THESIS FOR THE DEGREE OF LICENTIATE OF ENGINEERING

Data-Driven User Behavior Evaluation

JULIA ORLOVSKA

Division of Product Development

Department of Industrial and Materials Science

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2020

Data Driven User Behavior Evaluation
JULIA ORLOVSKA

Division of Product Development
Department of Industrial and Materials Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Cover picture: © Shutterstock

ABSTRACT

Automotive Original Equipment Manufacturers (OEMs) compete worldwide to stand out with new trends and technologies. Automated Driver Assistance Systems (ADAS) are an example of advanced solutions where a lot of effort is put into the development and utilization of vehicle data. ADAS systems range from different types of information/warning systems to adaptive functions designed to assist the driver in the driving tasks and ensure more efficient and comfortable driving. These types of systems have become standard at many OEMs, including Tesla, Cadillac, BMW, Mercedes, Volvo Cars, and others. Volvo Cars is well-known for the development of such ADAS functions as ACC (Adaptive Cruise Control) and PA (Pilot Assist). These functions offer lateral and/or longitudinal support, but leave the driver in full control and with responsibility for the driving task.

The ADAS systems are not fully automated. These systems have a number of limitations related to the context where they can operate. Previous studies have demonstrated that the drivers' understanding and adoption of these systems is not definite and may vary from full technology acceptance to complete ignorance. Therefore, in-depth understanding and interpretation of driver behavior and needs regarding the use of ADAS can significantly help developers to reflect on and improve the systems to meet the users' expectations.

Recently, the availability of data coming from the in-vehicle sensors network has increased significantly. The amount of received data potentially enables the in-depth quantitative driver behavior evaluation in a time-efficient and reliable way. Moreover, the ability of vehicle sensors and actuator data to synchronize the driver and system performance and assess the driving conditions in the moment of driver-system interaction can contribute to the comprehensive context-aware ADAS evaluation.

Developing methods for objective assessment of driver behavior is a task with a high level of complexity. This process requires (i) investigation of the driver behavior assessment area where vehicle data can be useful; (ii) identification of the influencing factors for evaluating ADAS functions; (iii) definition of the relevant data for the data-driven driver behavior evaluation; (iv) investigation of the ways to improve the feasibility of vehicle data.

The research presented in this thesis focuses on the understanding of vehicle data applicability in user-related studies. The core of this research is the methodology for objective ADAS evaluation and a mixed-method approach that helps to integrate the quantitative methodologies into existing, mainly qualitative, evaluation practices.

The conducted research revealed that vehicle data offers the possibility to determine individual user behavior, and to describe, categorize, and compare this to the average within a group. All of the above mentioned makes the applicability of vehicle data for user-related studies meaningful.

Keywords: vehicle data, data-driven evaluation, driver behavior, ADAS

ACKNOWLEDGEMENTS

Julia Orlovska
Gothenburg, January 2020

APPENDED PUBLICATIONS

The following research publications form the basis for the presented research.

Paper A.   **Orlovska, J.,** Wickman, C. and Söderberg, R., 2018. Big data analysis as a new approach for usability attributes evaluation of user interfaces: an automotive industry context. In *DS 92: Proceedings of the DESIGN 2018 15th International Design Conference (pp. 1651-1662).* DOI: https://dx.doi.org/10.21278/idc.2018.0243.

Paper B.   **Orlovska, J.**, Wickman, C. and Söderberg, R., 2018, January. Big Data Usage Can Be a Solution for User Behavior Evaluation: An Automotive Industry Example. In *Procedia CIRP 2018 51st CIRP Conference on Manufacturing Systems*. DOI: https://doi.org/10.1016/j.procir.2018.03.102

Paper C.   **Orlovska, J.**, Novakazi F., Wickman, C. and Söderberg, R., 2019. Mixed-Method Design for User Behavior Evaluation of Automated Driver Assistance Systems: an Automotive Industry Case. *International Conference on Engineering Design 2019 (ICED 2019)*, 05-08 August, Delft, Netherlands. DOI: https://doi.org/10.1017/dsi.2019.186

Paper D.   **Orlovska, J.**, Novakazi, F., Bligård, LO., Karlsson, I.C.M., Wickman, C. and Söderberg, R., 2019. Effects of the driving context on the usage of Automated Driver Assistance Systems (ADAS). Naturalistic Driving Study for ADAS Evaluation. *Transportation research Interdisciplinary perspectives (Submitted 07/10/19, accepted with minor revision).*

WORK DISTRIBUTION

The work in terms of writing, developing the ideas, collecting the data, analyzing results or producing core findings as well as the reviewing work for every appended paper was distributed among the authors according to the following.

Paper A.   **Orlovska, J.** performed data collection and data analysis, presented the core findings of the paper, and finally wrote the paper. Wickman C. and Söderberg R. contributed with ideas and provided review comments of the full paper.

Paper B.   **Orlovska, J.** wrote the paper. The author formed the initial ideas and designed the framework. The core findings were provided by **Orlovska J.**, while Wickman C. and Söderberg R. provided their feedback and contributed as reviewers.

Paper C.   **Orlovska, J.** wrote the paper. The initial idea of the presented method belonged to the author. The method design, as well as the core findings, were formulated by **Orlovska J.** Novakazi F. contributed in corroboration of methodology for qualitative evaluation. Wickman C. and Söderberg R. provided their feedback and contributed as reviewers.

