Designing and Evaluating a Head-Up Display with Focus on High Driver Performance for Non-Driving Related Features

Masters Thesis

Jeppe Skovby Bjørn,
Ronnie Svendsen

AALBORG UNIVERSITY
DENMARK

Aalborg University
Department of Computer Science
Selma Lagerløfs vej 300
DK-9220 Aalborg

AALBORG UNIVERSITY
STUDENT REPORT

Department of Computer Science
Cassiopeia
Selma Lagerløfs Vej 300
9220 Aalborg Øst
https://www.cs.aau.dk/

Title:
Designing and Evaluating a Head-Up Display with Focus on High Driver Performance for Non-Driving Related Features

Theme:
HCI - Master Thesis

Project Period:
Spring, 2022

Project Group:
cs-21-hci-10-21

Participant(s):
Jeppe Skovby Bjørn
jbjarn17@student.aau.dk
Ronnie Svendsen
rsvend17@student.aau.dk

Supervisor(s):
Jan Stage
jans@cs.aau.dk

Copies: 1
Page Numbers: 49
Date of Completion: 15-06-2022

Summary

During the preliminary research it was found that one-tenth of drivers are texting while driving in Denmark, even though it has been illegal since 1992. Therefore, it was chosen to look at possible alternatives to texting on a mobile phone while driving. A **H**ead-**U**p **D**isplay was found to have benefits to texting while driving. However, no research that tested a HUD with voice-to-text could be found. Therefore, the objective of this thesis was to explore the voice-to-text-based interaction with a HUD and compare it to using a phone while driving. This led to our main research question: *How can a HUD be designed and evaluated with focus on high driver performance, to support non-driving related features?*. To answer the main research question, we created two sub-questions: *What techniques are used in research literature to design and evaluate automotive HMI's?* and *What impact would a HUD-based text message system with voice-to-text as input, have on driver performance compared to a mobile phone?*. The first research question was chosen to be created since we had no prior experience or knowledge of how to develop an automotive HMI. To answer the research questions, two papers were written.

The first paper explored the techniques used for developing and evaluating automotive **H**uman-**M**achine **I**nterfaces(HMI). Through preliminary research of other papers, it was determined to use a **S**ystematic **L**iterature **R**eview(SLR) to guide the data collection. The search strategy was pilot tested and refined before studies were collected for the SLR. In total 907 studies were fetched and exclusion criteria were applied to determine which fit into the context of the paper. The final poll of studies was 87 which were published between 2005-2021 and came from 4 different databases. The results were quantified in a table before further analysis was conducted.

The most popular technique for developing and evaluating automotive HMI was observation followed by questionnaire, applied research, and interview. Observation and questionnaire are both used mainly to gather quantitative data. Often a combination of different techniques is used to gather different data and data types depending on the purpose of the usage.

In the second paper, we compared a HUD with voice-to-text-based communication with a mobile phone, to see if there was any difference in their respective impacts on driving performance. First, we designed a HUD and had eight participants evaluate it. The feedback from the participants lead to some design changes. To test the HUD and the phone while driving, we constructed a driving simulator. It consisted of a 50-inc monitor, a driver's set, G29 Logitech setting wheel, and pedals. The software used for the simulator was Carla Simulator[1] which is mainly written in C++ and Python. We invited 16 participants to use our simulator and perform tasks using both the HUD and the phone. The driving environment had no traffic congestion except for a leading vehicle the participants had to follow. The participants did two different tasks with both the HUD and the Phone. The two different tasks were to read a message and read and respond to a message. Each type of task was done three times with both the phone and the HUD, leading to a total of 12 tasks.

The results of the experiment showed that there was a difference in the driver performance when using a HUD for texting compared to a phone. The participants had significantly lesser lane crossings, were 3.5 times less likely to crash, saw a significant decrease in subjective workload, and could write longer messages with the HUD with the same completion time to a phone.

Contents

Summary ii

Master Thesis Details v

Preface vi

I Summary of Thesis 1

1 Introduction 2

1.1 Research Questions 3

2 Article Overview 4

2.1 Article One 4

2.2 Article Two 5

3 Research Method 6

3.1 Article One 6

3.2 Article Two 7

4 Limitations 9

4.1 Article One 9

4.2 Article Two 9

5 Conclusion & Future Work 10

References 11

II Papers **13**

Developing HMIs: A Literature Review of Automotive Development Techniques **1**

References 7

The Effects of Texting While Driving: Voice-To-Text Interaction with a HUD **1**

References 9

Master Thesis Details

Thesis Title: Designing and Evaluating a Head-Up Display with Focus on High Driver Performance for Non-Driving Related Features
Masters Students: Jeppe Skovby Bjørn, Ronnie Svendsen
Supervisor: Assoc. Prof. Jan Stage, Aalborg University

The main body of this thesis consists of a summary of the thesis and the two papers listed below. As seen on the list, the articles will be referred to as *Article One* and *Article Two* in the first part of the thesis.

[Article One] Jeppe Skovby Bjørn, Ronnie Svendsen, "Developing HMIs: A Literature Review of Automotive Development Techniques", 2022.
[Article Two] Jeppe Skovby Bjørn, Ronnie Svendsen, "The Effects of Texting While Driving: Voice-To-Text Interaction with a HUD", 2022.

These papers can be found in the appendix of the thesis.

Preface

This master thesis in software and appendixes have been written by two Software students in the tenth semester, at Aalborg University, in the period 1st February 2022 to 15th June 2022.

The report uses the Vancouver style for referencing, with the alteration that if a citation is before a period, the source belongs to that sentence, while if a citation is after a period, it indicates that the source is to the entire paragraph.

The master thesis group consists of: Jeppe Skovby Bjørn, and Ronnie Svendsen.

The master thesis follows the "Studeordningen for kandidatuddannelsen i software 2020" (Curriculum for the Master's program in Software 2020).

This master thesis was created to explore the text-based communication possibilities in a head-up display and how this could improve driver performance compared to a standard mobile phone. A systematic literature review was conducted to establish knowledge about the different techniques used to develop and evaluate automotive Human-machine interfaces.

In the 9th-semester pre-specialization project [2], a first draft of the systematic literature review was reported. The Systematic literature review in this paper is the second and final version.

A HUD for text-based communication was developed and evaluated based on the acquired knowledge.

We would like to thank our supervisor Jan Stage, for his guidance through the master thesis.

Jeppe Skovby Bjørn, Ronnie Svendsen

Aalborg University, June 14, 2022

Part I

Summary of Thesis

Chapter 1

Introduction

Using a mobile phone while driving harms the driver's performance by increasing the driver's reaction time, loss of correct position in a lane, and the speed of a vehicle decreases. [3]

As a result of this, rules have been created to reduce the impact mobile phones have on driving performance. One of these laws in Denmark is that one may not use a hand-held mobile phone while driving[4] . This law came into force in 1992[5].

However, 14% of the danish drivers think it is okay to do non-driving-related activities while driving, including but not limited to talking on the phone, reading text messages, and adjusting the GPS[6]. Drivers aged 18-35 are the most represented group as 27% think it is acceptable to do things that temporarily distract them from driving, such as the things mentioned above. Furthermore, every tenth driver in 2018 has written a text message on the phone while driving. [7]

In the USA, 24 states also restrict the usage of hand-held mobile phones as a primary law[8]. However, the number of crashes where the driver uses a mobile phone is stable in the USA with almost no reduction in fatal crashes from 2016 to 2019. Furthermore, people aged 20-29 are responsible for more than 1/3 of all mobile phone-related accidents. [9]–[11]

Even though it is illegal by law in most states in the USA and all of Denmark, it is evident that mobile phones hand-held is still used despite the drawbacks. One reason for this is that applications are designed to release dopamine that makes it feel rewarding every time one checks/answers a text message or notification[6].

Increasing the fine for using a hand-held mobile phone could possibly decrease the usage. However, this has not proven effective, and the total number of offenses is the same[12]. Another approach could be to make it safer to use mobile phones and find a way that makes it less distracting to perform tasks on the phone, such as calling, texting, or other tertiary driving tasks.

One possible solution could be to use the built-in voice assistant in mobile phones. Research on this topic showed that manual input was faster than using voice-to-text. However, the participants felt safer using the voice-to-text form, contrary to the manual. This research was conducted with a hand-held mobile phone, which is illegal by law in Denmark and almost half of the states in the USA. [13]

A study using voice assistants for performing tasks while driving has shown that the general driving performance is better, the workload is less, and the driver needs less time looking away from the road when using speech versus manual interfaces. Furthermore, interacting with a digital assistant also improved some driving performance parameters, such as lane-keeping and a better reaction time to hazardous situations[14]. Dependent on the type of task, the driver's workload, and the demographics of the driver, a manual interface can be better in some situations. An example of this could be in a task that required recognizer accuracy. [15]

Another possible solution to the problem could be to utilize a **H**ead-**U**p **D**isplay (HUD), with a 3x4 keypad mounted on the steering wheel, which is presented in [16]. The results showed improvements in different driver performance variables, such as lane deviation and glance period off-road compared to a QWERTY display. The study also commented that they had studied voice-to-text as a possibility. However, this was found to be inefficient, unreliable, and typically required visual attention to ensure that the message was as intended.

Since the publication of [16], voice assistants have grown exponentially. Google released Google Assistant, which by user experience, is reliable compared to other systems like Siri. [17], [18]

This gives rise to the question of the usage of advanced voice assistant systems, such as Google Assistant, Siri, etc. can be seen as a viable solution to the issue and help alleviate the apparent problem of people doing tertiary tasks, such as writing messages with a hand-held mobile phone, and make it a safer option.

This leads to the overall research question of the master thesis:

> **Overall Research Question:**
> *How can a HUD be designed and evaluated with focus on high driver performance, to support non-driving related features?*

1.1 Research Questions

This section will consist of the **R**esearch **Q**uestions (RQ), to be answered in this master thesis. RQ1 was chosen to be created, since we had no prior experience or knowledge of how to develop automotive HMI. Preliminary research revealed no conclusive results on how to develop automotive HMI and the techniques used varied.

Overall RQ: *How can a HUD be designed and evaluated with focus on high driver performance, to support non-driving related features?*

RQ1: *What techniques are used in research literature to design and evaluate automotive HMI's?*

RQ2: *What impact would a HUD-based text message system with voice-to-text as input, have on driver performance compared to a mobile phone?*

Chapter 2

Article Overview

This chapter brings an overview of the two articles written during this thesis.

2.1 Article One

This article was created to answer the first research question(RQ1): *What techniques are used in research literature to design and evaluate automotive HMI's?* To answer this question, a **S**tructured **L**iterature Review(SLR) was used to guide the review of research material.

Two SLRs were studied to get a better understanding of the setup and execution of an SLR. This provided the general structure of how the overall results should be represented in a table. Databases to gather literature were also identified via these SLRs.

The SLR consisted of initial pilot tests to refine the search query used. The search acquired 907 potential studies from four databases. In order to reduce the number of candidate studies and ensure relevancy, a selection process with exclusion criteria was utilized. The exclusion criteria consisted of five different criteria. The exclusion criteria were: published before 2005, duplicate article, not written in English, not peer-reviewed, and it does not design or evaluate automotive HMI. Of the 907 candidate studies, 87 of those became primary studies through the selection process based on the exclusion criteria.

In order to analyze the material and answer the research question, five types of information were extracted from each study. This included the year in which the article was published, which system is being designed/evaluated, which techniques are used to gather data, how users are involved, and in which part of the development users are involved. This provided a foundation from which additional in-depth investigations of the organized studies could be done.

The results showed that the spans of 2005-2009, 2010-2014, and 2015-2021 saw an increase in the number of publications in each span. Most of the techniques used for development were questionnaire and observation, which covered 73%. The most common way the techniques were used was in combination with each other and the most common combination was questionnaire and observation. Qualitative data was the most collected data type. The data collected was used to get an understanding of how the users used the automotive HMI. These data were then statistically analyzed. The qualitative data collected was used to get an understanding of why the users used the automotive HMI.

2.2 Article Two

The purpose of this article, was to answer the second research question(RQ2): *What impact would a HUD-based text message system with voice-to-text as input, have on driver performance?*

To answer the research question, we constructed a car simulator which had one 50-inch monitor, a driver's seat, and a G29 Logitech steering wheel and pedals. The simulator environment was developed using Carla Simulator[1], which needed minor changes. Furthermore, we also developed mock-ups of HUDs which were evaluated by eight participants. The participants' feedback lead to some changes in the HUD, and a prototype in the simulator was implemented. We invited 16 participants and asked them to perform a series of voice-to-text texting tasks with the HUD. To compare texting with the HUD, the participants also had to do the texting tasks with a phone. The participants had to follow a leading vehicle while doing their tasks. The leading vehicle followed the same path for every participant, to ensure consistency in what they went through.

The driving environment was designed so that the participants either drove in a city or freeway scenario. The freeway had high-speed limits whereas, the city scenario had traffic lights and slower speed limits.

The participants were asked to do two different tasks, which were to read a message and respond to a message. They had to do each task three times with the HUD and the phone, totaling 12 tasks. Every participant got the same messages.

Before the participants did the experiment with the HUD, they had to fill out a questionnaire for demographic data. At the end of the experiment, we did a 5-minute semi-structured interview with the participants.

In the experiment during the tasks, we collected data such as lane crossing, speed difference to the leading vehicle, eye glances of HUD, subjective workload, crashes, average words used pr. response, and time completion. The data was logged automatically through the simulator. The simulator screen was recorded using the software OBS.

The participants had a tendency when doing the tasks with the phone, that they looked at it, even though the signal to a task was not given. Only four of the participants in the experiment felt that the driving simulator was not realistic. Whereas 10 participants felt that it was realistic but had shortcomings, and two felt it was realistic.

The results showed that when the participants used the HUD over the phone, it had an impact on driver performance. The participants had fewer lance crossings, fewer crashes, and could write longer messages with the HUD. The subjective workload of each participant also saw a significant reduction. Results also showed that using the HUD for texting one would be 3.5 times less likely to be in a crash.

Chapter 3

Research Method

This chapter will encapsulate the research methods used in the articles for this thesis.

3.1 Article One

The research conducted in article one is an SLR, based on [19]. An SLR was chosen as the research method for the article, to identify, evaluate and interpret all relevant and available research for the topic and find possible gaps. Studies that become part of the SLR are called primary studies.

The reason for undertaking the SLR was the summarize the existing technology used in automotive research. Undertaking an SLR means having a well-defined methodology to reduce the possible publication bias of the studies. However, if bias is introduced in an SLR it can lead to misleading conclusions [20]. To reduce bias in the article, the methodology needed to be well documented and choices were to be explained with argumentation. By having a well-defined methodology, also makes the results of the SLR replicable.

An advantage to an SLR is that the studies can be quantified and making it a quantitative study [19]. In this SLR the studies found to be primary studies were quantified into a table based on the development techniques used in the study and then later analyzed.

There are some disadvantages when conducting an SLR according to [21] which are but not limited to:

- It requires access to a lot of articles which can be distributed across many databases.
  - This disadvantage was seen in the thesis when we discovered that we had no access to SAE Mobilus.
  - To compensate for this, the candidate studies from SAE Mobilus, which are in this review, was found with a complimentary search in other databases.
- Exclusion criteria can be applied a bit differently if the research team consists of many researchers.
  - A Cohens Kappa was calculated for each of the pilot tests where the exclusion criterion was applied. The Cohens Kappa value was high indicating that the researchers agreed on how to judge using this criterion.
- An SLR is very time-consuming and resource-demanding.

3.2 Article Two

Based on the definitions in [22] and [23] the research conducted in the second article is defined as a controlled laboratory experiment. This form of experiment gives the researchers control over what the participants do. In doing research in a controlled laboratory setting, the researchers manipulate the independent variables while keeping other factors constant and require that the participants are randomly assigned to the different experimental conditions.

The experiment conducted in the article was driver distraction research. Due to this, the experimental setup chosen was a laboratory setting over a field setting. This was due to safety reasons since the participants had to use a handheld mobile phone while driving, which is illegal in Denmark. Furthermore, experimenting with a laboratory setting is an advantage based on replicability. This could help against creating bias in the research. The experiment in the article was conducted as a within-subject design. The results of the experiment were analyzed using a two-paired samples t-test and Related-Samples Wilcoxon Signed-Rank Test.

Having a controlled laboratory experiment have advantages according to [23], such as:

- Having precise control over independent variables with randomized assignments, to isolate the effects of variables.
- Quantitative data can be used for statistical statements, by having it analyzed using inferential statistics.
- Having control over the environment also means that the experiment can be replicated.
  - It is possible in our experiment to reproduce the same scenarios since the leading vehicle will always follow the same route.
- Having a controlled environment also allows for precise data gathering by the equipment used.
  - This is applied to the simulator used in the thesis since it could measure completion time and lane deviations precisely as examples.

Controlled laboratory experiments, however, also hold some disadvantages according to [23]:

- When doing experiments in a controlled environment it is hard to verify the results when used in the field.
  - Conducting research in a controlled environment can make it hard to assess how it would behave in a more dynamic real-world context.
  - To compensate for this, the goal for the simulator was to be of high fidelity. Of the 16 participants in the experiment, only 4 of them said in their feedback, that it was not realistic. 10 of the participants said it was realistic, but had shortcomings, and two participants said it was realistic.

- Can also lead to artificial laboratory settings, which can have an impact on the validity of the results.
  - The participants can feel like they are inside a risk-free setting and would therefore react differently than they would in a real-world setting.
  - To reduce this possibility, we explained to the participants that they had to drive like they would in a real-world setting.

To gather data, automated logging in the simulator was used. This data was logged every half second of the specified driver performance parameters tested in the experiment. To signal when a task was started, while the participant was using a phone, a key would be pressed by one of the researchers so task-specific data could be found in the logging. The key was then pressed again when the task ended. When the HUD appeared the task-specific data was automatically marked in the logs. Alternatively, we had to manually log the performance parameters measured in this experiment, by reviewing video material from each participant and writing down everything related to it. Furthermore, this would also be a more time-consuming task than automated logging. Having more manual labor for the researchers could also potentially lead to more human errors.

The data gathered was tab separated into a text file, which could be parsed into an excel file. For statistical analysis, we used Excel and IBM SPSS statistics tools.

Chapter 4

Limitations

This section will present some limitations for the articles which should be taken into consideration. The limitations are sorted based on the affected article.

