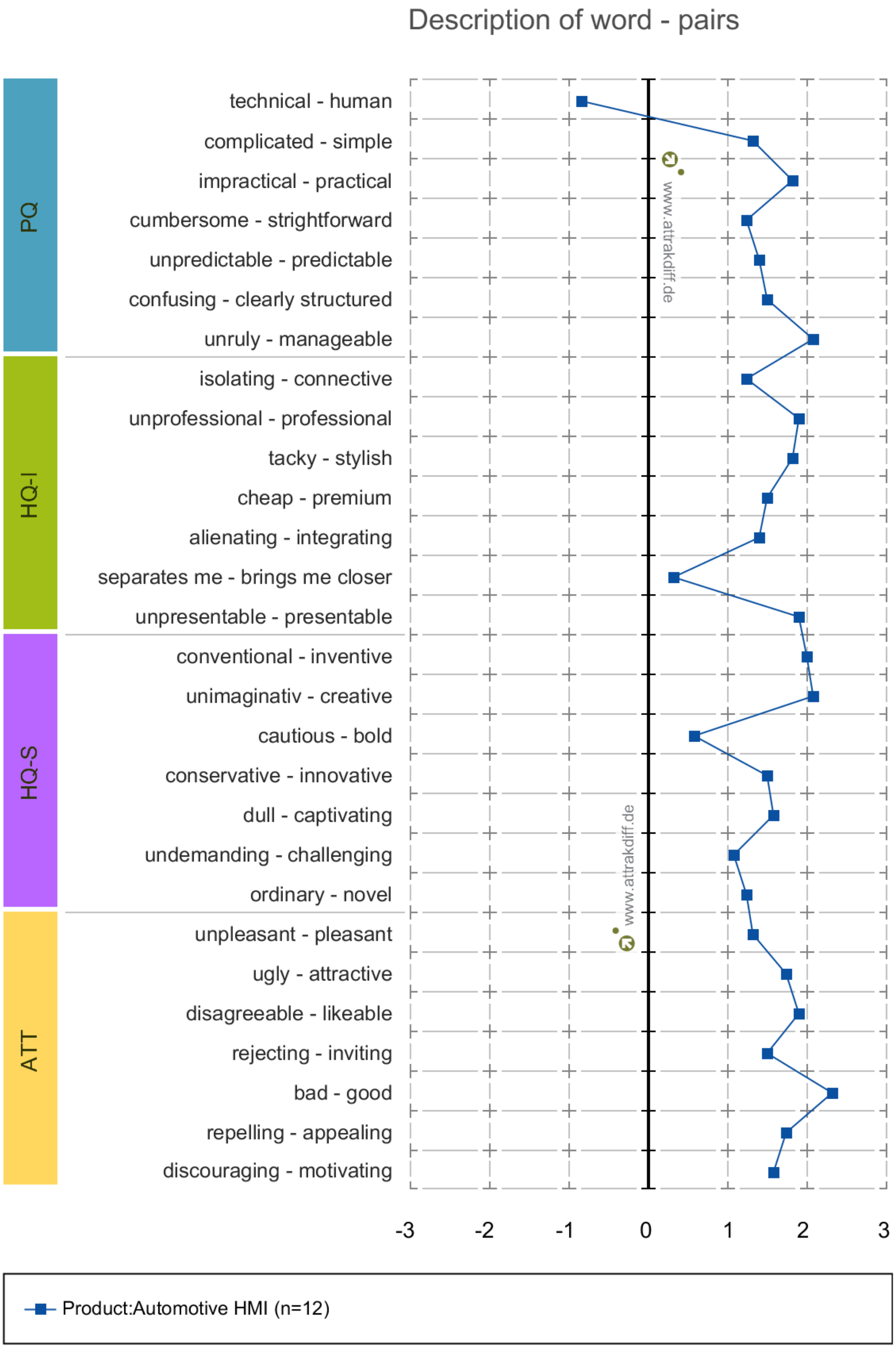


Figure F.1: Attrakdiff Usability study Scores : Semantic Differential