Paper D.   **Orlovska, J.** was the lead author of the paper and Novakazi F. wrote sections related to the qualitative evaluation. The initial idea of the paper belonged to **Orlovska J.** The method design was formulated by **Orlovska J.**, quantitative data collection and data analysis was performed by **Orlovska J.** Qualitative data collection and data analysis were performed together by **Orlovska J.** and Novakazi F., while Novakazi F. contributed in producing of qualitative findings. Bligård LO. contributed to conceptualization, representation of the results and reviewing the paper. Karlsson, I.C.M. contributed by structuring and reviewing the paper. Wickman C. and Söderberg R. provided their feedback and contributed as reviewers.

ADDITIONAL PUBLICATIONS

Paper E.   **Orlovska, J.**, Wickman, C. and Söderberg, R., 2019. Capturing Customer Profile Enables in-Vehicle User Identification: Design for Data-Based User Behavior Evaluation. In *Research into Design for a Connected World (pp. 665-675). Springer, Singapore*. Available at: https://doi.org/10.1007/978-981-13-5977-4_56

Orlovska, J. wrote the paper. The initial idea of the presented method belonged to Wickman C. The method design, as well as the core findings, were formulated by the author. Wickman C. and Söderberg R. provided their feedback and contributed as reviewers.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND TERMINOLOGY

ADAS
Automated Driver Assistance Systems - built-in vehicle assistance systems that support and facilitate the primary driving task, providing longitudinal and lateral support.

ACC
Adaptive Cruise Control – ADAS function providing longitudinal control of the vehicle

Big data
Describes a new generation of technologies and architectures, designed to extract *Value* from very large *Volumes* of a wide *Variety* of data, by enabling high-*Velocity* capture, discovery, and/or analysis.

CAN
Controller Area Network

CUE
Comparative User Testing

DC
Drive Cycle – the driving activity that starts when the ignition of the engine turns on and ends when the ignition of the engine turns off.

Driver behavior
The set of actions taken by a driver interacting with the system in order to reach a goal or complete a task.

Driver ID
User Identification number – few Driver IDs can be connected to one vehicle.

Driving context
The summary of external factors that affect driver behavior while using the evaluated system. For the ADAS evaluation, the driving context is defined as the aggregation of traffic, road, and weather conditions that in the association, encourage or discourage the usage of the ADAS.

DRM
Design Research Methodology

DSS
Driver Support Systems – the internal term for ADAS at the OEM. This abbreviation was used in the early stages of this research. However, to easily match Volvo DSS to other OEM's in the literature, the term ADAS that is broadly accepted was adopted.

GPS
Global Positioning System

HCI
Human-Computer Interaction

HMI
Human-Machine Interaction

ISO	International Organization for Standardization
ND study	Refers to a study where a strict experimental design does not constrain the data collection, and where the data is gathered in a natural driving context and under various driving conditions, closely resembling real-driving situations.
ND data	Naturalistic Driving Data – data collected in ND study
OEM	Original Equipment Manufacturer
PA	Pilot Assist – ADAS function providing longitudinal and lateral control of the vehicle
Primary Task	Driving the vehicle
SAE	Society of Automotive Engineers
Secondary Task	Task, unrelated to the driving, but requires the driver's attention from driving task. Example: turning on music, talking to the phone, etc.
UEMs	Usability Evaluation Methods
UI	User Interaction
Usability	The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use (ISO/IEC 9241-11).
UX	User Experience - defines as a person's perceptions and responses that result from the use or anticipated use of a product, system or service (ISO 9241-210).
Vehicle ID	Vehicle Identification number
WICE	Wireless communication and data acquisition system

I

INTRODUCTION

We live in a digital world where technologies become available at the touch of a screen, and the myriads of digital solutions are there to straighten our possibilities and automate our everyday tasks. This mainstream came from the Web and mobile applications development and has now reached the automotive sector. The automotive Original Equipment Manufacturers (OEMs), are forced to cope with the digitalization process and be engaged in the development of smart and connected automotive solutions. The automotive systems have begun the transformation from purely mechanical to "smart" programmable systems (Tornell et al., 2015). Hundreds of sensors support the performance of these systems to enable in-vehicle connectivity, providing new functionality, and automating the existed solutions.

As a result of this transformation, a number of solutions automating our driving activities emerged. Automated Driver Assistance Systems (ADAS) are good examples of corresponding solutions. ADAS are built-in vehicle support systems that provide longitudinal control of a vehicle through accelerating or braking in various traffic conditions, and/or lateral control through providing steering assistance (Naranjo et al., 2003). The main purpose of ADAS is to support and facilitate primary driver activities, providing driving assistance in real-time driving.

The development of systems like ADAS has changed the nature of driver activities. Nowadays the driver is cooperating with different automated functions offered by the vehicle. This cooperation presumes a good understanding of the system and the functions it provides. However, the systems are not fully automated, which means that they have limitations and are not able to provide support correctly in all driving conditions. The driver is ultimately responsible for monitoring whether the system with the existing limitations is capable of operating under dynamically changing driving conditions. This supervision role can be heavy (demanding) for the driver if the driver does not have a good understanding of how the system

works.