4.1 Article One

- The search query was only tested via Google Scholar before gathering candidate studies. This provided a general indication of the relevancy of the studies provided by the search query but could leave out relevant studies or include non-relevant studies depending on how the search query was interpreted by the databases.
- One of the exclusion criterion excluded non-peer-reviewed articles. The databases used in the article, did not have consistency when labelling studies. Therefore, some studies which could be peer-reviewed was excluded based on their label.

4.2 Article Two

- All of the results produced in the article are from a simulator and can therefore be hard to replicate the results in a real-world scenario.
- For testing it was chosen to use a driving environment, with no traffic congestion, except for the leading vehicle. This, therefore, does not reflect a real-world setting.
- Since the participants are driving in a simulated environment, it also meant that there would be no repercussion with some of the choices they could make. It could therefore be possible, that the participants would not react naturally as they would in certain situations.

Chapter 5

Conclusion & Future Work

This thesis, set out to answer the overall research question: *How can a HUD be designed and evaluated with focus on high driver performance, to support non-driving related features?*. To answer the overall research question, it was broken down into two sub-questions, that were answered in two articles.

The first research question, was associated with automotive development techniques and the research question was: *What techniques are used in research literature to design and evaluate automotive HMI's?*. The results showed that observations were used consistently more than the other techniques in each 5 year span followed by questionnaires. From the span 2005-2009 to the span 2010-2014, studies related to secondary systems saw a higher increase compared to tertiary systems. The span 2015-2021 sees a higher increase in tertiary systems than secondary systems. Observations and questionnaires were favorites when collecting quantitative data and had usage in all development phases without major differences. Interview, expert evaluation, and focus groups were used more for qualitative data and were used mostly in the earlier stages of development. A combination of multiple techniques is the most common way when developing and evaluating automotive HMI.

The second research question focused on finding out, what impact a voice-to-text-based HUD had as an impact on driver performance, and the research question was: *What impact would a HUD-based text message system with voice-to-text as input, have on driver performance?*. we showed by creating a prototype HUD in a simulator that using a HUD for voice-to-text based texting had an impact on driving performance, compared to texting with a phone in a low complexity driving setting. The participants showed a significant reduction in lane crossings, subjective workload, and had fewer crashes. The risk of crashing when using the HUD for texting over the phone was 3.5 times less likely. The completion time of tasks showed no significant differences, however, the participants were able to write longer messages when using the HUD.

Based on the results and the experience gained with the development techniques and testing of the HUD, new possibilities come to mind. The results of the HUD showed that it had a positive impact on the driving performance on some metrics, in a low congestion environment. As such, for future work, the HUD could be tested in a medium and high congestion environment to look at the impact of driver performance there. Another environment the HUD could be tested in is in sudden scenarios, where the participants have to react to sudden dangerous driving situations, such as a vehicle in front of the participants' car suddenly braking. Furthermore, the simulator used could be improved upon, by having screens as side windows in the car, while also applying sound, that mimics the environment the participants are driving in.

References

[1] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator”, in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[2] R. Svendsen and J. Skovby Bjørn, “Methods In Automotive HMI - A Systematic Literature Review”, English, Semester Project, Aalborg University, Jan. 2022. [Online]. Available: https://projekter.aau.dk/projekter/da/studentthesis/methods-in-automotive-hmi--a-systematic-literature-review(0383bcd4-807a-4777-8db5-99133744fa89).html.

[3] K. Lipovac, M. Đerić, M. Tešić, Z. Andrić, and B. Marić, “Mobile phone use while driving-literary review”, en, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 47, pp. 132–142, May 2017, ISSN: 1369-8478. DOI: 10.1016/j.trf.2017.04.015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847817302838 (visited on 03/10/2022).

[4] eLov, *eLOV.DK - Færdselsloven § 55a*, da. [Online]. Available: https://www.elov.dk/faerdselsloven/55a/ (visited on 03/09/2022).

[5] Transport- og Boligministeriet, *Vejledning om færdselslovens forbud mod anvendelse af håndholdte mobiltelefoner under kørsel*, May 1998. [Online]. Available: https://www.retsinformation.dk/eli/mt/1998/77 (visited on 03/09/2022).

[6] A. W. Jensen, *Derfor frister mobilen i trafikken*, da, Nov. 2019. [Online]. Available: https://fdm.dk/nyheder/trafik/2019-11-derfor-frister-mobilen-trafikken (visited on 03/07/2022).

[7] Jens, *Skriver du også sms’er bag rattet?*, da-DK, Jan. 2018. [Online]. Available: https://www.hvilkenbil.dk/skriver-du-ogsaa-smser-bag-rattet/ (visited on 03/09/2022).

[8] *| GHSA*. [Online]. Available: https://www.ghsa.org/state-laws/issues/distracted%5C%20driving (visited on 03/09/2022).

[9] National Highway Traffic Safety Administration, “Distracted Driving 2016”, ENG, *NHTSA*, p. 7, Apr. 2018. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812517.

[10] National Highway Traffic Safety Administration, “Distracted Driving 2018”, *NHTSA*, p. 7, Apr. 2020. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812926.

[11] National Highway Traffic Safety Administration, “Distracted Driving 2019”, ENG, *NHTSA*, p. 7, Apr. 2020, DOT HS 812 926. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813111.

[12] T. Arent, *Trods klip: Vi taler løs bag rattet – og bliver opdaget*, da, Mar. 2021. [Online]. Available: https://fdm.dk/nyheder/bilist/2021-03-trods-klip-vi-taler-loes-bag-rattet-bliver-opdaget (visited on 03/22/2022).

[13] *An Evaluation of the Effectiveness of Voice-to-Text Programs at Reducing Incidences of Distracted Driving | Blurbs New | Blurbs | Main*. [Online]. Available: https://www.trb.org/Main/Blurbs/168970.aspx (visited on 03/14/2022).

[14] D. Large, G. Burnett, V. Antrobus, and L. Skrypchuk, “Stimulating conversation: Engaging drivers in natural language interactions with an autonomous digital driving assistant to counteract passive task-related fatigue”, Mar. 2017.

[15] A. Baron and P. Green, “Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review”, Feb. 2006. [Online]. Available: https://trid.trb.org/view/806083 (visited on 03/18/2022).

[16] G. Villalobos-Zúñiga, T. Kujala, and A. Oulasvirta, “T9+HUD: Physical Keypad and HUD can Improve Driving Performance while Typing and Driving”, en, in *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ann Arbor MI USA: ACM, Oct. 2016, pp. 177–184, ISBN: 9781450345330. DOI: 10.1145/3003715.3005453. [Online]. Available: https://dl.acm.org/doi/10.1145/3003715.3005453 (visited on 03/18/2022).

[17] M. B. Hoy, “Alexa, siri, cortana, and more: An introduction to voice assistants”, *Medical Reference Services Quarterly*, vol. 37, no. 1, pp. 81–88, 2018, PMID: 29327988. DOI: 10.1080/02763869.2018.1404391. [Online]. Available: https://doi.org/10.1080/02763869.2018.1404391.

[18] A. Berdasco, G. López, I. Diaz, L. Quesada, and L. A. Guerrero, “User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana”, en, MDPI, Nov. 2019, p. 51. DOI: 10.3390/proceedings2019031051. [Online]. Available: https://www.mdpi.com/2504-3900/31/1/51 (visited on 03/22/2022).

[19] B. A. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering”, English, Keele University and Durham University Joint Report, Tech. Rep. EBSE 2007-001, Jul. 2007. [Online]. Available: https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf.

[20] *Assessing and addressing bias in systematic reviews*, en-GB, Apr. 2018. [Online]. Available: https://meta-evidence.co.uk/assessing-and-addressing-bias-in-systematic-reviews/ (visited on 06/10/2022).

[21] R. Mallett, J. Hagen-Zanker, R. Slater, and M. Duvendack, “The benefits and challenges of using systematic reviews in international development research”, *Journal of Development Effectiveness*, vol. 4, no. 3, pp. 445–455, Sep. 2012, ISSN: 1943-9342. DOI: 10.1080/19439342.2012.711342. [Online]. Available: https://doi.org/10.1080/19439342.2012.711342 (visited on 06/14/2022).

[22] J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human-Computer Interaction*, eng. San Francisco: Elsevier Science & Technology, 2017, ISBN: 0128053909.

[23] J. S. Olson and W. A. Kellogg, *Ways of Knowing in HCI*. Springer, 2014, vol. 2.

Part II

Papers

Developing HMIs: A Literature Review of Automotive Development Techniques

Jeppe Skovby Bjørn
Department of Computer Science
Aalborg University
Selma Lagerløfs Vej 300
Aalborg, Denmark
jbjarn17@student.aau.dk

Ronnie Svendsen
Department of Computer Science
Aalborg University
Selma Lagerløfs Vej 300
Aalborg, Denmark
rsvend17@student.aau.dk

ABSTRACT

When developing and evaluating automotive HMI, the techniques used differ from study to study. It can be hard to figure out what to use, and when. The objective of the study was to examine the current usage of design and evaluation techniques utilized when developing automotive HMI. A Structured Literature Review (SLR) was conducted to provide an overview of the current state of research in the area. A total of 907 candidate studies, published between 2005-2021, were gathered. After the selection process, 87 of the candidate studies became primary studies. The categorization of the primary studies revealed that observations and questionnaires are the most used techniques under development. Interview, focus group, self report, and expert evaluation only covered 13% of the total usage. The most common way the techniques are used is in combination with each other. Furthermore, 78% of the data collected from the studies was quantitative data. The most researched form of system was secondary and tertiary.

KEYWORDS

Systematic Literature Review; Automotive Engineering; Development Techniques; Human-Machine Interaction; Automotive

1 INTRODUCTION

Software in cars is evolving and so are the interfaces used to connect the driver with the car. These interfaces control how a driver can receive information and act accordingly. [1]

In the early years of automotive transportation, most **H**uman-**M**achine **I**nterfaces(HMI) were mechanically operated. However, with the introduction of entertainment systems, more complex and digital HMI's was implemented into the cars. [1]

The amount of hardware in cars is also increasing along with the amount of software. This can be seen by the number of electronic control units and the number of networks used to connect these in the vehicle. Because of this, most innovations in cars are more or less supported by the software. This is very evident in more high-end cars, which have more than 70 embedded platforms that implement various functionality. [2]

There are three categories that can be used to determine the responsibilities of in-vehicle systems: Primary, secondary, and tertiary. The purpose of the systems belonging to the primary category are to control the car. This could be the steering wheel, speeder, and brake pedal. The secondary category contains the systems which support safe driving like windshield wipers and mirrors. The third category, tertiary, is used for non-driving related interactions, like using the radio or changing the temperature in the car. [3]

In all three categories, driving performance must be considered since systems that interact with the driver impact this factor in some way. For the primary and secondary categories, this impact will most likely increase driving performance and safety. The steering wheel is an example of a primary system as it has a direct influence on the control of the car. Windshield wipers are considered secondary as it supports the vision and, thereby, enhance safety. Most tertiary systems impact the driving performance negatively if the attention of the driver is diverted towards these systems for too long [3]. Examples of these systems could be the radio, navigation, or phone calls through the in-car media system. It is, therefore, important to test these systems in order to verify that the HMI works as intended. A complex interface might be able to assist the designers' driving performance but have the opposite effect on potential users. These tests can be conducted in a variety of techniques that support different scenarios of development.

The development techniques used when developing automotive HMI differs from study to study. For example, [4] used observation when developing an augmented reality HUD which gives crash warnings to the driver, whereas [5] used expert evaluation when developing a road sign system HMI. When one wants to start researching automotive HMI, it can be hard to know what techniques are used, since it seems that when just browsing through some different papers.

The main objective of the study is the examine the current usages of design and evaluation techniques utilized by researchers when designing or evaluating HMI's. This means that we aim to look at peer-reviewed research to characterize the different design and evaluation techniques that are used by researchers and possible practitioners in the context of automotive software engineering for HMI's.

To reflect the main objective of the study, the main research question that is to be answered is: *What techniques are used in the research literature to design and evaluate automotive HMI's?*

To answer this question, a **S**tructured **L**iterature **R**eview(SLR) was conducted to provide an overview of the techniques used for developing automotive HMI's. The SLR will follow the guidelines presented in [6]. Candidate studies will be obtained through a search strategy. These will go through a selection process and become primary studies that will be used for information extraction.

2 RELATED WORK

This section will look deeper into two other articles which investigate similar topics. The methods used to conduct these SLRs will be described to further understand how an SLR can be conducted.

2.1 Automotive Software Engineering

An SLR was conducted in [7]. The paper aims to establish an overview of the research that has been done in the automotive domain. To guide the study, [7] introduces the following main research question: *What is the current state of the research pertaining to software engineering in automotive industry?*

Initial research revealed that in the late 70s the first software was introduced to the car industry. Since then, the software in cars has only increased rapidly and it is not expected that this growth will stop any time soon.

To establish this overview, research articles/papers from various databases were collected and used to provide an overview. A total of 679 papers were used as primary studies literature for the review.

The primary studies are categorized into a 2-dimensional diagram which provides an overview of the literature and which areas are more favorable and where the research gaps are located. The focus of these results is heavily on the overall technical part of software engineering which can make it harder to apply to the user experience part of the automotive research.

Our study differs from [7] since we are looking at the current development techniques used in the research of automotive HMI's.

2.2 UX & Evaluation Mehtods

A study from 2015[8] investigates user experience(UX) and how this term is used. They investigate the definition of UX, which knowledge is already present about UX, and what are the research trends among UX researchers. To answer the questions, the study utilizes a literature review to gather papers relevant to the study. In total, 90 papers were reviewed in the study with articles up to 2014. It did not seem to have a restriction on the year of publication, but only four of the included papers were published before the year 2000.

The overall results are aggregated in a 2-dimensional table with one axis representing the techniques used and the other axis representing the evaluation period, development phase, and what type of data was collected(qualitative and quantitative). This table provides a great overview of the techniques used and when they were used.

The study shows that some techniques, like interviews and self reports, have seen an increase in usage over time. In the opposite direction, expert evaluation and observation have seen a decrease in usage.

The results from the study are UX in every topic in software engineering and a more specific view of techniques used in the automotive spectrum is needed to answer the research question presented in this paper.

3 METHOD

To answer our research question, it was chosen to conduct an SLR. This section will go through the different aspects of an SLR, which include the selection criteria, the search strategy, the selection process, and information extraction.

3.1 Selection Criteria

To select and sort the different candidate studies, an inclusion criterion and a set of exclusion criteria were formulated. The inclusion criterion of the study was that if it was found by the search query refined in Section 3.2, it was to be included in the study, as a candidate study. The following exclusion criteria were applied to the candidate studies, to ensure that it was in the context of this study:

1. Published before 2005.
2. Duplicate article.
3. Not written in English.
4. Not Peer-reviewed.
5. Does not design or evaluate automotive HMI.

Exclusion criterion 1 was exploratory at implementation. However, based on the results the candidate studies that became primary studies from 2005-2021, averaged to be 11.67% each year. Based on this average, 2000-2004 would only retrieve a total of 7-8 studies in this SLR. Based on this, it was assumed that it would not shift the end result of the SLR, and it was therefore chosen to only include articles after 2005.

Exclusion criterion 2 ensured that the SLR would not contain several copies of the same article. Therefore, an article is a duplicate if that article shares the same list of authors and title with another.

Criterion 3 is to ensure that the studies in the SLR are readable for the researchers.

To ensure that the studies in the SLR are of some quality, everything not peer-reviewed is excluded as per criterion 4. Peer-reviewed studies typically mean conference papers, short papers, and technical papers.

The paper's and the SLR's focus is to look at techniques used when developing HMIs. As such, criterion 5 ensured that the objective of the paper, meaning that tools/techniques found in papers are only included in the SLR if used to develop an HMI. This means, that papers that only introduce techniques, but do not directly utilize the technique to develop automotive HMI in the study, would be excluded.

3.2 Search Strategy

To ensure the relevance of the articles, a search strategy was formed. According to the guidelines [6], four databases are sufficient for an SLR. Therefore, the top four databases of [9] are used to acquire articles for the SLR since this paper also looks at automotive research. These databases can be seen in Table 1. It is chosen to create a search query, to search for articles on the selected databases.

The search strategy used an iterative approach to refine the search query. An iteration occurred to refine the search query if it produced a result of over 1000 articles that had to be reviewed by the researchers in the period of the study.

Table 1: Searched databases and its results with the search query

Database:	Searched Date	Result
SAE Mobilus	14-12-2021	336
ACM	15-12-2021	93
SCOPUS	14-12-2021	390
IEEXPLORE	14-12-2021	88
Total result:		907

A pilot test [9] was used to refine the search query. The relevance of the results was found by looking at the first 100 articles on google scholar with the search query. Relevance is judged based on exclusion criterion 5 from the article's title, abstract, and keywords. If an article was irrelevant, it was examined further to identify why the search query included it. Cohen's kappa measures the reliability of the two researchers' rates the same, corrected by how often the researchers may agree by chance. Cohen's Kappa varies between 0-1, where 1 is the perfect agreement. [10], [11]

The search query went through two iterations. In the first iteration, the search query produced a result of 14,734 from the databases. This was considered too many articles and the pilot test was utilized. Each researcher deemed the relevance of the first 100 results from google scholar to be 60% and 67%, with a Cohen's Kappa of 0,76. It was discovered that the usage of a general term *Vehicle*, which could also refer to other types of vehicles, like an underwater vehicle was the reason for finding different articles that were not relevant.

The second iteration produced a total number of articles of 14,002 from the databases, which again was too many results, and the pilot test was conducted again. The second search query produced more relevant results from the first 100 results on google scholar, with a relevance factor of 84% and 85%´from the researchers and a Cohen's Kappa of 0,8. Examining the articles closer did not find any terms that could necessarily introduce irrelevant articles. Thus, it was chosen to remove typed-out words of abbreviations, such as *user interface* & *human machine interface* to narrow the total results from the databases.

This resulted in a total amount of articles of 907 from the databases which meant that it was down to a manageable amount for the two researchers of the study. The search query that is used on all of the databases is:

automotive AND (ui OR hmi) AND (design OR evaluation OR method)

However, when searching the ACM database some articles that were found in the preliminary research were not included in the search query. This was because the search string was too strict. As such, it was chosen to exploratory alter the search query by removing *(AND (design OR evaluation OR method)* when searching on ACM. This resulted in the articles being included in the study.

3.3 Selection Process

The results from using the search strategy resulted in 907 candidate studies, as seen in Table 1. The exclusion criteria were applied one by one, starting with 1 through 5, as seen in Figure 1. As seen in the figure, the first four exclusion criteria exclude 173, 133, 33, and 88, respectively. This leaves 480 candidate studies to be evaluated by criterion 5. Criterion 5 will be used over two iterations, where it first would be applied to the title & abstract and lastly a full-text review. Criterion 5 applied to the title & abstract excluded 297 candidate studies, leaving 183 for a full-text review. If an evaluation of an article on the title & abstract was inconclusive, meaning that the title & abstract indicated that the candidate study developed an automotive HMI, but not enough to conclude, it was included in the full-text review pool.

Due to limited access to SAE Mobilus from the university, not all articles could be included. This was since the university did not have a license to access articles from the database. Therefore, 65 candidate studies had to be excluded where there are 24 articles from SAE Mobilus remaining due to these articles being accessible from other databases.

This is not seen as a limitation to the study, since the articles from SAE Mobilus that became primary studies, did not alter the results. This left 118 candidate studies to be evaluated by criterion 5 full-text review.

Criterion 5 full-text review excluded a total of 31 articles. This means that 87 articles are primary studies in this study.

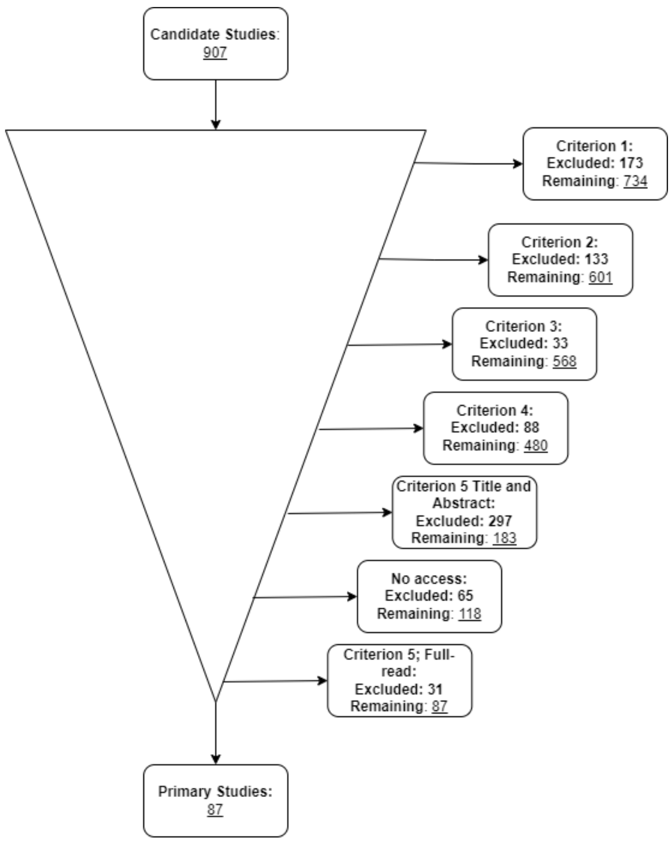

Figure 1: Selection process of the candidate studies

3.4 Information Extraction

Using the table from [8] as a foundation for the results of this paper enables the possibility of easy comparison between the two results.

Five types of data will be extracted from the primary studies in order to create a foundation of data. These five types of data are the year in which the article was published, which system is being designed/evaluated, which techniques are used to gather data, how users are involved, and in which part of the development users are involved.

Table 4 shows the results of this SLR and is created using inspiration from [8]. The paper brings six different techniques which will also be extracted from the primary studies. These are interview, questionnaire, observation, focus group, self report, and expert evaluation. Self report is more like a diary in this study where the participant will provide feedback over a period of time. In addition to those, applied research has been added to this poll of techniques. The techniques will be plotted into a table to provide an overview of the period of evaluation, the development phase, and the type of data acquired in the study.

4 RESULTS

This section will contain the findings of the data extraction from the primary studies. The findings were divided into subsections that examine different aspects of the results.

The primary studies (n=87) found by the selection process in 3.3 were used for data extraction. In the results, the format used is "study Nr. ##.." to refer to primary studies in the SLR. The citations for the primary studies can be found in *SLR References* in the appendix.

4.1 Publication and Technique Usage Trends

The number of primary studies found for each year is seen in Figure 2. An average of 11.67% (± 4.32) of the candidate studies per year became primary studies. The years with the least amount of studies are 2006, 2008, and 2017. This suggests that the area of research sees a fluctuation in popularity over the years. The year with the most amount of studies is 2019 which indicated that research in the area has seen a dip in interest since the number of studies from 2020 and 2021 is low compared to 2019. One could consider this dip a consequence of Covid-19 however, no evidence suggests this. The accumulated numbers of primary studies from 2005-2009, 2010-2014, and 2015-2021 are 16, 28, and 43, respectively, as calculated from Figure 1. This could indicate an increase in activity in this field of study over the years. A possible reason for the increased activity is that cars are more and more evaluated by the amount of software in a car and cars, in general, had more software in them leading to more activity. [2]

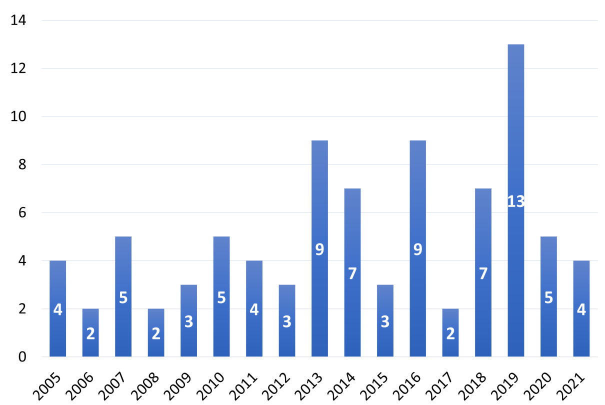

Figure 2: Primary studies pr. year

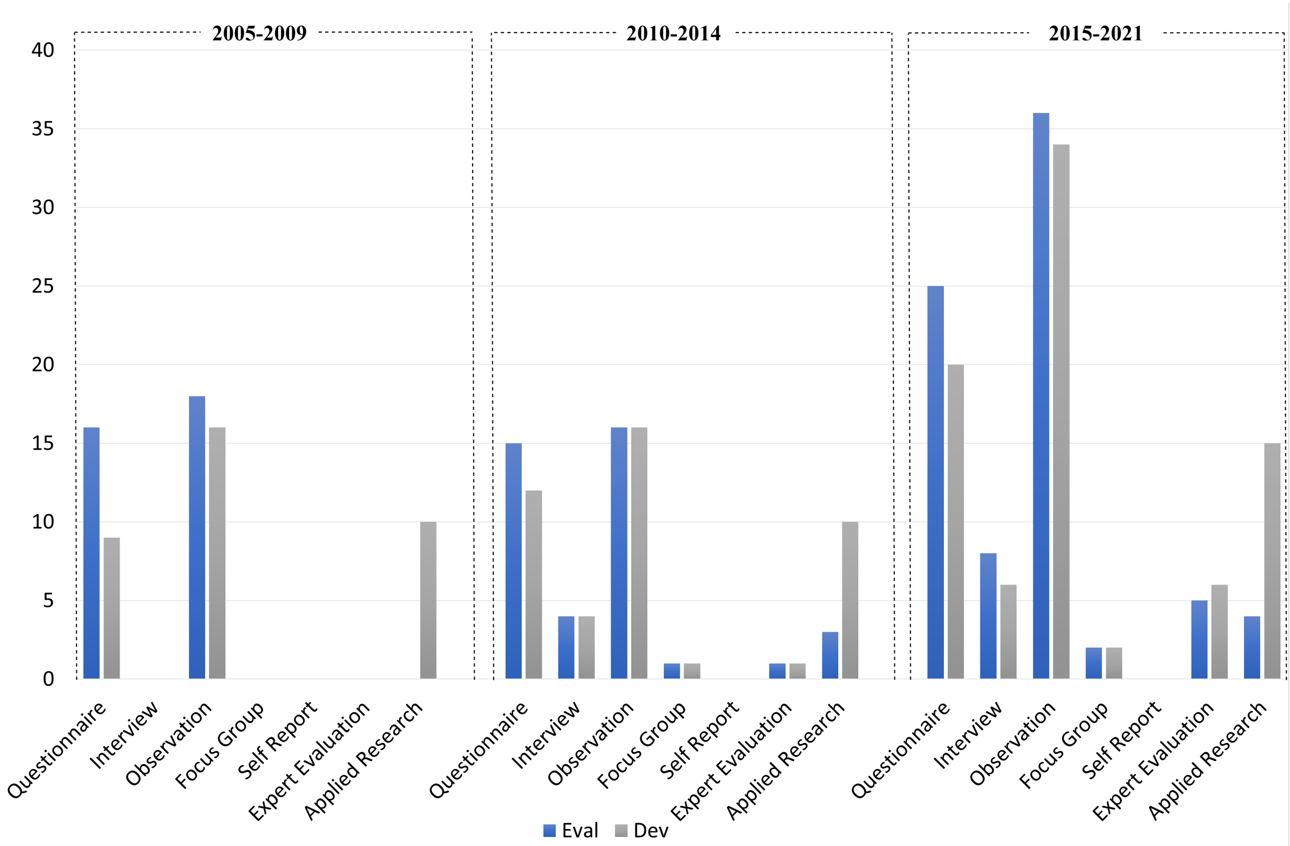

Figure 3: Method Usages in five year spans

Figure 3 depicts the usage trends, from 2005-2021. From 2005 to 2021, the most used techniques in evaluation and development are Applied Research(n=32), Questionnaire(n=52), and Observation(n=57). Looking at the techniques, divided into five-year spans from 2005 to 2021 also revealed that these techniques are the most used in every span, where the observation technique is the most used of them. Observation and Questionnaire cover 73% of the total technique usage for both evaluation and development, whereas applied research covers only 14%. This implies that user involvement is widely spread when developing automotive HMI.

Comparing the usage of questionnaires and observation in the three spans shows that the utilization is close. However, in 2015-2021 observation is pulling ahead of questionnaires. This suggests that questionnaires are losing popularity as a technique. In a contrast, the other development techniques, interview, focus group, and expert evaluation have seen a notable increase in the same span.

4.2 Technique Utilization

This section looks into how the primary studies utilize the techniques, to develop and evaluate their automotive HMI. This will be done by examining the primary studies in the SLR.

One of the most typical ways the primary studies use the techniques is in combination with other techniques. Study Nr.14 is an example of this, which use observation and questionnaire in combination to develop their automotive HMI. The participants in the study answered a questionnaire beforehand, for demographic data, whereas they go into the test, where the researchers observed the participants doing tasks. After each task, the participants were asked a set of questions and then they continued with the rest of the tasks. The questions used in the questionnaires were designed from the Likert scale [12]. Some primary studies, such as study Nr.40 that used a simulator to test HMI, also get the participants to fill out a questionnaire for simulator sickness from [13]. After the last task, the participants were given a more general questionnaire about the study. This can be seen in several different studies, such as study Nr. 52, 29, 39, and 40.

An example of the utilization of applied research is study Nr.81 that utilizes already known research and information to develop an early prototype automotive HMI. If the area is heavily researched, developers can benefit from this by using this knowledge to build the first draft. This can provide a more tangible product.

As seen in Table 4, expert evaluation has the most usage in the early phases of development. Study Nr. 105, which explored directional indicators for conditional driving automation, used three HMI designers to guide the concept design for the study. The quality of the concepts, which later was evaluated further, was increased using the feedback from the experts.

Interview, focus group, self report, and expert evaluation only cover 13% of techniques used.

The technique self report has not been used once in any of the primary studies. This indicates that self-reporting as a tool is not popular to use when developing automotive HMI.

Self report can be a very time-consuming process that takes several days to complete. This is very problematic when developing automotive HMI since it would require a lot of time, and resources, which may not be available during the development and evaluation process. Furthermore, efficiency is important in the product development context, where self report might not be seen as efficient since it is a slow process. [14]

Table 4 revealed that the technique Interview is only used with other techniques. The primary studies 2, 14, and 59 used interviews at the end of their experiments with automotive HMI to get more qualitative data from the participants, using a semi-structured interview.

Table 4: Every primary study(n=87) categorized

		Interview	Questionnaire	Observation	Focus group	Self report	Expert evaluation	Applied research	Total:
Period of evaluation	**Before**	11, 20, 43	11, 14, 23, 26, 30, 90, 94, 104		93		35, 93		**14**
	During	14, 55, 93	4, 21, 23, 26, 29, 35, 39, 40, 45, 52, 61, 62, 83, 90, 93, 94, 104	1, 2, 3, 4, 6, 11, 13, 14, 15, 17, 18, 29, 30, 35, 36, 39, 40, 42, 43, 45, 47, 51, 52, 54, 55, 57, 58, 59, 61, 62, 66, 68, 70, 72, 75, 77, 78, 80, 86, 87, 90, 91, 93, 94, 95, 96, 97, 99, 100, 101, 102, 103, 104, 105, 107, 108, 109			37, 105	10, 12, 16, 20, 64	**84**
	After	2, 14, 15, 26, 59, 64	1, 6, 11, 13, 14, 15, 18, 19, 23, 29, 40, 42, 52, 66, 70, 72, 75, 78, 80, 82, 86, 89, 91, 94, 104		27		88	16, 25	**35**
Total nr. of papers:		**12**	**50**	**57**	**2**	**0**	**5**	**7**	**Total Evaluation: 133**
Development phase	**Concepts**	14, 15, 43, 55	14, 15, 30, 72, 91, 101	14, 15, 30, 43, 55, 58, 68, 72, 77, 91, 95, 99, 100, 101, 103, 107, 108, 109	93		35, 37, 93, 105	4, 5, 9, 20, 25, 48, 49, 63, 64, 67, 69, 92, 98, 102,106	**48**
	Early prototypes	11, 26, 59, 64, 93	4, 6, 11, 13, 18, 26, 29, 35, 39, 40, 42, 45, 52, 61, 62, 66, 70, 75, 78, 80, 82, 83, 90, 93, 104	4, 6, 11, 13, 18, 29, 35, 36, 39, 40, 42, 45, 47, 51, 52, 54, 59, 61, 62, 66, 70, 75, 78, 80, 87, 90, 93, 97, 104			105	10, 26, 31, 35, 48, 60, 64, 81, 106	**69**
	Final prototypes or products	2	1, 19, 21, 23, 86, 89, 94	1, 2, 17, 57, 86, 94, 96, 101, 102, 105	27		88	3, 12, 31	**23**
Total nr. of papers:		**10**	**38**	**57**	**2**	**0**	**6**	**32**	**Total Development: 140**
Data	**Qualitative**	2, 14, 15, 20, 26, 55, 59, 93	18, 19, 83, 91, 93, 94	15, 17, 51, 55, 57, 91, 93	27, 93		35, 88, 105	20	**27**
	Quantitative	11, 14, 43, 59, 93	1, 4, 6, 11, 13, 14, 15, 18, 19, 21, 23, 26, 29, 30, 35, 39, 40, 42, 45, 52, 61, 62, 66, 70, 72, 75, 78, 80, 82, 86, 89, 90, 94, 101, 104	1, 2, 3, 4, 6, 11, 13, 14, 18, 29, 30, 35, 36, 39, 40, 42, 43, 45, 47, 51, 52, 54, 58, 59, 61, 62, 66, 68, 70, 72, 75, 77, 78, 80, 86, 87, 90, 94, 95, 96, 97, 99, 100, 101, 102, 103, 104, 105, 107, 108, 109			37, 93	10, 98	**95**
Total nr. of papers:		**13**	**41**	**58**	**2**	**0**	**5**	**3**	**Total Data: 122**

4.3 Data Acquisition

This section will explore how the different primary studies acquired data through their research and how the data was utilized based on it either being qualitative or quantitative. Table 4 revealed what type of data is collected when developing automotive HMI. A majority (78%) of the data collected is quantitative.

A downside to using quantitative data is that not all data is useful in a restricted environment. Furthermore, collecting opinion-based quantitative data only brings results of how the participants perceived their use and meaning of the automotive HMI. An example of this is study Nr.23 which is an intersection assistant based on inter-vehicle-communication, that used questionnaires as a means of assessment of their automotive HMI. These quantitative data are used to examine how the automotive HMI affects the driver of the vehicle based on gender and experience and the participants' opinions.

Quantitative data are manageable since they can be locked to a certain data type and measured equally. This allows for quick analysis using tools to provide quick results. Paper study Nr.45 is an example of this, which studied HUD augmented reality crash systems, and acquires more reliable quantitative data through observation in a simulator-based study on how the participants actually used the automotive HMI. The data collected are driver performance parameters such as mental workload and preferences of Driver vehicle interfaces tested. These data are applied to statistical analysis, such as one-way repeated measures to test for the statistical significance of the results.

Qualitative data opens the opportunity for a broader variation of data and feedback as it is not restricted in the same way as quantitative data. Study Nr.27, as an example, used focus groups to acquire data about the problems the participants had while driving with the interior interaction design. It revealed problems with the vehicle's interactive ergonomics that cause the users to double their workload due to the placements of switches.

Another example is Study Nr.2, which used interviews to collect qualitative data on three different interaction forms; physical, digital, and voice. The data collected is to obtain the users' views on the interaction forms and their possible inconveniences. The results are summarized feedback from the interviews.

Some primary studies chose to use a mixed-method and acquired both quantitative and qualitative data. Mixed-method adds to the complexity of the study and increases the time needed[15]. study Nr.59 uses a mixed-method by having interviews and observations for each participant in the study to find more nuanced answers to some of the results. By acquiring quantitative data through observation, the paper gets an understanding of how the participants used the automotive HMI, by a set of measurable parameters. With the qualitative data acquired through interviews, it was possible to get an understanding of what factored into the results, such as possible design decisions of the automotive HUD that had an impact on the participants' driver performance.