Previous studies show that a significant number of drivers do not fully understand the limitations of driving support systems (Llaneras, 2006). In many cases, drivers expect the system to be able to handle on-road situations when the system activation preconditions were not fulfilled. Moreover, the level of automation can be different between systems in the vehicle. This means that the driver can misinterpret the system capabilities and not engage when the system requires intervention from the driver. The study conducted by Jenness et al. (2007) demonstrated that drivers' expectations regarding the system capabilities were higher than actual capabilities of system performance. A wrong understanding of the system's limitations creates misconceptions between the driver and the system. Consequently, this misinterpretation of the ADAS capabilities harms driver's trust and reliance on technology (Itoh, 2012; Kazi et al., 2007) and may decrease technology use and acceptance.

For the OEM, the successful implementation of ADAS means the adoption of the system, and its usage in the way developers intended, taking into consideration the limitations of the systems. Therefore, an in-depth understanding and interpretation of driver behavior and needs regarding the use of ADAS can significantly help developers to reflect on and improve the systems to meet the users' expectations.

Traditionally, the evaluation of driver behavior has been performed with qualitative methods, trying to understand and explain the root causes of specific user behavior (Creswell, 2014). However, the use of qualitative methods for driver behavior evaluation of systems such as ADAS, where the complexity of driver-system interaction is high, and the driver comprehension is uncertain, often makes driver input non-relevant. Moreover, the driver behavior and usage of these systems is not solely dependent on human factors and system performance but is also influenced by the continually changing driving context. If a driver attempts to activate the system in a driving context that does not comply with the intended design requirement, the system performance can be unstable and lead to rejection by the user. In these cases, the driver is not always able to interpret the situation correctly and, therefore, cannot provide the developers with the correct feedback. In addition to this, a qualitative researcher's own perception can bias participants' view on the evaluated object through specifically designed tasks and questions.

The improvement of data feasibility that comes together with the development of sensors-based smart technologies brings new abilities for "smart" systems evaluation. Any vehicle today generates a large amount of real-time data coming from sensors that enable the performance of ADAS functionalities. The same sensors can potentially provide user-related data on driver performance. This information can contribute to the understanding of the effectiveness of driver-system collaboration. Furthermore, the ability of vehicle data to identify the driving event and assess the driving conditions in the moment of driver-system interaction can contribute to comprehensive context-aware ADAS evaluation.

Vehicle data analysis can be particularly helpful for (1) measuring the overall ADAS use level; (2) categorizing the drivers, according to their use and performance of the evaluated functions; (3) identifying the key differences in user behavior that depend on the specificity of driving context; (4) identifying potential problem areas in driver-system interaction and measuring its severity; (5) identification of various trends and patterns in the recorded use scenarios.

Therefore, the use of vehicle data in the evaluation loop of how the system is understood and used by drivers is a critical activity for the ADAS developers. This approach is able to provide the developers with automated input on how the system can be improved to meet the user's expectations.

2

1.1 DEFINING THE PROBLEM

The usage of signal data to understand driver behavior and fulfilment of design intent of different systems began to develop in the areas such as internet services (Angelini, M. et al., 2018; Carta, T. et al., 2011), smartphone apps development (Visuri, A. et al., 2017), and the gaming industry (Kim J.N. et al., 2008). For the automotive industry, this area is relatively new. The main reason is that the automotive software platforms were initially were designed with a limited ability to obtain user-related vehicle data. The vehicle data was mainly used for system performance verification.

The positive examples of utilizing user-related data in other domains made the OEMs start to think of possible solutions for better in-vehicle connectivity. The easiest way to override this problem would be the development of a new automotive platform with a high level of connectivity between all in-vehicle systems. However, due to the high complexity of a vehicle as a system, the development of such a platform has high costs for the automotive sector. As a solution, most OEMs chose the step-wise development of the existing platforms, building up new add-ons for the existing platforms. This approach slows the speed of the development of in-vehicle connectivity and affects data feasibility/compatibility among various developed solutions. This often limits the acquisition of relevant data, which is one of the main problems for this research.

Another issue is the lack of established approaches regarding the use of vehicle data for user-related studies. This area is relatively new for the automotive industry. Despite the positive examples from other fields that demonstrate the potential for data-driven user behavior evaluation, the automotive practices revealed that user evaluation is mostly connected to the traditional, mainly subjective methods. Therefore, the main challenge is to define the scope where the vehicle data can be used for driver behavior evaluation and identify a reliable way that can contribute to a better understanding of driver behavior under the various driving conditions.

1.2 RESEARCH FOCUS

Tracking of driver behavior with the help of vehicle data will provide developers with quick and reliable user feedback on how drivers are using the system. This input will include all interrelations between driver behavior, system performance, and other vehicle systems and components. This integration of the results will help to consider driver input and better understand how the system under evaluation can be improved to meet driver needs.

According to this aim, the research focuses on two main aspects:
- Finding effective solutions for vehicle data use in user-related studies.
- Identifying the extent to which vehicle data can contribute to driver behavior understanding.

Additionally, special consideration is given to the structuring and developing and improvement of the feasibility of vehicle data. This means that constant work on re-evaluating data requirements to improve a vehicle-based dataset for driver behavior evaluation needs to be done.

1.2.1 Scientific goals

From the academic point of view, vehicle data utilization is a reasonably new area with the ongoing development of methods for logging, processing, analysis, and visualization of interaction data (Visuri A. et al., 2017; Vuillemot R. et al., 2016).

Specific research should also be addressed to the developing of methodology for incorporating newly generated data-driven knowledge in the existing methods for user behavior evaluation and into the decision-making process.

Today, there is no single method that helps to capture the complexity of user behavior. Therefore, the research should be focused on the effective combination of existing methods for user behavior evaluation. The academic acknowledgment of such methodology is needed. Thus, the scientific goal of this research is to design methods for effective user behavior evaluation utilizing vehicle data and to study how these methods can be incorporated into the existing methodologies for user behavior evaluation.

1.2.2 Industrial goals

This research project has been carried out in close collaboration with Volvo Cars. The overall purpose of the project is to learn how to utilize vehicle signals at the company level to understand the usage of advanced in-vehicle support systems. This data-driven evaluation can subsequently contribute to understanding the implications, advantages, and limitations of the system from a user point of view.

To achieve this goal, the following company resources need to be examined in the first place: (i) the technical feasibility of vehicle data and its limitations; (ii) the possibilities for further development of the dataset; (iii) the means and methods for data acquisition, data pre-processing and data storage. The above-described actions will help to build in robust infrastructure for data support at the company level.

Another goal for this research is to define the scope where the vehicle data contributes to the understanding of user behavior since results solely based on vehicle data cannot uncover all aspects of driver behavior. In this step, the validation of obtaining results is one of the primary industrial goals.

The ultimate goal is to contribute to practical solutions for effective vehicle data acquiring, data processing, and data utilization, taking part in developing the relevant dataset, improving vehicle sensors, and the data acquisition system. The above-described improvements will facilitate the future use of vehicle data in real-time driver support.

1.2.3 Research questions and hypothesis

This research is based on the hypothesis that the correct specification of vehicle data and consequent collection and analysis of these data can provide a better understanding of user interaction between the driver and the evaluated system. The assumption that it would be possible to evaluate the driver behavior regarding the evaluated system or function in various driving context was set.

As a result of the set assumption the following research questions were identified:

RQ 1) How can vehicle data be used for data-driven user behavior evaluation?

RQ 2) How can the data-driven approach be incorporated into existing methods for driver behavior evaluation?

RQ 3) How can the validity of the data-based results be achieved?

RQ 4) How can vehicle data be used for the data-driven design of vehicle systems?

1.3 RESEARCH SCOPE

Operating the vehicle is a complex and multi-tasking activity. On the road, a driver communicates with a variety of the subsystems and functions of the vehicle controls and interacts with the external environment. Current technologies do not allow us to track and understand the complete driver behavior with respect to all in-vehicle systems. To limit the scope for the driver behavior evaluation to a manageable area, the driver behavior evaluation in this research was focused on the driver behavior assessment of two ADAS functions, namely Adaptive Cruise Control (ACC) and Pilot Assist (PA). The ACC function uses vehicle cameras

and radar to automatically adjust the vehicle's speed with regard to other objects moving in front or keeps a set speed. In other words, ACC provides longitudinal control of a vehicle through accelerating or braking according to pre-set speed and time interval to the vehicle in front. PA offers the functionalities ACC offers, but also provides steering assistance, helping to keep the vehicle in its lane at the set speed and preselected time interval to the vehicle in front as long as there are clear markings on the road. Thus, the PA provides both longitudinal and lateral control of a vehicle. (Volvo Cars, 2019).

Although ACC and PA provide lateral and/or longitudinal support, they are semi-autonomous systems. This means that these systems leave the driver in full control and with the responsibility for the driving task. According to the SAE classification (SAE standard J3016, 2018), six levels of driving automation are defined, ranging from level 0 (complete manual driving) to level 5 (fully autonomous driving). Figure 1 provides a detailed description of the driving automation levels.

LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4	LEVEL 5
NO AUTOMATION	DRIVER ASSISTANCE	PARTIAL AUTOMATION	CONDITIONAL AUTOMATION	HIGH AUTOMATION	FULL AUTOMATION
Manual control. The driver performs all driving tasks like steering, braking, or acceleration and others.	The vehicle features a single automated system (e.g. ACC that keeps pre selected speed through braking&acceleration)	The vehicle can perform steering and acceleration. The human still monitors all tasks and can take control at any time.	The vehicle can perform most driving tasks, but human override is still required. Environmental detection capabilities.	The vehicle performs all driving tasks under specific circumstances. Geofencing is required. Human override is an option.	The vehicle performs all driving tasks under all conditions. No human attention or interaction is required.

Figure 1. Levels of driving automation (SAE standard J3016).

ACC is defined as a Level 1, Driver Assistance system, according to the SAE classification. ACC is designed to be a supplementary driving aid and is not intended to replace the driver's attention and judgment. PA is classified as a Level 2, *Partial Automation* system. Level 2 of the SAE classification means that the driver has full responsibility for the driving task even though the system is able to provide steering assistance as well as braking and acceleration support. The driver has to monitor the driving environment and be prepared to take back control of the system at any time.

Different levels of driving automation expect a different level of driver involvement in ADAS performance, starting from full control over the ADAS performance and ending up with zero interaction with the system. These levels of driver involvement have a significant effect on driver behavior. At Levels 1-2, driver interaction with the automation is the highest, since the automation leaves the driver with full responsibility for the driving activity. Since ACC and PA are classified as Level 1-2, driver behavior in this research is a specific behavior that is connected to Levels 0-2 of automation. The results of the driver behavior evaluation cannot be applied to the same systems with a higher or lower level of automation.