4.4 Categorization of Automotive HMIs

Table 3 depicts the results of the kind of systems developed over five-year spans, from the definition of primary, secondary, and tertiary categorization in [3]. Secondary systems(n=56) and tertiary systems (n=51), which is not a notable difference. The total number of categorized systems is higher than the total number of primary studies in the SLR. This is because an automotive HMI can be in more than one category, dependent on the system. An example of this is study Nr.59. This system can be seen as both secondary and tertiary because the live traffic preview on the GPS gives safety, and the GPS part of the system only brings navigation. In each of the spans, examples will be used to look at the tendencies of the typically developed systems from each category in that span.

From 2005 to 2009, a total of 20 systems were developed. Only one primary system was developed, which is study Nr.104. This study compared three different steering interfaces, a steering wheel with pedals, a joystick, and a robotic grip when driving a vehicle. This study found that driver experience can be enhanced by 76.3% when using a robotic grip. study Nr.68 is a secondary system in the span that examined **M**ixed **R**eality (MR) sickness, which is a motion sickness caused by a sensory conflict between the real world, and the virtual world. The study focused on reducing MR sickness in connection to HUDs, modeled after ADAS systems.

Study Nr.23 is another secondary system in the span, developed an intersection assistant based on inter-vehicle-communication, which was created to help drivers assess situations and prevent accidents. The results showed that the system can assist drivers and improve safety.

An example of a tertiary system in the span is study Nr.13, which studied haptic feedback from a touch screen interface. The results indicated user acceptance of the haptic feedback and preference for multi-modal feedback. Another tertiary system in the span is the study Nr.21 that evaluated push switches both in and out of a vehicle that investigated the effect of context. The result of the study was that customers partially discriminated in the feel of the switches.

From 2010 to 2014, a total of 32 different systems were developed. Study Nr.51 is a secondary system developed in the span. This study focused on developing a full-windshield HUD interface that aimed to increase the driver's awareness and response time in low visibility conditions. The HUD represents a minimalist form of real objects that offers a form of guidance on the motorway. study Nr.37 is another secondary system developed in the span. This study designed an automotive HMI that displays road signs in the tachometer and split-screen on the navigation map to protect children at a school bus stop. A tertiary system, study Nr.49, developed in the span, designed a climate control system with automatic climate control. Another tertiary system study Nr.63 examined ways to optimize automotive displays under heavy light and creates an image enhancement algorithm that optimizes the perception of automotive HMI for bright ambient light.

In 2010-2014, secondary and tertiary systems had a small difference, in the amount of systems developed. However, in 2015-2021, there was no notable difference to developed tertiary and secondary systems developed. This indicates that more systems are being developed in automotive HMI for entertainment while driving and corresponds with how the evolution of automotive HMI has been as described in [16].

Table 3: Category of systems five year spans

	2005-2009	2010-2014	2015-2021	Total:
Primary	1	0	2	3
Secondary	10	19	27	56
Tertiary	9	13	29	51
Total:	20	32	58	

From 2015-to 2021, a total of 58 systems were developed. An example of a secondary system being developed in the span of 2015-2021 is study Nr.34. This HMI looks at an Automotive Advanced Driver Assistance System (ADAS) that warns drivers of pending dangers. This study tests an emotionally adaptive voice alert system that changes the voice alerts depending on the drivers' mood. This is done using a Deep Learning-based emotion recognition system. Another secondary system from the span is study Nr.72. This study examines how to design the take-over-requests when an automated vehicle tells the driver to control the vehicle.

An example of a tertiary system is study Nr.14, which developed and tested four different voice assistant personalities for in-car speech interfaces in a real-world driving study. The results of the study showed that users had higher likability and trust in assistants that matched their personalities. Another tertiary system in the 2015-2021 span is study Nr.2, which compares three different interaction forms for automotive HMI in a vehicle. These three interaction forms are physical, digital, and voice. From a user satisfaction point of view, the users preferred voice interaction. However, users needed a longer time to finish tasks with this interaction form.

As seen in Table 3, the number of developed systems sees a notable increase in each span. An indication of why is that automotive vehicles in general, from low-end to luxury cars, have more software, and it is possible to have more systems embedded. [17]

Over the spans, primary systems have little activity wherein 2010-2014 it did not see any. However, from 2015 to 2021 two primary systems have been developed. study Nr.26 is one of those primary systems which creates a haptic feedback pedal for better driving performance to reduce battery consumption in an electric vehicle. study Nr.91 is the other primary system developed in 2015-2021. This system is a user-defined voice and mid-air gesture command to maneuver an automated vehicle.

5 Discussion

The results of this study show that observation is the most used technique for including the users in the development phases as seen in Table 4. It is followed by questionnaire, applied research, and interview respectively. Compared to [8], it can be seen that some of the same techniques are present in the top four most used techniques, which can be seen in Table 2. However, techniques like observation were seen noticeably more used in the results of this study compared to the results in [8].

Table 2: Top Four Used Techniques

This Article (N=140)	Results from[8] (N=134)
Observation (n=57)	Questionnaire (n=66)
Questionnaire (n=38)	Interview (n=24)
Applied Research (n=32)	Observation (n=18)
Interview (n=10)	Self Report (n=17)

Observation is present in both results but the number of times it has been utilized is different. It is utilized 57 times in the automotive research field according to the results of this paper but only 18 times in the overall results from [8]. The difference between these two studies is that [8] studies everything which revolves around UX in the software engineering field.

For some applications or software, this seems to be a minor inconvenience as the risk of testing software in the field can be rather risk-free or the consequences can be somewhat manageable depending on the risk of the system. Systems, which are developed for the driver of a vehicle, can affect the driver's performance and, thereby, affect the reaction time from an incident that happens to the driver taking action to avoid events leading to accidents involving humans.

By introducing simulations the developers have the opportunity to deploy the product in a more or less real-life scenario to test and verify some of the requirements set for the system. If a system is built as a primary or a secondary category, the results should improve the driver's performance or at least keep it steady. In the case of the tertiary systems, the driver performance can be impacted but should be in accordance with the guidelines in [18].

Even though observation seems to be the way to go for most software solutions, the catch of using it needs to be considered. If the system developed is very simple, and the risk of a system failure is considered low, observations of the product in a realistic environment can be expensive and time-consuming.

6 CONCLUSION

This paper's main objective, was to answer the research question; *What techniques are used in the research literature to design and evaluate automotive HMI's?*. To answer this, we conducted an SLR, to identify primary studies that were relevant to the research question.

A total of 87 primary studies were found for our study. These primary studies were then categorized based on their data extracted.

Looking at the accumulated number of primary studies from 2005-2009, 2010-2014, and 2015-2021, revealed that each span saw an increase in the number of publications in the context of the SLR. The usage of the techniques showed that observation and questionnaire covered 73% of the total technique usage in the span of 2005-2021. Furthermore, interview, focus group, self report, and expert evaluation covers a total of 13% and applied research covers 14%. The rest is covered by These two techniques were roughly used the same number of times in 2005-2009 and 2010-2014, however, in 2015-2021 observations pulled notable ahead. The most common way the techniques are used is in combination with other techniques, such as observation and questionnaire. Interviews were only used coupled with other techniques.

The primary studies also revealed that 78% of the data collected was quantitative data. The data was used to get an understanding of how the users used the respective automotive HMI based on driver performance parameters. These data were then analyzed with statistical data analysis. The qualitative research collected was used to understand why the users used the automotive HMI as they did.

The most researched systems in the primary studies were secondary and tertiary with no notable difference. The kind of secondary and tertiary systems developed varied from ADAS systems to evaluating interaction forms. Primary systems saw little to no development, however, from 2015 to 2021 two primary systems were developed. One of these primary systems was a user-defined mid-air gesture command to maneuver an automated vehicle and a haptic feedback pedal for better driving performance.

Future Work

The technique usage discovered in this article, revealed that the techniques are used intertwined. Future research in the area, could look into the pros and cons of using the different techniques with each other, and find the most optimal mix, with respect to time, capital, and efficiency. As such, one would have to look at what the different techniques bring to the development of automotive HMI and how they are used.

The search query was made as a "one type fits all". A future study could manipulate the searching process to include multiple search queries, one for each database, in order to see if this could have an impact on the results.

References

[1] G. Meixner, C. Häcker, B. Decker, *et al.*, "Retrospective and Future Automotive Infotainment Systems—100 Years of User Interface Evolution", en, in *Automotive User Interfaces: Creating Interactive Experiences in the Car*, ser. Human–Computer Interaction Series, G. Meixner and C. Müller, Eds., Cham: Springer International Publishing, 2017, pp. 3–53, ISBN: 9783319494487. DOI: 10.1007/978-3-319-49448-7_1. [Online]. Available: https://doi.org/10.1007/978-3-319-49448-7_1 (visited on 01/05/2022).

[2] A. Pretschner, M. Broy, I. H. Kruger, and T. Stauner, "Software Engineering for Automotive Systems: A Roadmap", in *Future of Software Engineering (FOSE '07)*, May 2007, pp. 55–71. DOI: 10.1109/FOSE.2007.22.

[3] C. U. Okere, "Improving Automobile Driver Safety and Experience when Performing a Tertiary Task using Visual, Hands- free and Gestural-based Interfaces", English, in *International Journal of Information Technology (IJIT)*, vol. Volume 7 Issue 2, Department of Computer Science, The State University of New York at Oswego, Feb. 2021, p. 13. [Online]. Available: http://ijitjournal.org/volume-7/issue-2/IJIT-V7I2P2.pdf.

[4] H. Kim, X. Wu, J. L. Gabbard, and N. F. Polys, "Exploring head-up augmented reality interfaces for crash warning systems", in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ser. AutomotiveUI '13, New York, NY, USA:

Association for Computing Machinery, Oct. 2013, pp. 224–227, ISBN: 9781450324786. DOI: 10.1145/2516540.2516566. [Online]. Available: https://doi.org/10.1145/2516540.2516566 (visited on 06/10/2022).

[5] M. Ganzhorn, J. P. F. Diederichs, and H. Widlroither, “Enhancing the reproduction of road signs in automotive hmis to protect children”, in *2011 11th International Conference on ITS Telecommunications*, 2011, pp. 338–344. DOI: 10.1109/ITST.2011.6060078.

[6] B. A. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering”, English, Keele University and Durham University Joint Report, Tech. Rep. EBSE 2007-001, Jul. 2007. [Online]. Available: https://www.elsevier.com/__data/promis_misc/525444systematicreviewsguide.pdf.

[7] A. Haghighatkhah, A. Banijamali, O.-P. Pakanen, M. Oivo, and P. Kuvaja, “Automotive software engineering: A systematic mapping study”, en, *Journal of Systems and Software*, vol. 128, pp. 25–55, Jun. 2017, ISSN: 0164-1212. DOI: 10.1016/j.jss.2017.03.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0164121217300560 (visited on 04/08/2022).

[8] K. Bang, M. A. Kanstrup, and A. Kjems, *What are They Measuring? A Literature Review of Empirical Studies of UX Evaluation Methods*, English, 2015. [Online]. Available: https://projekter.aau.dk/projekter/files/214022189/UX_Evaluation_Methods.pdf.

[9] A. Haghighatkhah, A. Banijamali, O.-P. Pakanen, M. Oivo, and P. Kuvaja, “Automotive software engineering: A systematic mapping study”, *Journal of Systems and Software*, vol. 128, pp. 25–55, 2017, ISSN: 0164-1212. DOI: https://doi.org/10.1016/j.jss.2017.03.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0164121217300560.

[10] Stephanie, *Cohen's Kappa Statistic*, en-US, Dec. 2014. [Online]. Available: https://www.statisticshowto.com/cohens-kappa-statistic/ (visited on 04/12/2022).

[11] K. Pykes, *Cohen's Kappa*, en, Jan. 2021. [Online]. Available: https://towardsdatascience.com/cohens-kappa-9786ceceab58 (visited on 04/12/2022).

[12] S. Mcleod, “[Likert Scale Definition, Examples and Analysis]”, en, 2008. [Online]. Available: https://www.simplypsychology.org/likert-scale.html (visited on 04/29/2022).

[13] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness”, *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993. DOI: 10.1207/s15327108ijap0303_3. [Online]. Available: https://doi.org/10.1207/s15327108ijap0303_3.

[14] F. Naujoks, K. Wiedemann, N. Schömig, S. Hergeth, and A. Keinath, “Towards guidelines and verification methods for automated vehicle hmis”, *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 60, pp. 121–136, 2019, ISSN: 1369-8478. DOI: https://doi.org/10.1016/j.trf.2018.10.012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847818301748.

[15] A. Shorten and J. Smith, “Mixed methods research: Expanding the evidence base”, *Evidence-Based Nursing*, vol. 20, no. 3, pp. 74–75, 2017, ISSN: 1367-6539. DOI: 10.1136/eb-2017-102699. [Online]. Available: https://ebn.bmj.com/content/20/3/74.

[16] G. Meixner, C. Häcker, B. Decker, *et al.*, “Retrospective and Future Automotive Infotainment Systems—100 Years of User Interface Evolution”, en, in *Automotive User Interfaces: Creating Interactive Experiences in the Car*, ser. Human–Computer Interaction Series, G. Meixner and C. Müller, Eds., Cham: Springer International Publishing, 2017, pp. 3–53, ISBN: 9783319494487. DOI: 10.1007/978-3-319-49448-7_1. [Online]. Available: https://doi.org/10.1007/978-3-319-49448-7_1 (visited on 01/05/2022).

[17] *Automotive Software Architectures*, en. [Online]. Available: https://link.springer.com/book/10.1007/978-3-030-65939-4 (visited on 05/29/2022).

[18] *81 FR 87656 - Visual-Manual NHTSA Driver Distraction Guidelines for Portable and Aftermarket Devices*, en, publisher: Office of the Federal Register, National Archives and Records Administration, May 2016. [Online]. Available: https://www.govinfo.gov/app/details/FR-2016-12-05/2016-29051/summary (visited on 01/11/2022).

SLR References

1 Nora Broy et al. "3D Displays in Cars: Exploring the User Performance for a Stereoscopic Instrument Cluster". In: Association for Computing Machinery, 2014, pp. 1–9. ISBN: 9781450332125. DOI: 10.1145/2667317.2667319. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85116180685&origin=resultslist&sort=plf-f&src=s&st1=3D+Displays+in+Cars%3a+Exploring+the+User+Performance+for+a+Stereoscopic+Instrument+Cluster&sid=8777c155d69503613f718aa94012d8cd&sot=b&sdt=b&sl=104&s=TITLE-ABS-KEY%283D+Displays+in+Cars%3a+Exploring+the+User+Performance+for+a+Stereoscopic+Instrument+Cluster%29&relpos=0&citeCnt=12&searchTerm=

2 Q. Li, J. Wang, and T. Luo. "A Comparative Study of In-Car HMI Interaction Modes Based on User Experience". In: vol. 1421. 2021, pp. 257–261. ISBN: 9783030786441. DOI: 10.1007/978-3-030-78645-8_32. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85112050179&origin=resultslist&sort=plf-f&src=s&st1=A+Comparative+Study+of+In-Car+HMI+Interaction+Modes+Based+on+User+Experience&sid=0c3528c343401a2d76db1e5980dc2022&sot=b&sdt=b&sl=91&s=TITLE-ABS-KEY%28A+Comparative+Study+of+In-Car+HMI+Interaction+Modes+Based+on+User+Experience%29&relpos=0&citeCnt=0&searchTerm=

3 S Han, J Jeong, and S Seo. "A Methodology of UX Design of Touch-based Automotive Headunit by Minizing Driver's Distraction". In: 2018, pp. 276–280. DOI: 10.1109/ICITE.2018.8492566. URL: https://ieeexplore.ieee.org/document/8492566

4 M. Feld et al. "A model-based approach to the design and rendering of adaptive automotive user interfaces". In: vol. 220. 2013, pp. 2634–2648. ISBN: 9783885796145. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85083242770&origin=resultslist&sort=plf-f&src=s&st1=A+model-based+approach+to+the+design+and+rendering+of+adaptive+automotive+user+interfaces&sid=b168f4a5630e9f0a6425b9168a4b4089&sot=b&sdt=b&sl=104&s=TITLE-ABS-KEY%5C%28A+model-based+approach+to+the+design+and+rendering+of+adaptive+automotive+user+interfaces%5C%29&relpos=0&citeCnt=0&searchTerm=

5 C.A. Pickering, K.J. Burnham, and M.J. Richardson. "A review of automotive human machine interface technologies and techniques to reduce driver distraction". In: 2007, pp. 223–228. ISBN: 9780863418631. DOI: 10.1049/cp:20070468. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-67649823419&origin=resultslist&sort=tp-t&src=s&st1=A+review+of+automotive+human+machine+interface+technologies+and+techniques+to+reduce+driver+distraction&sid=f6d5ac2ff1a59d567ee87c0a6fbbbea7&sot=b&sdt=b&sl=118&s=TITLE-ABS-KEY%5C%28A+review+of+automotive+human+machine+interface+technologies+and+techniques+to+reduce+driver+distraction%5C%29&relpos=0&citeCnt=27&searchTerm=

6 X. Qian and Y. Sun. "A Study of Simplified Chinese Typeface Design applied to Automotive User Interface". In: 2019, pp. 23–29. ISBN: 9781450372473. DOI: 10.1145/3332169.3333572. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85069731152&origin=resultslist&sort=plf-f&src=s&st1=A+Study+of+Simplified+Chinese+Typeface+Design+applied+to+Automotive+User+Interface&sid=c0d6e4f9328b77d2e6c5eaf141121d4d&sot=b&sdt=b&sl=97&s=TITLE-ABS-KEY%28A+Study+of+Simplified+Chinese+Typeface+Design+applied+to+Automotive+User+Interface%29&relpos=1&citeCnt=0&searchTerm=

9 Simon Gansel et al. “An Access Control Concept for Novel Automotive HMI Systems”. In: Association for Computing Machinery, 2014, pp. 17–28. ISBN: 9781450329392. DOI: 10.1145/2613087.2613104. URL: https://doi.org/10.1145/2613087.2613104

10 N. Valverde et al. “An engineering perspective on the quality of the automotive push-buttons’ haptic feedback in optimal and suboptimal interactions”. In: *Journal of Engineering Design* 30 (8-9 2019), pp. 336–367. DOI: 10.1080/09544828.2019.1656802. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85071164732&origin=resultslist&sort=plf-f&src=s&st1=An+engineering+perspective+on+the+quality+of+the+automotive+push-buttons%e2%80%99+haptic+feedback+in+optimal+and+suboptimal+interactions&sid=bf84af88aa34aa6d572b9c815edeb956&sot=b&sdt=b&sl=143&s=TITLE-ABS-KEY%28An+engineering+perspective+on+the+quality+of+the+automotive+push-buttons%e2%80%99+haptic+feedback+in+optimal+and+suboptimal+interactions%29&relpos=0&citeCnt=0&searchTerm=