This research is based on data from one automotive company. The ADAS functions are offered today by most car manufacturers, however, in this research, only two ADAS functions designed by a single OEM were investigated. The limited scope is justified by the poor data feasibility and low utilization of vehicle data at the OEM at the start of this research. As a result, significant time was spent to design and ensure the feasibility of the relevant dataset needed for user behavior evaluation. Moreover, almost a year of data collection was required to be able to

achieve the first results.

Thus, this research is grounded on the utilization of vehicle data. Data from CAN-bus (Controller Area Network) and the GPS data are the primary data sources in this research. Many of the ND studies, due to the subjective nature of driver behavior, try to capture personal driver data by including the use of advanced technologies, such as eye-tracking technology, reading of biological data, and measuring driver's psychosomatic parameters. These approaches often require complicated and expensive equipment and additional adjustments to the vehicle's configuration. The major drawback of this approach is the limited number of cars available for evaluation, due to the specific equipment requirements and legal limitations regarding the generation and processing of sensitive data.

Moreover, the complexity of such a dataset would increase the volume of data collected from the vehicle significantly, without providing any benefits for the data analysis, since the outcomes and meanings of the video stream or eye-tracking data are often dubious. To compare explicit and implicit user opinion eye-tracking studies must be still supported by the questionnaires (Köhler et al., 2015). Considering the above, the main focus of this research is given to the design of simple and reliable solutions that enable the evaluation of the bigger vehicle pool. In the future perspective, this will allow the implementation of developed solutions for driver behavior evaluation in all vehicle models entering the market.

2

FRAME OF REFERENCE

This chapter explains the concepts, phenomena, and context to which this research relates. Moreover, this chapter provides a brief overview of the research field to familiarize the reader with the existing terminology, approaches, and methods used in the area of driver behavior evaluation.

2.1 VEHICLE DATA

Recently, the importance of data and its applicability for different tasks has become a prominent discussion topic. The rapid development of this domain has been incepted by significant progress in the areas such as mobile applications development, web services, and the game industry (Visuri et.al., 2017; Angelini et al., 2018; Carta et al., 2011; Kim et al., 2008). Speaking about automotive OEMs, the complexity of products and therefore the volume of data that is needed, is much higher comparing to other products. The continuously increasing in-vehicle connectivity opens new capabilities for obtaining new objective usability data (Tornell et al., 2015), demonstrating the great potential of utilizing vehicle data for diverse purposes. Nowadays, automotive OEMs generate enormous amounts of data in volume, with high velocity, in real-time, to support the development of automated processes.

Along with the system performance assessment, the vehicle data analysis gives us the possibility for context-aware user behavior evaluation and indicates how well the user understands the system functionality. Vehicle data also offers the ability to determine certain trends in user behavior, as well as identify specific use errors, the usage of a particular function, and other usability issues (Orlovska et al., 2019). Additionally, the ongoing research of the quality of vehicle data and its applicability have a positive effect on the feasibility of this data, which has been gradually improving over recent years.

2.1.1 Big data

Big data is a relatively new term that usually refers to a significant volume of data that is difficult to store, process and analyze using traditional database technologies (Manyika et al., 2011). The definition of big data can vary from a large volume of data for scientific visualization

(Berman, 2013) to a large volume of data that is beyond the technical capability to store, manage and process efficiently (Gantz and Reinsel, 2011). The most traditional way to describe big data is using the "three V's" - big data characteristics: Volume, Variety, and Velocity (Chen et al., 2014). However, some researchers argue that Value (the fourth "V" characteristic), is the most important dimension of big data (Alpaydin, 2010; Johanson, 2017). Value extraction is the main purpose of big data processing. It highlights the importance of big data as a source of knowledge and refers to the process of discovering hidden values from large datasets (Tullis and Albert, 2013).

The definition provided by Gantz and Reinsel (Alpaydin, 2010) was initially adopted in this research: "*Big data describes a new generation of technologies and architectures, designed to extract value from very large volumes of a wide variety of data, by enabling high-velocity capture, discovery, and/or analysis.*" Big data can include transactional data, warehoused data, metadata, and other data that could be captured from information on product consumption or utilization. Thus, in early papers, the author used the big data term.

However, recently, the existing interpretation of bid data has been changed. Some researchers started to distinguish *big data* and *large data*, adding that *big data* requires real-time data analysis, where the Machine Learning (ML) algorithms are used. Acknowledging this newly added interpretation, in the later papers, the term big data is avoided. The author instead uses the terms vehicle data, sensors-based data, Naturalistic Driving (ND) data, objective data, or quantitative data (when conducting mixed-method studies that include quantitative and qualitative approaches).

Despite potential advantages, the domain of using vehicle data in the automotive industry is poorly explored mainly because the automotive platforms are slower in the development of data technologies. The specificity of automotive platforms is that initially they were not designed to support data logging for data-driven evaluation in the first place. The development of a new automotive platforms supporting the new trend of data-driven design and assessment, is extremely costly today. Thus, automotive companies chose step-wise development of the existing platforms, adding layers and layers of new lines of code to the existing one.

2.1.2 External data acquisition solutions

External acquisition systems are developing, in parallel with the improvement of the automotive platforms, as intermediate solutions to support the current needs for data-driven evaluation. Today, to achieve the required data collection, the external wireless communication and data acquisition unit needs to be installed in all test vehicles. It enables the management of the data from the vehicle fleet, by keeping track of map-based positioning, mileage, uptime and diagnostic codes. In this research, the WICE-system is an external wireless communication and data acquisition unit that requires installation in the test vehicle. It supports the testing and validation stages of automotive development by efficient use of telematics technology and global coverage (Johanson, 2017). The WICE system consists of two major parts: (i) *Wireless Communication Unit (WCU)* - the hardware unit that supports communication interfaces for data logging and measuring, including telematics services. Types of logged data: CAN bus and FlexRay bus, analog inputs, digital inputs, USB and Ethernet data; (ii) *Back-end server infrastructure* - includes the web-based front-end user interface, including data storage units and database with meta-information.

Overall, the system provides metrology services from connected vehicles, including a collection of measurement data signals of various types (logs, signals, images, video, etc.). The WICE system is able to manage information regarding vehicle fleet by keeping track of map-based positioning, mileage, uptime, Diagnostic Trouble Codes, etc. Figure 2 shows the high-level architecture for WICE data logging and the real-time data processing system.
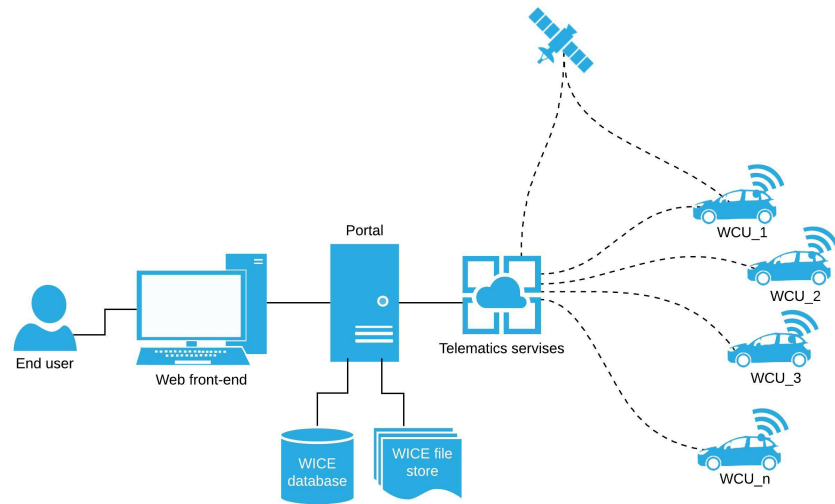
Figure 2. A high-level overview of the WICE-system communication infrastructure

The WICE portal is a complex software providing server-side functionality for vehicle testing, verification, and development. The WICE users interact with the system through the Web front-end that gives users access to the WICE application services and data. The WICE portal implements the core functionality of the supported services including fleet management of connected vehicles, tasks and data management, user management, as well as administration. The Telematic services provide the communication interface to the connected vehicles. Every connected vehicle has a Wireless Communication Unit (WCU) installed in the car. The WCU hardware unit contains monitoring and diagnostics modules and enables in-vehicle data capture, including the GPS positioning and vehicle status information. The state of the WICE system is kept in the WICE database. The measurement data logged from vehicles is stored in the WICE file store, a large volume storage based on the data lake concept.

However, this approach has its limitations. To provide the required data, every vehicle needs to be additionally instrumented with an external acquisition system. This does not allow the OEM to expand the study to the whole vehicle fleet of real users. The OEM's employees who use instrumented vehicles and share the data might cause a bias, being often far more experienced in using support systems due to their work tasks and engineering background.

Moreover, currently no systematic approach regarding the use of vehicle sensors data has been developed, and vehicle data is not used to its full capacity. While vehicle data is extensively used for system performance verification, it is less used for driver performance evaluation and driving context assessment. Therefore, this research project can benefit engineers, further developing the tools and methods for more effective ways of data collection, data processing and data applicability.

2.2 USER BEHAVIOR

The definition of user behavior can vary from one research area to another. For example, in the field of mobile device research, user behavior is tied to user data collected from users' mobile devices. This data can in turn be used to reflect on user behavior.

In this research, the term user behavior is limited to the driver behavior. Under the driver behavior, we understand the set of actions taken by a driver interacting with the system in order to reach a goal or complete a task. Driver behavior can also be defined in relation to the characteristics of the context and the particular actions that are expected in that context in order to achieve a desirable outcome.

2.2.1 Usability

Usability is one of the important concepts when talking about driver behavior and driver interaction with the product. Interaction with the product that is easy to use and understand, increase users' productivity, decreases the learning process, and enhances the user satisfaction of the product itself. The main advantage for users is that they can perform their tasks easily and efficiently. Good usability of a particular product when comparing with similar products usually means the user would choose that product in the future. Good usability positively increases the reputation of the product and most likely would lead to an increase in sales. Therefore, the main goal for usability engineers is to construct a system or product that people find usable and will use (Ovaska, 1991).