11 Bryan Reimer et al. “An Exploratory Study on the Impact of Typeface Design in a Text Rich User Interface on Off-Road Glance Behavior”. In: Association for Computing Machinery, 2012, pp. 25–32. ISBN: 9781450317511. DOI: 10.1145/2390256.2390260. URL: https://doi.org/10.1145/2390256.2390260

12 M. Bradley, P.M. Langdon, and P. John Clarkson. “An inclusive design perspective on automotive HMI trends”. In: vol. 9739. 2016, pp. 548–555. ISBN: 9783319402376. DOI: 10.1007/978-3-319-40238-3_52. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84978872477&origin=resultslist&sort=plf-f&src=s&st1=An+inclusive+design+perspective+on+automotive+HMI+trends&sid=7fb99ad8e0ed6b5c09e96ab67012c4ed&sot=b&sdt=b&sl=71&s=TITLE-ABS-KEY%28An+inclusive+design+perspective+on+automotive+HMI+trends%29&relpos=0&citeCnt=4&searchTerm=

13 Matthew J Pitts et al. “Assessing Subjective Response to Haptic Feedback in Automotive Touchscreens”. In: Association for Computing Machinery, 2009, pp. 11–18. ISBN: 9781605585710. DOI: 10.1145/1620509.1620512. URL: https://doi.org/10.1145/1620509.1620512

14 Michael Braun et al. “At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces”. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (2019), pp. 1–11. URL: https://doi.org/10.1145/3290605.3300270

15 H. Kim and E.C. Jung. “Automotive cluster UI design based on time and path from the perspective of the ’Passenger’”. In: *Archives of Design Research* 32 (3 2019), pp. 5–17. DOI: 10.15187/adr.2019.08.32.3.5. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85073592592&origin=resultslist&sort=plf-f&src=s&st1=Automotive+cluster+UI+design+based+on+time+and+path+from+the+perspective+of+the+%27Passenger%27&sid=9cf07a1218eb831be5112bec7d8002bd&sot=b&sdt=b&sl=106&s=TITLE-ABS-KEY%28Automotive+cluster+UI+design+based+on+time+and+path+from+the+perspective+of+the+%27Passenger%27%29&relpos=0&citeCnt=0&searchTerm=

16 D. Wilfinger et al. “Automotive HMI Test Package: An Exploitable Approach to Study In-Car HMIs”. In: vol. 2013-April. 2013, pp. 2377–2380. ISBN: 9781450318990. DOI: 10.1145/2468356.2468776. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84910623736&origin=resultslist&sort=plf-f&src=s&st1=Automotive+HMI+Test+Package%3a+An+

Exploitable+Approach+to+Study+In-Car+HMIs&sid=255b12e6e9e730637be23e344dabfa19&sot=b&sdt=b&sl=88&s=TITLE-ABS-KEY%28Automotive+HMI+Test+Package%3a+An+Exploitable+Approach+to+Study+In-Car+HMIs%29&relpos=0&citeCnt=1&searchTerm=

17 C.-X. Jin, F. You, and J.-M. Wang. "Automotive HUD interaction design based on lane changing Scenario". In: vol. 9733. 2016, pp. 277–284. ISBN: 9783319395128. DOI: 10.1007/978-3-319-39513-5_26. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84978842998&origin=resultslist&sort=plf-f&src=s&st1=Automotive+HUD+interaction+design+based+on+lane+changing+Scenario&sid=a2bbea53a607d4d8fcf9cd6209b535ea&sot=b&sdt=b&sl=80&s=TITLE-ABS-KEY%28Automotive+HUD+interaction+design+based+on+lane+changing+Scenario%29&relpos=1&citeCnt=0&searchTerm=

18 Michael Braun et al. "Automotive Research in the Public Space: Towards Deployment-Based Prototypes For Real Users". In: Association for Computing Machinery, 2018, pp. 181–185. ISBN: 9781450359474. DOI: 10.1145/3239092.3265964. URL: https://doi.org/10.1145/3239092.3265964

19 T. Wellings, M.J. Pitts, and M.A. Williams. "Characterising the experience of interaction: An evaluation of automotive rotary dials". In: *Ergonomics* 55 (11 2012), pp. 1298–1315. DOI: 10.1080/00140139.2012.708057. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84867841454&origin=resultslist&sort=plf-f&src=s&st1=Characterising+the+experience+of+interaction%5C%3a+An+evaluation+of+automotive+rotary+dials&sid=e76f997d64ce484791d495cbfbb15011&sot=b&sdt=b&sl=101&s=TITLE-ABS-KEY%5C%28Characterising+the+experience+of+interaction%5C%3a+An+evaluation+of+automotive+rotary+dials%3C%5C%29&relpos=0&citeCnt=5&searchTerm=

20 M. Kunur et al. *Creating inclusive HMI concepts for future cars using visual scenario storyboards through design ethnography.* Vol. 9178. 2015, pp. 139–149. ISBN: 9783319206868. DOI: 10.1007/978-3-319-20687-5_14. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84947251099&origin=resultslist&sort=plf-f&src=s&st1=Creating+inclusive+HMI+concepts+for+future+cars+using+visual+scenario+storyboards+through+design+ethnography&sid=ec9456015e536085c9dcfb2357c036be&sot=b&sdt=b&sl=123&s=TITLE-ABS-KEY%5C%28Creating+inclusive+HMI+concepts+for+future+cars+using+visual+scenario+storyboards+through+design+ethnography%5C%29&relpos=0&citeCnt=2&searchTerm=

21 T. Wellings, M.A. Williams, and M. Pitts. "Customer perception of switch-feel in luxury sports utility vehicles". In: *Food Quality and Preference* 19 (8 2008), pp. 737–746. DOI: 10.1016/j.foodqual.2008.03.004. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-49149110977&origin=resultslist&sort=tp-t&src=s&st1=Customer+perception+of+switch-feel+in+luxury+sports+utility+vehicles&sid=29220884f17212dd630253129aceef1f&sot=b&sdt=b&sl=83&s=TITLE-ABS-KEY%5C%28Customer+perception+of+switch-feel+in+luxury+sports+utility+vehicles%5C%29&relpos=0&citeCnt=25&searchTerm=

23 A Benmimoun, J Chen, and T Suzuki. "Design and Practical Evaluation of an Intersection Assistant in Real World Tests". In: 2007, pp. 606–611. ISBN: 1931-0587. DOI: 10.1109/IVS.2007.4290182. URL: https://ieeexplore.ieee.org/document/4290182

25 Y.-J. Kwon et al. "Design of Automotive Digital Instrument Cluster Adjustable to Driver's Cognitive Characteristics". In: 2019, pp. 461–463. ISBN: 9781728108926. DOI: 10.1109/ICTC46691.

2019.8939993. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85078274086&origin=resultslist&sort=plf-f&src=s&st1=Design+of+Automotive+Digital+Instrument+Cluster+Adjustable+to+Driver%5C%27s+Cognitive+Characteristics&sid=2026730cfba2ecf38791daba1cea1b46&sot=b&sdt=b&sl=111&s=TITLE-ABS-KEY%5C%28Design+of+Automotive+Digital+Instrument+Cluster+Adjustable+to+Driver%5C%27s+Cognitive+Characteristics%5C%29&relpos=0&citeCnt=1&searchTerm=

26 Alex de Ruiter and Miguel Bruns Alonso. "Designing Haptic Effects on an Accelerator Pedal to Support a Positive Eco-Driving Experience". In: Association for Computing Machinery, 2019, pp. 319–328. ISBN: 9781450368841. DOI: 10.1145/3342197.3344532. URL: https://doi.org/10.1145/3342197.3344532

27 Z. Wang, F. Bian, and P. Li. *Designing of vehicle operation and control based on semiotics.* Vol. 644-650. 2014, pp. 539–542. ISBN: 9783038352464. DOI: 10.4028/www.scientific.net/AMM.644-650.539. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84915751475&origin=resultslist&sort=plf-f&src=s&st1=Designing+of+vehicle+operation+and+control+based+on+semiotics&sid=82ecc6605396aa8d525a0be716856cea&sot=b&sdt=b&sl=76&s=TITLE-ABS-KEY%5C%28Designing+of+vehicle+operation+and+control+based+on+semiotics%5C%29&relpos=0&citeCnt=0&searchTerm=

28 C. Knoll, R. Vilimek, and I. Schulze. *Developing the HMI of electric vehicles on the necessity of a broader understanding of automotive user interface engineering.* Vol. 8519 LNCS. 2014, pp. 293–304. ISBN: 9783319076348. DOI: 10.1007/978-3-319-07635-5_29. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84958543950&origin=resultslist&sort=plf-f&src=s&st1=Developing+the+HMI+of+electric+vehicles+on+the+necessity+of+a+broader+understanding+of+automotive+user+interface+engineering&sid=579e2dd9da1906e3429e2540b9cd0266&sot=b&sdt=b&sl=139&s=TITLE-ABS-KEY%5C%28Developing+the+HMI+of+electric+vehicles+on+the+necessity+of+a+broader+understanding+of+automotive+user+interface+engineering%5C%29&relpos=0&citeCnt=3&searchTerm=

29 C. Kataoka et al. "Development and effectiveness evaluation of interactive voice HMI system". In: *IEICE Transactions on Information and Systems* E104.D (4 2021), pp. 500–507. DOI: 10.1587/TRANSINF.2020EDP7068. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85105077426&origin=resultslist&sort=plf-f&src=s&st1=Development+and+effectiveness+evaluation+of+interactive+voice+HMI+system&sid=a400a44c8921303e099b549b505
sot=b&sdt=b&sl=87&s=TITLE-ABS-KEY%5C%28Development+and+effectiveness+evaluation+of+interactive+voice+HMI+system%5C%29&relpos=0&citeCnt=0&searchTerm=

30 B. Schwambach et al. "Development and usability of a novel holographic 3-dimensional user interface using hand-gesture control for a future vehicle application by older adults". In: vol. 2. 2018, pp. 992–996. ISBN: 9781510889538. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85072751948&origin=resultslist&sort=plf-f&src=s&st1=Development+and+usability+of+a+novel+holographic+3-dimensional+user+interface+using+hand-gesture+control+for+a+future+vehicle+application+by+older+adults&sid=73ceda2378e711ec580f5615a0614d8a&sot=b&sdt=b&sl=168&s=TITLE-ABS-KEY%5C%28Development+and+usability+of+a+novel+holographic+3-dimensional+user+interface+

using+hand-gesture+control+for+a+future+vehicle+application+by+older+adults%5C%29&relpos=0&citeCnt=0&searchTerm=

31 M. Liu et al. "Development of curved in cell TFT-LCD module for automotive display". In: vol. 49. 2018, pp. 1339–1342. DOI: 10.1002/sdtp.12161. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85054033086&origin=resultslist&sort=plf-f&src=s&st1=Development+of+curved+in+cell+TFT-LCD+module+for+automotive+display&sid=edc0459525fd22e1f9e6a2e2e4800fa0&sot=b&sdt=b&sl=82&s=TITLE-ABS-KEY%5C%28Development+of+curved+in+cell+TFT-LCD+module+for+automotive+display%5C%29&relpos=0&citeCnt=1&searchTerm=

34 S M Sarala, D H Sharath Yadav, and A Ansari. "Emotionally Adaptive Driver Voice Alert System for Advanced Driver Assistance System (ADAS) Applications". In: 2018, pp. 509–512. DOI: 10.1109/ICSSIT.2018.8748541. URL: https://ieeexplore.ieee.org/document/8748541

35 Marie-Christin Ostendorp et al. "Engineering Automotive HMIs That Are Optimized for Correct and Fast Perception". In: Association for Computing Machinery, 2016, pp. 293–298. ISBN: 9781450343220. DOI: 10.1145/2933242.2935869. URL: https://dl.acm.org/doi/10.1145/2933242.2935869

36 B Wortelen et al. "Engineering Optimized Control Panel Designs for Electric Driving". In: 2019, pp. 125–130. ISBN: 2380-7350. DOI: 10.1109/CogInfoCom47531.2019.9089913. URL: https://ieeexplore.ieee.org/document/9089913

37 M Ganzhorn, J P F Diederichs, and H Widlroither. "Enhancing the reproduction of road signs in automotive HMIs to protect children". In: 2011, pp. 338–344. DOI: 10.1109/ITST.2011.6060078. URL: https://ieeexplore.ieee.org/document/6060078

39 D.R. Large et al. "Evaluating secondary input devices to support an automotive touchscreen HMI: A cross-cultural simulator study conducted in the UK and China". In: *Applied Ergonomics* 78 (2019), pp. 184–196. DOI: 10.1016/j.apergo.2019.03.005. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85063115413&origin=resultslist&sort=plf-f&src=s&st1=Evaluating+secondary+input+devices+to+support+an+automotive+touchscreen+HMI%5C%3a+A+cross-cultural+simulator+study+conducted+in+the+UK+and+China&sid=740e1e9c5ebf6c6e6f46c4e2f02bf030&sot=b&sdt=b&sl=154&s=TITLE-ABS-KEY%5C%28Evaluating+secondary+input+devices+to+support+an+automotive+touchscreen+HMI%5C%3a+A+cross-cultural+simulator+study+conducted+in+the+UK+and+China%5C%29&relpos=0&citeCnt=6&searchTerm=

40 P. George et al. "Evaluation of smartphone-based interaction techniques in a CAVE in the context of immersive digital project review". In: vol. 9012. 2014. ISBN: 9780819499295. DOI: 10.1117/12.2038237. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84901801545&origin=resultslist&sort=plf-f&src=s&st1=Evaluation+of+smartphone-based+interaction+techniques+in+a+CAVE+in+the+context+of+immersive+digital+project+review&sid=870b79546fb1766247de1b178249e369&sot=b&sdt=b&sl=129&s=TITLE-ABS-KEY%5C%28Evaluation+of+smartphone-based+interaction+techniques+in+a+CAVE+in+the+context+of+immersive+digital+project+review%5C%29&relpos=0&citeCnt=2&searchTerm=

42 J P F Diederichs, B Weber, and M Ganzhorn. “Experimental evaluation of ITST based school bus stop notifications for car drivers”. In: 2011, pp. 345–350. DOI: 10.1109/ITST.2011.6060080. URL: https://ieeexplore.ieee.org/document/6060080

43 N. Broy et al. “Exploring design parameters for a 3D head-up display”. In: 2014, pp. 38–43. ISBN: 9781450329521. DOI: 10.1145/2611009.2611011. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84904479911&origin=resultslist&sort=plf-f&src=s&st1=Exploring+design+parameters+for+a+3D+head-up+display&sid=274b3f06458f78d6cf515c13522f204 sot=b&sdt=b&sl=67&s=TITLE-ABS-KEY%5C%28Exploring+design+parameters+for+a+3D+head-up+display%5C%29&relpos=0&citeCnt=16&searchTerm=

45 Hyungil Kim et al. “Exploring Head-up Augmented Reality Interfaces for Crash Warning Systems”. In: Association for Computing Machinery, 2013, pp. 224–227. ISBN: 9781450324786. DOI: 10.1145/2516540.2516566. URL: https://doi.org/10.1145/2516540.2516566

47 Burkay Sucu and Eelke Folmer. “Haptic Interface for Non-Visual Steering”. In: Association for Computing Machinery, 2013, pp. 427–434. ISBN: 9781450319652. DOI: 10.1145/2449396.2449451. URL: https://doi.org/10.1145/2449396.2449451

48 Sebastian Osswald, Pratik Sheth, and Manfred Tscheligi. “Hardware-in-the-Loop-Based Evaluation Platform for Automotive Instrument Cluster Development (EPIC)”. in: Association for Computing Machinery, 2013, pp. 323–332. ISBN: 9781450321389. DOI: 10.1145/2494603.2480305. URL: https://doi.org/10.1145/2494603.2480305

49 A Degani, A Gellatly, and M Heymann. “HMI aspects of automotive climate control systems”. In: 2011, pp. 1795–1800. ISBN: 1062-922X. DOI: 10.1109/ICSMC.2011.6083932. URL: https://ieeexplore.ieee.org/document/6083932

51 V. Charissis and S. Papanastasiou. “Human-machine collaboration through vehicle head up display interface”. In: *Cognition, Technology and Work* 12 (1 2010), pp. 41–50. DOI: 10.1007/s10111-008-0117-0. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-76349097816&origin=resultslist&sort=plf-f&src=s&st1=Human-machine+collaboration+through+vehicle+head+up+display+interface&sid=433232edd7cbb475971676f2c57d8c1e&sot=b&sdt=b&sl=84&s=TITLE-ABS-KEY%5C%28Human-machine+collaboration+through+vehicle+head+up+display+interface%5C%29&relpos=0&citeCnt=66&searchTerm=

52 Verena Broy, Frank Althoff, and Gudrun Klinker. “IFlip: A Metaphor for in-Vehicle Information Systems”. In: Association for Computing Machinery, 2006, pp. 155–158. ISBN: 1595933530. DOI: 10.1145/1133265.1133297. URL: https://doi.org/10.1145/1133265.1133297

54 C. Jeong et al. “In-vehicle display HMI safety evaluation using a driving simulator”. In: *International Journal of Automotive Technology* 14 (6 2013), pp. 987–992. DOI: 10.1007/s12239-013-0108-x. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84887448916&origin=resultslist&sort=plf-f&src=s&st1=In-vehicle+display+HMI+safety+evaluation+using+a+driving+simulator&sid=da0bd1c1a0adc2328562787925ee3e7b&sot=b&sdt=b&sl=81&s=TITLE-ABS-KEY%5C%28In-vehicle+display+HMI+safety+evaluation+using+a+driving+simulator%5C%29&relpos=0&citeCnt=9&searchTerm=

55 Q. Zeng, B. Jiang, and Q. Duan. “Integrated evaluation of hardware and software interfaces for automotive human-machine interaction”. In: *IET Cyber-Physical Systems: Theory and Applications* 4 (3 2019), pp. 214–220. DOI: 10.1049/iet-cps.2019.0002. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85073163451&origin=resultslist&sort=plf-f&src=s&st1=Integrated+evaluation+of+hardware+and+software+interfaces+for+automotive+human-machine+interaction&sid=8a2dcb9df0ece7baab4e33b1b386ee6f&sot=b&sdt=b&sl=113&s=TITLE-ABS-KEY%5C%28Integrated+evaluation+of+hardware+and+software+interfaces+for+automotive+human-machine+interaction%5C%29&relpos=0&citeCnt=6&searchTerm=

57 Andrew W Gellatly et al. “Journey: General Motors’ Move to Incorporate Contextual Design into Its next Generation of Automotive HMI Designs”. In: Association for Computing Machinery, 2010, pp. 156–161. ISBN: 9781450304375. DOI: 10.1145/1969773.1969802. URL: https://doi.org/10.1145/1969773.1969802

58 J. Vieira et al. “Kansei engineering as a tool for the design of in-vehicle rubber keypads”. In: *Applied Ergonomics* 61 (2017), pp. 1–11. DOI: 10.1016/j.apergo.2016.12.019. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85008323130&origin=resultslist&sort=plf-f&src=s&st1=Kansei+engineering+as+a+tool+for+the+design+of+in-vehicle+rubber+keypads&sid=bbdb0d1b3b407bdc82f606743cf1aae7&sot=b&sdt=b&sl=87&s=TITLE-ABS-KEY%5C%28Kansei+engineering+as+a+tool+for+the+design+of+in-vehicle+rubber+keypads%5C%29&relpos=1&citeCnt=40&searchTerm=

59 P Fröhlich et al. “Live traffic camera content in the connected car: Experiences from motorway experiments”. In: 2013, pp. 396–403. ISBN: 2378-1297. DOI: 10.1109/ICCVE.2013.6799826. URL: https://ieeexplore.ieee.org/document/6799826

60 S.H. Jeong et al. “Low cost design of parallel parking assist system based on an ultrasonic sensor”. In: *International Journal of Automotive Technology* 11 (3 2010), pp. 409–416. DOI: 10.1007/s12239-010-0050-0. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-77953037216&origin=resultslist&sort=plf-f&src=s&st1=Low+cost+design+of+parallel+parking+assist+system+based+on+an+ultrasonic+sensor&sid=ceae0f2452d75b811c02b4482485d944&sot=b&sdt=b&sl=94&s=TITLE-ABS-KEY%5C%28Low+cost+design+of+parallel+parking+assist+system+based+on+an+ultrasonic+sensor%5C%29&relpos=0&citeCnt=39&searchTerm=

61 P Clement et al. “Measuring trust in automated driving using a multi-level approach to human factors¡sup¿*¡/sup¿”. In: 2021, pp. 410–417. DOI: 10.1109/DSD53832.2021.00069. URL: https://ieeexplore.ieee.org/document/9556399

62 Florian Schaub et al. “On Credibility Improvements for Automotive Navigation Systems”. In: *Personal Ubiquitous Comput.* 17 (5 June 2013), pp. 803–813. ISSN: 1617-4909. DOI: 10.1007/s00779-012-0519-0. URL: https://doi.org/10.1007/s00779-012-0519-0

63 K. Blankenbach et al. “Optimization and evaluation of automotive displays under bright ambient light using novel image enhancement algorithms”. In: vol. 45. 2014, pp. 595–598. DOI: 10.1002/j.2168-0159.2014.tb00156.x. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84905029877&origin=resultslist&sort=plf-f&src=s&st1=Optimization+and+evaluation+of+automotive+displays+under+bright+ambient+light+using+novel+image+

enhancement+algorithms&sid=e97a362d85db0b9bac36af0e5c6b4e7f&sot=b&sdt=b&sl=133&s=TITLE-ABS-KEY%5C%28Optimization+and+evaluation+of+automotive+displays+under+bright+ambient+light+using+novel+image+enhancement+algorithms%5C%29&relpos=0&citeCnt=0&searchTerm=

64 K. Blankenbach et al. "Optimizing and evaluating new automotive HMI image enhancement algorithms under bright light conditions using display reflectance characteristics". In: *Journal of the Society for Information Display* 22 (5 2014), pp. 267–279. DOI: 10.1002/jsid.246. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84923183674&origin=resultslist&sort=plf-f&src=s&st1=Optimizing+and+evaluating+new+automotive+HMI+image+enhancement+algorithms+under+bright+light+conditions+using+display+reflectance+characteristics&sid=9403e36c2676333cc094d06d924caa3c&sot=b&sdt=b&sl=160&s=TITLE-ABS-KEY%5C%28Optimizing+and+evaluating+new+automotive+HMI+image+enhancement+algorithms+under+bright+light+conditions+using+display+reflectance+characteristics%5C%29&relpos=0&citeCnt=10&searchTerm=

66 T. Helldin et al. "Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving". In: 2013, pp. 210–217. ISBN: 9781450324786. DOI: 10.1145/2516540.2516554. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84888184123&origin=resultslist&sort=plf-f&src=s&st1=Presenting+system+uncertainty+in+automotive+UIs+for+supporting+trust+calibration+in+autonomous+driving&sid=1d105dd5338947fe53541b86c4cc8ae1&sot=b&sdt=b&sl=117&s=TITLE-ABS-KEY%5C%28Presenting+system+uncertainty+in+automotive+UIs+for+supporting+trust+calibration+in+autonomous+driving%5C%29&relpos=0&citeCnt=145&searchTerm=

67 T Matsuura and K Sato. "Proposal of next-generation HMI utilizing context and personality theories". In: 2018, pp. 1–3. ISBN: 2158-4001. DOI: 10.1109/ICCE.2018.8326154. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85048831744&origin=resultslist&sort=plf-f&src=s&st1=Proposal+of+next-generation+HMI+utilizing+context+and+personality+theories&sid=610969846eb6ad41b309a42c5cc86ff3&sot=b&sdt=b&sl=89&s=TITLE-ABS-KEY%5C%28Proposal+of+next-generation+HMI+utilizing+context+and+personality+theories%5C%29&relpos=0&citeCnt=0&searchTerm=

68 S Nakajima et al. "Proposal of reduction method of Mixed Reality sickness using auditory stimuli for advanced driver assistance systems". In: 2009, pp. 1–5. DOI: 10.1109/ICIT.2009.4939696. URL: https://ieeexplore.ieee.org/document/4939696

69 A O Yontem et al. "Prospective Immersive Human-Machine Interface for Future Vehicles: Multiple Zones Turn the Full Windscreen Into a Head-Up Display". In: *IEEE Vehicular Technology Magazine* 16 (1 2021), pp. 83–92. ISSN: 1556-6080. DOI: 10.1109/MVT.2020.3013832. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85090239428&origin=resultslist&sort=plf-f&src=s&st1=Prospective+Immersive+Human-Machine+Interface+for+Future+Vehicles%5C%3a+Multiple+Zones+Turn+the+Full+Windscreen+Into+a+Head-Up+Display&sid=6a7e9977a585880f6cfa4826da2fad62&sot=b&sdt=b&sl=144&s=TITLE-ABS-KEY%5C%28Prospective+Immersive+Human-Machine+Interface+for+Future+Vehicles%5C%3a+Multiple+Zones+Turn+the+Full+Windscreen+Into+a+Head-Up+Display%5C%29&relpos=0&citeCnt=5&searchTerm=

70 B.I. Ahmad et al. "Selection facilitation schemes for predictive touch with mid-air pointing gestures in automotive displays". In: 2018, pp. 21–32. ISBN: 9781450359467. DOI: 10.1145/3239060.3239067. URL: https://dl.acm.org/doi/10.1145/3239060.3239067

72 J Chen et al. "Study on Safety Analysis Method for Take-over System of Autonomous Vehicles**Research supported by the National Key Research and Development Program of China (Grant No. 2018YFB0105101; 2018YFB0105103)." In: 2020, pp. 1972–1977. ISBN: 2642-7214. DOI: 10.1109/IV47402.2020.9304599. URL: https://ieeexplore.ieee.org/document/9304599

75 V. Melcher et al. "Take-Over Requests for Automated Driving". In: *Procedia Manufacturing* 3 (2015), pp. 2867–2873. DOI: 10.1016/j.promfg.2015.07.788. URL: https://doi.org/10.1016/j.promfg.2015.07.788

77 J. Dobres, B. Reimer, and N. Chahine. "The effect of font weight and rendering system on glance-based text legibility". In: 2016, pp. 91–96. ISBN: 9781450345330. DOI: 10.1145/3003715.3005454. URL: https://dl.acm.org/doi/10.1145/3003715.3005454

78 T. Hecht, S. Kratzert, and K. Bengler. "The effects of a predictive HMI and different transition frequencies on acceptance, workload, usability, and gaze behavior during urban automated driving". In: *Information (Switzerland)* 11 (2 2020). DOI: 10.3390/info11020073. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85081136448&origin=resultslist&sort=plf-f&src=s&st1=The+effects+of+a+predictive+HMI+and+different+transition+frequencies+on+acceptance%5C%2c+workload%5C%2c+usability%5C%2c+and+gaze+behavior+during+urban+automated+driving&sid=6a4d7350ce97c09698bedd2f624d5462&sot=b&sdt=b&sl=168&s=TITLE-ABS-KEY%5C%28The+effects+of+a+predictive+HMI+and+different+transition+frequencies+on+acceptance%5C%2c+workload%5C%2c+usability%5C%2c+and+gaze+behavior+during+urban+automated+driving%5C%29&relpos=0&citeCnt=5&searchTerm=

80 B. Niedermaier et al. *The new BMW iDrive - Applied processes and methods to assure high usability*. Vol. 5620 LNCS. 2009, pp. 443–452. ISBN: 9783642028083. DOI: 10.1007/978-3-642-02809-0_47. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-70350335721&origin=resultslist&sort=r-f&src=s&st1=The+new+BMW+iDrive+-+Applied+processes+and+methods+to+assure+high+usability&sid=b28699f8a99f14d0368ee0b15b39672f&sot=b&sdt=b&sl=80&s=ALL%5C%28The+new+BMW+iDrive+-+Applied+processes+and+methods+to+assure+high+usability%5C%29&relpos=0&citeCnt=15&searchTerm=

81 S. Rodriguez Garzon and M. Poguntke. *The personal adaptive in-car HMI: Integration of external applications for personalized use*. Vol. 7138 LNCS. 2012, pp. 35–46. ISBN: 9783642285080. DOI: 10.1007/978-3-642-28509-7_5. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-84857524006&origin=resultslist&sort=plf-f&src=s&st1=The+personal+adaptive+in-car+HMI%5C%3a+Integration+of+external+applications+for+personalized+use&sid=c7e25eb79fd4ccda606d7ea5db3ddf5d&sot=b&sdt=b&sl=106&s=TITLE-ABS-KEY%5C%28The+personal+adaptive+in-car+HMI%5C%3a+Integration+of+external+applications+for+personalized+use%5C%29&relpos=1&citeCnt=12&searchTerm=

82 M. Walter, V. Seitz, and K. Bengler. *The virtual penetrating the physical and the implication for augmented reality head-up displays*. Vol. 795. 2019, pp. 3–10. ISBN: 9783319946184. DOI:

10.1007/978-3-319-94619-1_1. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85049358733&origin=resultslist&sort=plf-f&src=s&st1=The+virtual+penetrating+the+physical+and+the+implication+for+augmented+reality+head-up+displays&sid=b78a90d840db66bb95ead2a296a9942d&sot=b&sdt=b&sl=110&s=TITLE-ABS-KEY%5C%28The+virtual+penetrating+the+physical+and+the+implication+for+augmented+reality+head-up+displays%5C%29&relpos=0&citeCnt=0&searchTerm=

83 Sebastian Feuerstack et al. "Theater-System Technique and Model-Based Attention Prediction for the Early Automotive HMI Design Evaluation". In: Association for Computing Machinery, 2016, pp. 19–22. ISBN: 9781450345330. DOI: 10.1145/3003715.3005466. URL: https://doi.org/10.1145/3003715.3005466

86 A. Amditis et al. "Towards the automotive HMI of the future: Overview of the AIDE - Integrated project results". In: *IEEE Transactions on Intelligent Transportation Systems* 11 (3 2010), pp. 567–578. DOI: 10.1109/TITS.2010.2048751. URL: https://ieeexplore-ieee-org.zorac.aub.aau.dk/document/5504222

87 David R Large et al. "Twist It, Touch It, Push It, Swipe It: Evaluating Secondary Input Devices for Use with an Automotive Touchscreen HMI". in: Association for Computing Machinery, 2016, pp. 161–168. ISBN: 9781450345330. DOI: 10.1145/3003715.3005459. URL: https://doi.org/10.1145/3003715.3005459

88 M. Lee and Y. Lee. *UI proposal for shared autonomous vehicles: Focusing on improving user's trust*. Vol. 12213 LNCS. 2020, pp. 282–296. ISBN: 9783030505363. DOI: 10.1007/978-3-030-50537-0_21. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85088900227&origin=resultslist&sort=plf-f&src=s&st1=UI+proposal+for+shared+autonomous+vehicles%5C%3a+Focusing+on+improving+user%5C%e2%5C%80%5C%99s+trust&sid=c0e83a9ecdf9d832a10fd6f609cce9bb&sot=b&sdt=b&sl=93&s=TITLE-ABS-KEY%5C%28UI+proposal+for+shared+autonomous+vehicles%5C%3a+Focusing+on+improving+user%5C%e2%5C%80%5C%99s+trust%5C%29&relpos=1&citeCnt=0&searchTerm=

89 T. Wellings, M. Williams, and C. Tennant. "Understanding customers' holistic perception of switches in automotive human-machine interfaces". In: *Applied Ergonomics* 41 (1 2010), pp. 8–17. DOI: 10.1016/j.apergo.2009.03.004. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-70349778819&origin=resultslist&sort=plf-f&src=s&st1=Understanding+customers%5C%27+holistic+perception+of+switches+in+automotive+human-machine+interfaces&sid=c3624b21217a6f631fb2971de3b50505&sot=b&sdt=b&sl=110&s=TITLE-ABS-KEY%5C%28Understanding+customers%5C%27+holistic+perception+of+switches+in+automotive+human-machine+interfaces%5C%29&relpos=0&citeCnt=33&searchTerm=

90 SeungJun Kim et al. "Usability of Car Dashboard Displays for Elder Drivers". In: Association for Computing Machinery, 2011, pp. 493–502. ISBN: 9781450302289. DOI: 10.1145/1978942.1979011. URL: https://doi.org/10.1145/1978942.1979011

91 Henrik Detjen et al. "User-Defined Voice and Mid-Air Gesture Commands for Maneuver-Based Interventions in Automated Vehicles". In: Association for Computing Machinery, 2019, pp. 341–348. ISBN: 9781450371988. DOI: 10.1145/3340764.3340798. URL: https://doi.org/10.1145/3340764.3340798

92 F. Wehr and M. Luccarelli. "Using personas in the design process. Towards the development of green product personality for in-car user interfaces". In: vol. 2019-Augus. 2019, pp. 2911–2920. DOI: 10.1017/dsi.2019.298. URL: https://www.scopus.com/record/display.uri?eid=2-s2.0-85079751468&origin=resultslist&sort=plf-f&src=s&st1=Using+personas+in+the+design+process.+Towards+the+development+of+green+product+personality+for+in-car+user+interfaces&sid=d1d08d72b909229f052bfc485505ec3e&sot=b&sdt=b&sl=132&s=TITLE-ABS-KEY%5C%28Using+personas+in+the+design+process.+Towards+the+development+of+green+product+personality+for+in-car+user+interfaces%5C%29&relpos=0&citeCnt=0&searchTerm=

93 Michael Braun et al. "What If Your Car Would Care? Exploring Use Cases For Affective Automotive User Interfaces". In: Association for Computing Machinery, 2020. ISBN: 9781450375160. DOI: 10.1145/3379503.3403530. URL: https://doi.org/10.1145/3379503.3403530

94 Amnon Shashua, Itay Gat, and Meny Benady. "A Monocular Vision Advance Warning System for the Automotive Aftermarket". In: *SAE 2005 World Congress & Exhibition*. SAE International, Apr. 2005. DOI: https://doi.org/10.4271/2005-01-1470. URL: https://doi.org/10.4271/2005-01-1470

95 Kazuya Itoh et al. "A Study on Estimating the Variation of Driver's State by EEGs and EOGs". In: *SAE 2006 World Congress & Exhibition*. SAE International, Apr. 2006. DOI: https://doi.org/10.4271/2006-01-0575. URL: https://doi.org/10.4271/2006-01-0575

96 Kazutoshi Ebe et al. "Additional Findings on the Multi-Modal Demands of "Voice-Command" Interfaces". In: *SAE 2016 World Congress and Exhibition*. SAE International, Apr. 2016. DOI: https://doi.org/10.4271/2016-01-1428. URL: https://doi.org/10.4271/2016-01-1428

97 Louis Tijerina et al. "An Examination of Driver Eye Glance Behavior, Navigational Errors, and Subjective Assessments While Using In-Vehicle Navigational Systems With and Without Landmark Enhancements". In: *WCX™ 17: SAE World Congress Experience*. SAE International, Mar. 2017. DOI: https://doi.org/10.4271/2017-01-1375. URL: https://doi.org/10.4271/2017-01-1375

99 Linda Angell et al. "Bench-Marking Drivers' Visual and Cognitive Demands: A Feasibility Study". In: *SAE International Journal of Passenger Cars - Mechanical Systems* 8.2 (Apr. 2015), pp. 584–593. ISSN: 1946-3995. DOI: https://doi.org/10.4271/2015-01-1389. URL: https://doi.org/10.4271/2015-01-1389

100 Dean P. Chiang, Aaron M. Brooks, and David H. Weir. "Comparison of Visual-Manual and Voice Interaction With Contemporary Navigation System HMIs". In: *SAE 2005 World Congress & Exhibition*. SAE International, Apr. 2005. DOI: https://doi.org/10.4271/2005-01-0433. URL: https://doi.org/10.4271/2005-01-0433

101 Naoki Kamiya et al. "Development and Evaluation of a Multi-Functional Steering Wheel Switch". In: *SAE 2005 World Congress & Exhibition*. SAE International, Apr. 2005. DOI: https://doi.org/10.4271/2005-01-0424. URL: https://doi.org/10.4271/2005-01-0424