Usability definition

Although the term usability is widely in use, there is no agreement on the exact definition (Paige et al., 2017). Different opinions regarding the way to measure usability together with the different fields where usability is practiced bring many similar definitions together (Folmer and Bosch, 2004). Nielsen (1993) describes usability as an aspect that influences product acceptance. He classified usability through five usability attributes: learnability, efficiency, memorability, errors and satisfaction. Nielson's classification, together with the definition from ISO/IEC 9241-11 standards (1998) that defines usability as *"the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use"* - are two of the most widely accepted definitions in practice. Thus initially, usability was more focusing on an outcome of interaction rather than on the quality of the product the user was interacting with. In the later standards (ISO/IEC 9241-210, 2010) usability is defined from two perspectives: as the quality of the product and as the outcome of interaction related to its quality in use (Paige et al., 2017).

According to the classification made by Bruno and Al-Qaimari (2004), usability consists of four common factors that have an impact on the whole interactive system: the user, the technology, the task, and the context of use. Consequently, the author adopted this definition in the research. The author assumes that only understanding of the user's behavioral model and technical limitations of the interactive systems within the specific context of the system's use, including the analysis of the influence of external conditions, can lead to the successful development of an interactive system that meets user's requirements.

Furthermore, according to Peham et al. (2014), usability could be described by the two following processes:

Learning process - the dynamic process that could be described as a process of gaining knowledge by studying, practicing and improving specific skills. Learning cannot be developed instantly but develops over time as experience increases. When the *learning process* comes to an end (when a user has reached stable user performance that doesn't change significantly over time), the *usage process* takes place.

Usage process - presumes that the driver has learned how to use the product and the Usage process measures how easy the product is to use once it has been learned. Figure 3 represents the improvement of user performance skills during the learning and usage processes.
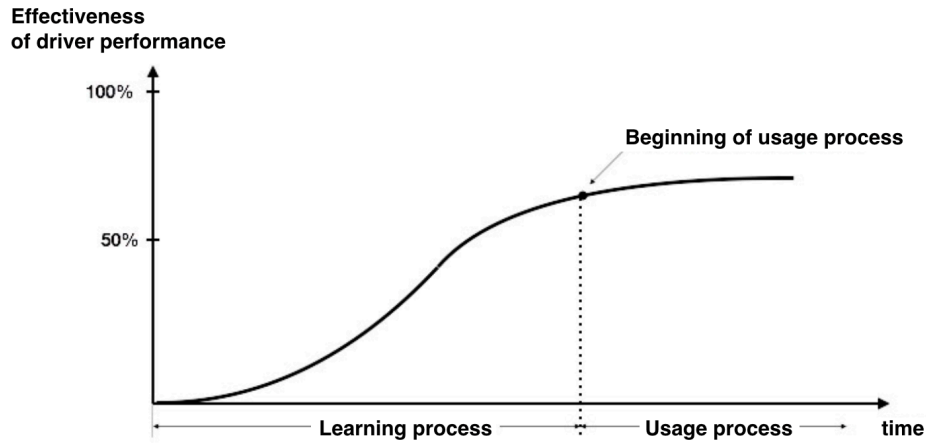
Figure 3. Learning and Usage processes in a User Performance development curve.

In this research, the author only focused on the usage process evaluation, since the learning process requires a different setup for the study where the previous driver experience is considered, and data collection starts from the first interaction with the evaluating system.

Usability attributes

To be able to measure usability, a definition of usability attributes is required. Many attempts have been made to define the list of attributes for usability assessment. However, no agreement regarding a unified view on usability attributes has so far been reached. The digitalization trend for complex products, such as mobile phones, computers or cars, has consequently increased the convergence between the computer science, telecommunication, and engineering fields. This has only boosted the complexity level and introduced new types of interaction intended to help a user in communication with technology (e.g., touch screens, voice commands, gesture interfaces). Such a dramatic change in the interface frameworks forced usability engineering to engage a large number of specialists within various disciplines. As a result, the usability attributes list can vary in different fields and for different products.

2.2.2   User experience

Another concept highly related to the presented research is the concept of User Experience (UX). According to ISO 9241-210, UX defines as *"a person's perceptions and responses that result from the use or anticipated use of a product, system or service"*. UX is the umbrella-term that *"takes a broader view, looking at the individual's entire interaction with the thing, as well as the thoughts, feelings, and perceptions that result from that interaction"* (Tullis and Albert, 2013). Roto et al. (2011) have a similar view on UX, simply defining it as experience generated by interacting with the system. Forlizzi and Bettarbee (2004) describe UX as people's interaction with the product and the overall emotions resulting from this interaction. The NN/g (2008) illustrated the UX scope by encompassment of the utility, usability, and desirability of the product (see Figure 4).

Although this research is heavily focused on interaction with the system, at this stage, the author avoids using the UX term since vehicle data does not support an assessment of driver perception, which is the part of overall emotions resulting from the interaction.