102 Junichi Yukawa et al. "Development of the HMI System That Improves the Safety and Operational Convenience by the Combination of Steering Wheel Switch and Head Up Display". In: *SAE 2005*

World Congress & Exhibition. SAE International, Apr. 2005. DOI: https://doi.org/10.4271/2005-01-0438. URL: https://doi.org/10.4271/2005-01-0438

103 Hiroaki Kosaka et al. "Evaluation of a New In-vehicle HMI System Composed of Steering Wheel Switch and Head-Up Display when a Driver Encounters Sudden Danger". In: *SAE World Congress & Exhibition*. SAE International, Apr. 2007. DOI: https://doi.org/10.4271/2007-01-0439. URL: https://doi.org/10.4271/2007-01-0439

104 Chengshi Wang, Yue Wang, and John R. Wagner. "Evaluation of a Robust Haptic Interface for Semi-Autonomous Vehicles". In: *SAE International Journal of Connected and Automated Vehicles* 2.2 (May 2019), pp. 99–114. ISSN: 2574-0741. DOI: https://doi.org/10.4271/12-02-02-0007. URL: https://doi.org/10.4271/12-02-02-0007

105 Wei Wang, Hongnan Lin, and Pranav Nair. "Lookie Here! Designing Directional User Indicators across Displays in Conditional Driving Automation". In: *WCX SAE World Congress Experience*. SAE International, Apr. 2020. DOI: https://doi.org/10.4271/2020-01-1201. URL: https://doi.org/10.4271/2020-01-1201

106 Paul F. Smith, Jie Chen, and Hongxing Hu. "Model-Based Design Study and Evaluation of New HMI Concepts for Vehicle Multimedia, Climate Control and Navigation Systems". In: *SAE World Congress & Exhibition*. SAE International, Apr. 2007. DOI: https://doi.org/10.4271/2007-01-0702. URL: https://doi.org/10.4271/2007-01-0702

107 Richard Young, Li Hsieh, and Sean Seaman. "The Dimensional Model of Driver Demand: Extension to Auditory-Vocal and Mixed-Mode Tasks". In: *SAE International Journal of Transportation Safety* 4.1 (Apr. 2016), pp. 72–106. ISSN: 2327-5626. DOI: https://doi.org/10.4271/2016-01-1427. URL: https://doi.org/10.4271/2016-01-1427

108 Jonathan Dobres et al. "The Influence of Driver's Age on Glance Allocations during Single-Task Driving and Voice vs. Visual-Manual Radio Tuning". In: *SAE 2016 World Congress and Exhibition*. SAE International, Apr. 2016. DOI: https://doi.org/10.4271/2016-01-1445. URL: https://doi.org/10.4271/2016-01-1445

109 Colleen Serafin et al. "User Experience in the U.S. and Germany of In-Vehicle Touch Screens with Integrated Haptic and Auditory Feedback". In: *SAE World Congress & Exhibition*. SAE International, Apr. 2007. DOI: https://doi.org/10.4271/2007-01-0451. URL: https://doi.org/10.4271/2007-01-0451

The Effects of Texting While Driving: Voice-To-Text Interaction with a HUD

Jeppe Skovby Bjørn
Department of Computer Science
Aalborg University
Selma Lagerløfs Vej 300
Aalborg, Denmark
jbjarn17@student.aau.dk

Ronnie Svendsen
Department of Computer Science
Aalborg University
Selma Lagerløfs Vej 300
Aalborg, Denmark
rsvend17@student.aau.dk

ABSTRACT

The use of handheld mobile phone while driving has been illegal in Denmark since 1998. Still, every tenth driver in Denmark has written a text message while driving. The objective of the study, is to investigate the possible effects of texting while driving using a HUD with voice-to-text interaction compared to a mobile phone in a simulator. A first low fidelity prototype was created using existing state-of-the-art HUDs from Audi, BMW, and Nissan. The prototype was evaluated with 8 participants and the design was corrected according to their feedback. 16 participants were invited to drive the simulator with no traffic congestion other than a leading vehicle, while receiving and answering text messages using a mobile phone and the designed HUD. Results show that using a HUD enables the driver to maintain their lane better than when they used a mobile phone. The participants also wrote longer messages when replying through the HUD.

KEYWORDS

Head-Up display; Human-Machine Interaction; Automotive Engineering; Automotive; Simulator

1 INTRODUCTION

The usage of handheld mobile phones while driving has been illegal in Denmark since 1998 and is also illegal in 24 states in the USA[1], [2]. However, the use of handheld mobile phones is still a problem today. Furthermore, in Denmark, every tenth driver in 2018 has written a text message on the phone while driving[3]. Mobile-phone-related crashes in the USA have seen no reduction from 2016 to 2019. Furthermore, 1/3 of all mobile-phone-related accidents in the USA are made by people aged 20-29. [4]–[6]

A possible solution to mobile-phone-related accidents and texting while driving can be seen in [7]. The paper explored the effects a **H**ead-**U**p **D**isplays(HUD) with a 3x4 keypad mounted on the steering wheel had on driver performance. The HUD showed improvements to some driver performance variables, such as lane crossings and glance period away from the road, when comparing it to a QWERTY display. The paper also explored the possibility of using voice-to-text, however, at the moment of the article they found them unreliable and inefficient. However, since this research, voice assistant systems have seen exponential growth in understanding the users and what they want. [8], [9]

A HUD displays information on the windshield in the line of vision from the driver. More traditional displays, also called **H**ead-**D**own **D**isplays(HDD), requires the driver to look down and away from the traffic environment. An HDD will most likely be placed in the centre counsel together with the radio and climate controls. [10]

This paper explores how voice-to-text-based interaction with a HUD during simulated driving affects the driver's performance.

This paper will first design mock-ups of a HUD. Different elements which affect the readability and usability will be explored to single out a recommended design for future HUDs used for text-based communication. These mock-ups are then presented to users to gather qualitative data about the presented mock-ups. The conducted experiment, with the refined mock-up, will then be described. Lastly, the results of the experiment will be presented and discussed.

2 DESIGNING THE HUD

This sections describes the method used to for gathering quantitative and qualitative data about the participants and the presented designs.

2.1 Prototypes

The first draft created can be seen in Figure 1. It was based on HUD designs from Audi, BMW, and Nissan. They all shared similar design features like speedometer and GPS. The HUD presented in a high-end BMW included a large additional section containing information regarding the condition of the vehicle. This area will be used for text-based communication in the designed HUD.

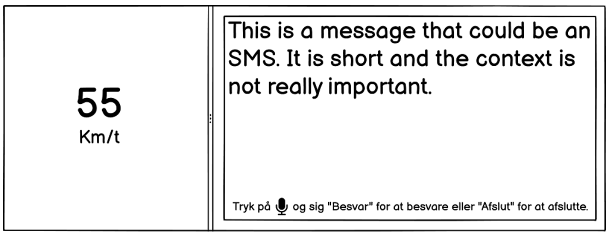

Figure 1: First draft for the HUD

Figure 1 represents the overall blueprint for the HUD. This contained the necessary elements and the general placement of there. The speed will be on the left side of the HUD and the text-based communication will be on the right side of the HUD. The interaction options are placed in the bottom part of the text. The early prototypes had a coloured background to emphasize options the user would have to interact with the HUD. The refined prototypes can be seen in Figures 2-3. The prototypes had three different variants with three different background transparencies.

Figure 2: Refined version of the HUD with all information present

Figure 3: Refined version of the HUD with limited information present

The main difference between Figure 2 and Figure 3 is the amount of information presented in the HUD and the guiding text in the lower part of the HUD. The first type only had the most necessary information(Who sent the message) for the user to understand that a text message had been received(not displaying the text itself) and the other contained everything. Figure 2 displays the speed on the left and the received message on the right. The two commands "Besvar"(Answer) and "Afslut"(Cancel) is supported by a guiding word "Kommando"(Command) to emphasize that these are voice commands for interaction. Figure 3 displays the speed on the left, the name of the sender on the right, and the commands at the bottom have been trimmed down a bit. Figure 2-3 was shown to the participants, and based on their feedback a final prototype was created to be implemented in the simulator used for the experiment.

2.2 Participant

Eight participants evaluated the prototypes. The participants were distributed evenly between males and females. The average age of the participants was 25.63 (SD=2.62). The driving experience of the participants ranged from 0-3 years(n=1), 5-7 years (n=3), 7-10 years(n=3), and 10+ years(n=1) of experience.

Table 1: Demographic Data of the Participants

Number of Participants	*8*
Males	4
Females	4
Mean Age	25.63
Standard Deviation	2.62
Driving Experience	n
0-3 Years	1
3-5 Years	0
5-7 Years	3
7-10 Years	3
10+ Years	1

2.3 Data Collection & Analysis

The demographic questionnaire was collected using Google Analyse. It allowed for a quick overview of the collected data via auto-generated diagrams. The data was exported to a spreadsheet for further analysis.

The participants was interviewed in order to provide their preferences about the HUD. Windows Voice Recorder was used to record the interviews. One participant could not make it for a physical interview. This interview was recorded using Open Broadcaster Software[11].

The recordings were played back, and notes were taken about the participant's preferences. These were used to determine flaws in the design and improve the design.

2.4 Procedure & Setting

The participants were instructed to fill in a questionnaire to gather demographic data.

The participants were then informed about how the interview would be executed. This meant explaining to the participants that the interview would be recorded(sound only) and how their data would be handled and anonymized.

The first part of the interview was used to gather information about the participants' usage of different systems while driving and how much time and effort they put into using these systems.

In the second part of the interview, the participants were introduced to different designs of a heads-up display for texting. Here the participants were instructed to provide feedback about each design and an overall impression of the design. This part used a semi-structured interview to provide more qualitative feedback about the designs. The total time used per participant was around 25 minutes($\pm$ 5 minutes).

The questionnaires and interviews were conducted in a project room at Aalborg University.

2.5 Results

This section contains the results from conducting the interviews and further analysing of the outcome. It is split into two sub-section, each containing the feedback for different part of the HUD.

2.5.1 HUD Interactions

Every participant understood that the bottom bar of the prototype presented the options for interaction. It was unnecessary with a word specifying that it was commands, the user could choose. Two participants suggested that the design could be customizable as the user got more familiar with the flow of the interactions.

The green colour received positive feedback from seven participants and the last participant could see an issue by using colours often seen in traffic. When asked about improvements to this system, all participants agreed that using colours to indicate the options could be improved by introducing multiple colours to point out different options. The colour layout could then mimic the colours from answering or denying a phone call. This would make the "View text" and the "Answer" field green and the "Cancel" field red.

One participant wanted to move this bar to the top of the HUD instead of having it at the bottom. It caught his eye too much and preferred it to be placed closer to the traffic instead.

2.5.2 HUD Visuals

The participants gave some examples of messages that could be used in one of the two prototypes. If a text message is advertising something or can be ignored, the prototype with limited information would be preferable. One participant stated that if they wanted to see the message most of the time, this could be annoying to have this extra interaction. A solution to this was presented by a participant. A participant suggested that a list of contacts could be made to allow messages from these to be displayed in the HUD. Every other message from contacts not on the list will not be displayed.

The prototypes displaying all information were missing information about the contact sending the message. Seven participants suggested a title above the message itself to contain this information.

Seven participants preferred 50% transparency when it comes to the background. Only one preferred it to be around 25%. The 50% transparency was preferred as the text was easier to read while not dimming the traffic behind it too much. Two participants stated that if the sun was shining, lower transparency might be preferable. A participant was concerned about the visibility of vehicles or alike behind the HUD if the transparency got too low. The ability to adjust the transparency to solve this was first brought up by a participant.

2.5.4 Final HUD

Figure 4 represents the first interaction the user sees when receiving a message. The overall placement of elements has not changed. In addition to the prototypes mentioned above, the final prototype includes a heading to display the sender. The interaction menu at the bottom splits into two options with designated colours for each option: One for "Besvar"(Answer in English), which has the green colour and one for "Afslut"(Finish in English).

Figure 4: Refined version of the HUD with all information present

Figure 5 is the interaction if the user chooses to answer a message. The design is similar to that in Figure 4 but with the addition of the "Rediger"(Edit in English) option.

Figure 5: Refined version of the HUD with limited information present

3 EVALUATING THE HUD

This section goes through the experiment conducted for the designed HUD: It will consist of the method used and the results of the experiment.

3.1 Method

This section introduces the participants in the study, the apparatus and setup used in the experiment, the task the participants had to do, and the experimental procedure.

3.1.1 Participants

As seen in Table 2, eight females and eight males participated in the experiment. All participants were required to fill out demographic data using a questionnaire. The mean age of the participants was 25.38 (± 2.22) age. Every participant had either normal or corrected-to-normal vision. Most of the participants had 5-10 years of driving experience.

Table 2: Demographic Data of the Participants

Number of Participants	*16*
Males	8
Females	8
Mean Age	25.38
Standard Deviation	2.22
Normal Vision	7
Corrected Vision	9
Driving Experience	n
1-2 Years	1
3-4 Years	2
5-10 Years	11
11-20 Years	2

3.1.2 Apparatus & Setup

The apparatus for the experiment consisted of a self-developed driver-centred environment, as shown in Figure 6. The Simulator was constructed in a group room at Cassiopeia, The University of Aalborg. One 50-inch monitor, a driver's seat, G29 Logitech steering wheel and pedals were used for the environment. The resolution of the monitor used was 1920 x 1080. The computer used in the experiment had a 3.60 GHz Intel Core i9-9900K processor and 16GB 2666MHz DDR4 memory. The GPU specifications were NVIDIA GeForce RTX 2080 8GB. To track the participant's eyes on the screen, Tobii Eye Tracker 5 was used. To record the screen with the eye tracker active (without the participants seeing the eye tracker), OBS was used.

The software used for the simulator was Carla Simulator[12] which is mainly written in C++ and Python. The Carla Simulator uses a server-client architecture where the server is written in C++ and the clients use Python. The server side was used unchanged at version 0.9.13. Two clients were built to match the needs of this experiment: (1)A leading vehicle and (2)A client vehicle controlled by the participants.

The leading vehicle utilized the existing self-driving abilities in the Carla Simulator and only minor changes to controls and starting position had to be made. Since the Carla Simulator is optimized for developing self-driving agents, the Python script used for the client vehicle had to be built from scratch.

In addition to the two clients, a message script was developed to control the HUD. This was created using Python and used a local area network to communicate.

3.1.3 Tasks

The task was conducted using the developed HUD and a mobile phone with Messenger(Meta Platforms Inc) installed. For each of the two methods, the participants had to go through 2 types of tasks regarding text-based communication: (1)Read aloud and (2)Read aloud and reply.

For the first type of task, the participants were instructed to read the received message aloud. By reading aloud, the observers could verify that the message received was understood and not just skipped.

Figure 6: Driving simulator used in the experiment, showing HUD and phone usage

For the second type of task, the participants were instructed to read the received message aloud and reply to it as well. This puts an additional load on the participants by forcing them to reply to a given message. The reply had to fit into the context of the message received. If the message stated "Where are you right now?" a reply which was related to any location would be accepted, and the task would be marked as completed. To reply to a message on the phone, the participants had to write on the touch keyboard in messenger and send it. For the HUD, the participants had to say "Besvar"(Reply in English) and then say aloud their reply. When their answer appeared in the HUD, they had the choice of either sending the message by saying "send"(Send in English) or editing the message by saying "Rediger"(Edit in English). By editing a message, they had to say the full message again. The button "Afslut"(Close in English) was also available at this time, but the participants were told to not to use it when replying to a message.

Each type of task, was performed three times for both the HUD and the mobile phone giving a total of 12 tasks per participant.

A task would start when the participants grabbed the phone, or the HUD appeared on the screen. To signal a task was ready on the phone, the phone notified the participants with a sound and vibration. When a participant finished the task, they had to put the phone down or say "afslut" when it was the HUD. However, when the participants had replied to a message on the HUD, they would need to say "send", and the HUD would disappear after 3 seconds.

The researchers used Wizard-of-Oz to send the messages that the participants had to read aloud and read aloud and reply with an external computer. Furthermore, the reply to the messages that the participants would say to a received message on the HUD, was typed in by a researcher where it would then appear in the HUD.

3.1.4 Procedure

Participants first filled in their demographic information, driving experience, and quality of vision by questionnaire when starting the experiment. Participants were asked to sit in the driving simulator and get comfortable with the seating position. When the participant was comfortable with the driving position, they were asked to accustom themselves to the driving simulator. When they were comfortable with driving in the simulator, they would tell us with self-report. To give the same explanation to every participant, a script was followed, so every participant got the same information about what they had to go through in the experiment. Every participant practised using the HUD and the phone before the experiment started. The practice consisted of having a test task on the HUD and the phone as it would be in the experiment.

While doing the tasks, the participants were instructed to follow a leading vehicle while staying in the same lane and maintaining the same speed as the leading vehicle. It was also explained that they had to keep the same speed as the leading vehicle throughout the experiment. A suburban area with highways, traffic lights, and no traffic congestion, except for the leading vehicle, was selected as the environment for the experiment. This was to isolate the effects of variables to only the leading vehicle and not other vehicles.

The participants all followed the same route with the leading vehicle. However, it was balanced whether they started with the phone or the HUD and if they had to start with read aloud or read aloud and answer. When the participants felt ready, they would start with their assigned task. After the participants had completed their task, they were asked to assess their subjective workload using the modified Cooper-Harper rating scale. They were then provided with rest time to avoid fatigue and simulator sickness. The participants self-reported when they had enough rest and were ready to continue the second task. After the participants finished the experiment, a semi-structured interview was conducted to provide user feedback.

3.1.5 Data Collection & Analysis

Video material from the simulator screen was produced of the participant's eye movements when looking at the screen. This video material was produced by OBS and an example can be seen in Figure 7. The eye movement of the participants was manually analyzed for eye glance behaviour. To guide the data collection for eye glances, two patterns were identified. For data analysis, the videos with the eye glances will be examined, and common eye glance behaviour will be grouped together.

The data collected from the simulator were lane crossings, distance to the lead vehicle, the speed difference between the leading vehicle and the participant's vehicle, crashes, and task completion times. This data was logged automatically by the simulator with half a second interval. Task completion time is defined by the description of starting and finishing a task in section 3.1.3 for the mobile phone and HUD. A crash is defined as when the participant drives into the leading vehicle or objects in the environment. Lane crossings are when the participants unintentionally cross the stripes in the road. The participants' subjective workload, using the HUD and the Phone, was collected with the modified Cooper-Harper rating scale, where the participants rated their workload from 1 to 10, where 10 is the highest possible workload. Furthermore, the average word's Pr. response when answering a message was also collected and analyzed. The semi-structured interview at the end of the experiment was to get the participant's opinion of interacting with the HUD and the Phone and what the main differences between the two were.

The data collected during the tasks were cleaned manually and organized using Excel. For statistical data analysis on the data, Excel and IBM Statistic was used. We used Shapiro-Wilk to test for normality distribution in the data. We performed Related-Samples Wilcoxon Signed Rank Test and Paired-samples t-test on non-normally distributed data and normally distributed data, respectively, with the condition of the device used as a within-subject factor.

Figure 7: Screen Snip of A Participants eyes Tracked on Screen

3.2 Results

This section describes the results of the experiment with the HUD, grouped into the different data collected in the experiment. The data from speed difference when receiving a message, lane crossing when receiving a message, average words used pr. message, and subjective workload were not normally distributed. The rest of the data was normally distributed.

3.2.1 Results of Speed Difference

A Related-samples Wilcoxon Signed Rank Test and paired sampled t-test was calculated to see if the speed difference between the leading vehicle and the participants' vehicle had a statistical significance while doing a task. The speed difference, when only receiving a message, was not statistically significant ($z = -0.57, p = .57, r = 1$) for Phone (M = 11.90, SD = 5.88) and HUD (M = 11.96, SD = 3.23) as seen in Table 3. The difference in mean between the experimental conditions was 0.06. The variance for the HUD and the phone is 10.42 and 34.56, respectively.

Table 3: Results of Related-samples Wilcoxon Signed Rank Test for Speed difference When Receiving a Message

N	*Phone* *16*	*HUD* *16*
Mean	11.9	11.96
Variance	34.56	10.42
Standard Deviation	5.88	3.23
Standardized Test Statistics	-.57	
Total N	16	
P-Value (2-sided test)	.57	
Cohen r	1	

The speed difference when receiving and answering a message, for Phone (M = 11.63, SD = 4.02) and HUD (M = 11.99, SD = 2.69) was not statistically significant ($t(15) = .31, p = .76$) as seen in Table 4. The mean difference between the HUD and the Phone is 0.33, with HUD having the higher mean. The variance of the HUD and phone is 7.24 and 16.19, respectively.

Table 4: T-test for Speed Difference when Receiving and Answering a Message

N	*Phone* *16*	*HUD* *16*
Mean	11.63	11.99
Variance	16.12	7.24
Standard Deviation	4.02	2.69
Pearson-Correlation	.12	
Df	15	
T-Statistical	.31	
P(T<=t) Two-tail	.76	
T-Critical two-tail	2.13	

3.2.2 Results of Lane Crossings

As seen in Table 5, lane crossings when receiving a message only are statistical significant ($z = 2.27, p = .03, r = 0.4$) for Phone (Md = 1.29, SD = 1.05) and HUD (M = 0.67, SD = 0.21). This implies that the participants had fewer lane crossings when using the HUD to receive messages. Furthermore, the variance for HUD and Phone is 0.21 and 1.05, which implies that the participants, in general, had fewer lane crossings when using the HUD.

Table 5: Results of Related-samples Wilcoxon Signed Rank Test for lane Crossing When Receiving a Message

	Phone	*HUD*
N	*16*	*16*
Mean	1.29	0.67
Variance	1.05	0.21
Standard Deviation	1.03	0.46
Standardized Test Statistics	2.27	
Total N	16	
P-Value (2-sided test)	.02	
Cohen r	0.4	

The results of the statistical analysis of lane crossings when receiving and answering messages can be seen in Table 6. The result showed, that Phone (M = 2.25, SD = 1.63) and HUD (M = 1.21, SD = .68) it was statistical significant ($t(15) = -2.39, p = .03$). The variance of HUD and Phone was 0.46 and 2.64, respectively. This implies, as with when only receiving a message, that the participants had fewer lane crossings with the HUD when responding to a message.

Table 6: T-test for Lane crossings when Receiving and Answering a Message

	Phone	*HUD*
N	*16*	*16*
Mean	2.25	1.21
Variance	2.64	.46
Standard Deviation	1.63	.68
Pearson-Correlation	.02	
Df	15	
T-Statistical	-2.39	
P(T<=t) Two-tail	.03	
T-Critical two-tail	2.13	

A reason why lane crossings were significant when using the HUD to text can be found in the interviews from the experiment. Six of the participants reported in the interview that it had an impact on their performance that they could have both hands on the steering wheel while texting, as it was easier for them to hold the vehicle in its lane.

3.2.3 Results of Task Completion Time

To see if participants were faster when doing tasks on the HUD than on the phone, the task completion time was used. A two-sample t-test was conducted to see if the completion time for receiving or receiving and answering had a statistical significance. The statistical analysis for receiving a message only can be seen in Table 7. The results show that Phone (M = 8.7, SD = 3.28) and HUD (M = 7.06, SD = 1.75) task completion time, when only receiving a message, is not statistical significant ($t(15) = -2.40, p = .053$). Furthermore, the variance for the HUD and the phone are 3.07 and 10.75, which is significantly higher.

Table 7: T-test for Completion Time when Receiving a Message

	Phone	*HUD*
N	*16*	*16*
Mean	8.7	7.06
Variance	10.75	3.07
Standard Deviation	3.28	1.75
Pearson-Correlation	.37	
Df	15	
T-Statistical	-2.10	
P(T<=t) Two-tail	.053	
T-Critical two-tail	2.13	

Table 8 shows the results of the statistical analysis for completion time when receiving and answering a message. The statistical analysis for Phone (M = 14.98, SD = 3.30) and HUD (M = 13.93, SD = 4.27) revealed that task completion time was not statistical significant ($t(15) = -0.93, p = .36$) when receiving and answering a message.

Table 8: T-test for Completion Time when Receiving and Answering a Message

	Phone	*HUD*
N	*16*	*16*
Mean	14.98	13.93
Variance	10.91	18.24
Standard Deviation	3.30	4.27
Pearson-Correlation	.32	
Df	15	
T-Statistical	-.93	
P(T<=t) Two-tail	.36	
T-Critical two-tail	2.13	

3.2.4 Average Words Pr. Response

To get a better understanding of how many words a participant chose to respond to a message, a Related-samples Wilcoxon Signed Rank Test was conducted. The results show that Phone (M = 1.79, SD = .34) and HUD (M = 3.1, SD = 3.52) average words, when answering a message, is statistical significant ($z = -2.67, p = .008, r = 0.47$), as seen in Table 9. The variance for phone and HUD is .34 and 3.52, respectively. This implies that, in general, a participant would use longer sentences when answering a message using the HUD.

Table 9: Results of Related-samples Wilcoxon Signed Rank Test for Average Words used pr. Answer

N	*Phone* *16*	*HUD* *16*
Mean	1.79	3.1
Variance	0.34	3.52
Standard Deviation	.58	1.88
Standardized Test Statistics	-2.67	
Total N	16	
P-Value (2-sided test)	$.008$	
Cohen r	0.47	

3.2.5 Results of Glance Patterns - HUD

Reviewing the footage of the eye tracker revealed some recurrent patterns for glance times when using the HUD. The usage of the patterns can be seen in Table 10.

The first pattern observed was a two-second initial glance followed by 2-4 glances using one second or less. The first initial glance was generally used to acquire an understanding of the text message as a whole. This was followed by up to 4 glances which seemed to be used for ensuring that keywords from the text were understood correctly before terminating the text or replying. The second pattern was used if a text message consisted of multiple sentences or was split into multiple lines. The participants used two glances(one for each sentence/line) followed by 2-4 glances per sentence/line to clarify some words.

Table 10: Demographic Data of the Participants

Glance Patterns	Receive	Receive & Answer
First Pattern	27	33
Second Pattern	13	8
Other	6	5
Unknown	2	2

The first pattern was used 62.5% of the time(n=60), shared between the two types of tasks, while the second pattern was used in 21.9% of the time(n=21) In Table 10, the *Other* pattern is covering for other patterns observed while reviewing the footage. Examples of *Other* patterns are reading for longer periods without focusing on the traffic(More than 2 seconds) or only using small glances repeatably. Four of the tasks are unknown with regards to the pattern as the eye tracker did not fully obtain the information needed.
The interviews in the experiment revealed that every participant reported that the HUD had an advantage over mobile phone. This advantage was that texting with the HUD while driving, the participants reported that they could still follow what was going on with the leading vehicle. Whereas, with the phone, they had to glance away from the road and would therefore lose their situational awareness.

3.2.6 Results of Subjective Workload Score

Since the collected data for the subjective workload in the modified Cooper-Harper scale was not normally distributed, it was chosen to do a Paired Related-Samples Wilcoxon Signed Rank Test for significance. The results of the statistical analysis can be seen in Table 11. The test shows that the median differences are statistically significant for the subjective workloads of phone and HUD ($z = 3.54, p =< .001, r = 0.62$).

Table 11: Results of Related-samples Wilcoxon Signed Rank Test for Subjective Workload

	Subjective Workload
Standardized Test Statistics	3.54
Total N	16
P-Value	$< .001$
Cohen r	0.62

Comparing the medians of the subjective workload for HUD and Phone seen in table 12 reveals that the median for HUD and Phone is two and eight, respectively. This implies that the subjective workload decreases when using the HUD for text-based communication. Furthermore, the upper whisker for the HUD is 2, whereas the lower whisker of the phone is 4. This implies that the perceived subjective workload is higher when using the phone for texting than using the HUD.

Table 12: Box Plot Data for Subjective Workload

N	*Phone* *16*	*HUD* *16*
Upper Whisker	10	2
3rd Quartile	9	2
Median	8.5	2
1st Quartile	6.5	1.5
Lower Whisker	4	1

3.2.7 Number of Crashes

An important thing to do while driving a vehicle is not to crash the vehicle.

Table 13 shows the number of crashes during the different experimental conditions. Every different driving task had crashes. The driving tasks with the most number of crashes were *Phone - Receive & Answer* with four crashes. The overall crash rate during the experiment was 12.5%, with eight crashes in 64 experiments.

Table 13: Number of crashes according to experimental conditions

Experimental Condition	N	*Number of crashes*
HUD - Receive	16	1
HUD - Receive & Answer	16	1
Phone - Receive	16	2
Phone - Receive & Answer	16	4

The number of crashes shows that it happens when driving and doing a task simultaneously. Calculating the likelihood ratio for crashing with the HUD rather than the phone showed one is 3.5 less likely to crash.
The video material was further analyzed, and the reasons for accidents were grouped. Analysis of the video material revealed two main reasons for crashing. These reasons are; not paying attention to the leading vehicle and not paying attention to the environment. As Table 14 shows, six of the crashes were caused by the participants not paying attention to what the leading vehicle did. The last two crashes were caused by the participants not paying attention to where the road was going.

Table 14: Reasons for crashes

	HUD	Phone
Not paying attention to leading vehicle	2	4
Not paying attention to the road	0	2

4 DISCUSSION

Overall the results indicate that comparing the HUD for texting rather than the phone has some impact on the participants driving performance. The participants were better to steer the vehicle and have significantly fewer lane crossings when receiving & receiving and answering a message. Furthermore, the variance of the lane crossings was also significantly lower when using the HUD. The subjective workload for the participants was also significantly lower than using a phone.

It was found by Virginia Research [13] that one texting while driving was 23 times more likely to crash. In this experiment, we found that using the phone over the developed HUD for texting one was 3.5 times more likely to crash. This could imply that using the HUD for texting is safer than the phone when driving. However, the testing environment differs from the one used in our experiment, where Virginia Research used real vehicles for testing and we used a simulator with no traffic congestion other than the leading vehicle. This could have had an influence on the results of this experiment since higher traffic congestion could lead to the participants needing to focus differently.

While observing the video material, and the observation of the participants under their tasks, it was discovered that the participants had tendencies of looking at their phones, even though the signal that a task was ready for the participants was not given. This could suggest that it had an impact on the participants in the experiment, that they were always aware and thought of the phone during tasks that involved the phone. However, the participant's eyes were not tracked when they looked away from the screen of the simulator.

In the interview, the participants were asked about the usage of a mobile phone while driving. Table 15 depicts the results of the interview, regarding what degree the participants use their phones while driving. Only three of the participants in the experiment have had heavy usage of their phones while driving, whereas five had no usage at all. We tried looking at the collected average of the participants in each grouping in Table 15, based on the data collected on lane crossing, task completion time, crashes, and speed difference. However, this revealed no clear tendencies in driver performances based on their experience using a mobile phone. This suggests that experience did not have a factor in how a participant performed in our simulator with the driving tasks.

Table 15: Degree of Phone usage of Participants

N	**16**
Heavy Usage (Texting etc.)	3
Light Usage (Change Song, See Notifications etc.)	5
Does Not Use	5
Used Legally (Phone In a Stand)	3

During the interview, the participants were asked about how they felt driving the simulator and if they could relate it to real-world driving. Table 16 depicts the participants' answers grouped into categories. As the results of the interview show over half of the participants felt that the simulator was realistic, but had some shortcomings, such as missing the mirrors on the side of each car. However, the participants who expressed this still believed that the simulator had a high enough fidelity for the objective of the experiment. The participants, that felt it was not realistic, explained it by stating they felt like they were in a video game. One participant expressed that they would have been more careful if driving a real car. This could have had an influence on the results, that some of the participants did not believe it was realistic and therefore possible made some choices, that they would not have normally.

Table 16: Realistic Degree of Simulator from Participants View

N	**16**
Not Realistic	4
Realistic, But Had Shortcomings	10
Realistic	2

The completion time of the tasks between the phone and the HUD showed no significance in the statistics test since the mean completion time for both tasks when using the HUD or the phone had a small difference. However, there is a significant difference in how the participants answered the messages. The results revealed that the participants had a tendency to give longer answers when using the HUD over the phone. This could imply that the participants feel safer when using the HUD and therefore gives more detailed answers.

5 CONCLUSION

In this paper, we showed that using a HUD over a phone for texting has an impact on driving performance. This impact is seen in lane crossings, where participants had fewer of those when using the HUD. Furthermore, the risk of being in a crash with the HUD was 3.5 times less likely. Completion time of the tasks showed no significance when using the HUD or the Phone, however, the participants used more words to write their messages when using the HUD. The participants reported that their subjective workload was significantly lessened when using the HUD.

5.1 Limitations

A Limitation to this study, is that the experiment was conducted in a controlled environment using a simulator. Therefore, the results of the study can first be confirmed, if the HUD is being tested in a real-life setting, with a real automotive vehicle. Furthermore, the environment used in the study for testing the HUD had no traffic congestion other than the leading vehicle. To further confirm our results, more research is needed with more life like traffic congestion, with more vehicles, pedestrians, and cyclists. Furthermore the HUD should also be tested with more variation to road types.

5.2 Future Work

For future study, it could be interesting to test the HUD with more variables, such as glance behaviour with both the phone and the HUD and see what effect the HUD would have on the user's reaction time. Furthermore, it could also be interesting to test the HUD in a setting with different intensities of traffic congestion, such as low, medium, and high, to see what effects it would have on the driver's performance.

ACKNOWLEDGEMENTS

This paper was supported financially by the University of Aalborg, which provided the Apparatus used in this experiment. We would also like to thank all the participants who participated in the project.

References

[1] Transport- og Boligministeriet, *Vejledning om færdselslovens forbud mod anvendelse af håndholdte mobiltelefoner under kørsel*, May 1998. [Online]. Available: https://www.retsinformation.dk/eli/mt/1998/77 (visited on 03/09/2022).

[2] / *GHSA*. [Online]. Available: https://www.ghsa.org/state-laws/issues/distracted%5C%20driving (visited on 03/09/2022).

[3] Jens, *Skriver du også sms'er bag rattet?*, da-DK, Jan. 2018. [Online]. Available: https://www.hvilkenbil.dk/skriver-du-ogsaa-smser-bag-rattet/ (visited on 03/09/2022).

[4] National Highway Traffic Safety Administration, "Distracted Driving 2016", ENG, *NHTSA*, p. 7, Apr. 2018. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812517.

[5] National Highway Traffic Safety Administration, "Distracted Driving 2018", *NHTSA*, p. 7, Apr. 2020. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812926.

[6] National Highway Traffic Safety Administration, "Distracted Driving 2019", ENG, *NHTSA*, p. 7, Apr. 2020, DOT HS 812 926. [Online]. Available: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813111.

[7] G. Villalobos-Zúñiga, T. Kujala, and A. Oulasvirta, "T9+HUD: Physical Keypad and HUD can Improve Driving Performance while Typing and Driving", en, in *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Ann Arbor MI USA: ACM, Oct. 2016, pp. 177–184, ISBN: 9781450345330. DOI: 10.1145/3003715.3005453. [Online]. Available: https://dl.acm.org/doi/10.1145/3003715.3005453 (visited on 03/18/2022).

[8] M. B. Hoy, "Alexa, siri, cortana, and more: An introduction to voice assistants", *Medical Reference Services Quarterly*, vol. 37, no. 1, pp. 81–88, 2018, PMID: 29327988. DOI: 10.1080/02763869.2018.1404391. [Online]. Available: https://doi.org/10.1080/02763869.2018.1404391.

[9] A. Berdasco, G. López, I. Diaz, L. Quesada, and L. A. Guerrero, "User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana", en, MDPI, Nov. 2019, p. 51. DOI: 10.3390/proceedings2019031051. [Online]. Available: https://www.mdpi.com/2504-3900/31/1/51 (visited on 03/22/2022).

[10] Y.-C. Liu and M.-H. Wen, "Comparison of head-up display (HUD) vs. head-down display (HDD): Driving performance of commercial vehicle operators in Taiwan", en, *International Journal of Human-Computer Studies*, vol. 61, no. 5, pp. 679–697, Nov. 2004, ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2004.06.002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581904000497 (visited on 06/10/2022).

[11] *Open Broadcaster Software | OBS*. [Online]. Available: https://obsproject.com/ (visited on 06/08/2022).

[12] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator", in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[13] *New data from Virginia Tech Transportation Institute provides insight into cell phone use and driving distraction*, en. [Online]. Available: https://vtx.vt.edu/content/vtx_vt_edu/en/articles/2009/07/2009-571.html (visited on 06/09/2022).