Nevertheless, the author positions this research in the field of UX. However, the author does not claim to have performed a data-driven assessment of UX. Primarily, it is the driver's interaction with the system in the long term that has been assessed, which allows us to reflect

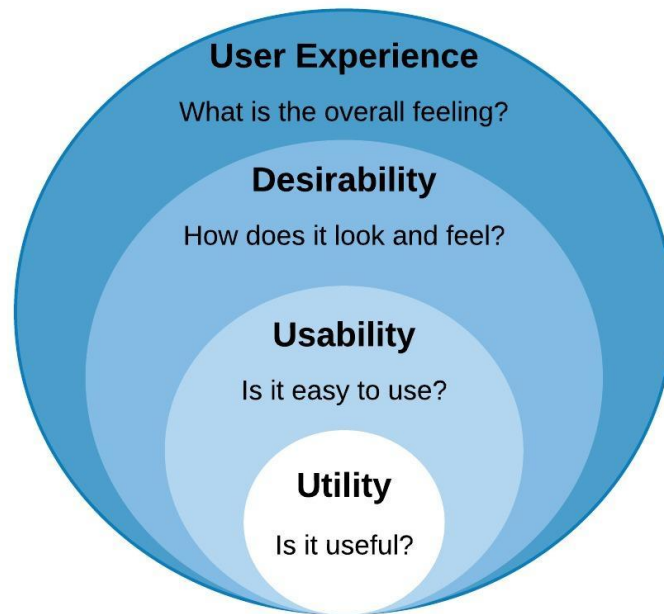on specific UX metrics that focus on driver behavior or attitude towards the evaluated subject/object.

Figure 4. The scope of User Experience (adapted from NN/g Conference, 2008)

2.3 METHODS FOR DRIVER BEHAVIOR ASSESSMENT

According to Ivory and Hearst (2001), there are 132 documented usability evaluation methods (UEMs), which were derived mainly for web user interfaces assessment. These methods are divided into five classes: testing, inspiration, inquiry, analogical modeling and simulation. However, if compared to the complex interfaces that include a physical interface in combination with a graphical interface placed in multi-mode screens, the number of applicable methods is very limited. In this research, the author is dealing with the already launched product, without the ability to change its design. Methods for this type of evaluation are usually narrowed to inquiry methods such as surveys, interviews or user feedback.

Nielsen (1994) suggested using several evaluation methods that increase findings regarding different usability issues and cross-checking the evaluation results. Two Comparative User Testing studies CUE-1 and CUE-2 (Molich et al., 1999) confirm Nielsen's suggestion by demonstrating the lack of consistency and systematic approach to usability evaluations. Those studies demonstrated that usability findings performed for the same project could vary dramatically, depending on the usability team's area of expertise and the methods that usability experts chose for the evaluation.

Moreover, in the engineering world, it is difficult to support the decisions on the results of subjective evaluation. Usability experts often feel undervalued in comparison with the other engineers able to support their decisions with objective evaluation. Having all these challenges in mind, usability experts are continually looking for ways to improve the usability assessment quality. In particular, they are most interested in bringing objective methods into the field. For that reason, the idea to utilize explicit knowledge at the data level is attractive to usability engineers. The objective data analysis can provide an understanding of users in a better way, by looking at the learnability or usage dynamics, evaluating individual or group behavior, detecting the usability issues, and measuring their severity.

2.3.1 Qualitative research approach

In the studies related to user evaluation traditionally a qualitative approach is applied. Qualitative research methods focus on the quality of things, trying to explain, describe and discover the root causes of user behavior (Creswell, 2014; Merriam, 2009). Denzin and Lincoln (2005) describe this research approach as an attempt to understand things in their natural environment, by interpreting the phenomena based on the meaning that a particular user or group of users bring to them. Qualitative methods usually focus on gathering subjective impressions regarding system usage, rather than targeting specific user tasks or identifying the variables that cause specific user behavior (Orlovska, 2018).

Main advantages of qualitative research

- Qualitative research is the most appropriate for situations when we need an explanation of why different things are happening, what their nature is and how they can be described.
- Deep, widespread evaluation is possible. Participants are usually able to freely express their opinions, which helps to build a discussion and elaborate on what they mean.
- The human factors in the form of user perception are the primary interest of qualitative studies.
- Occurring events can be observed in their natural context without reducing the complexity of system, processes or tasks.
- Qualitative approaches have a well-established methodology, based on UEMs, summarized by Ivory and Hearst (2001). These methods allow receiving user-related input at different stages of product development.

Limitations of qualitative research

- Due to the relatively low amount of participants, qualitative methods have no statistical significance, which means that the findings from the qualitative study cannot be extrapolated to the larger population sets with the same confidence level (Atieno, 2009).
- Frequencies of different issues detected through qualitative research are difficult to measure. As a result, rare phenomena can receive the same attention from the researcher as more frequent aspects (Atieno, 2009). A low amount of participants also reduces the possibility of classifying users or issues they experience.
- A qualitative researcher cannot be seen as an independent individual. (Rovai et al., 2013). Research techniques and environments (the lab or the questionnaire), as well as the researcher's own perception, can bias participants' view on the evaluated object and affect the interpretation of the results.
- Qualitative methods are often criticized for their low reliability. Different results may be achieved with various participants or at a different time.
- Qualitative studies are time-consuming. If stakeholders need to take an urgent decision, then probably the qualitative study that takes months to be administrated is not an option (Sallee and Flood, 2012).

2.3.2 Quantitative research approach

Quantitative research often focuses on measurements that test hypotheses, determine an outcome and generalize conclusions (Denzin and Lincoln, 2008). Quantitative studies may produce valid and reliable data due to the possibility to control the measurements with the help of specifically created technical solutions. Quantitative data could be obtained by quantifying subjective user input, taking from extensive user surveys, or using an automated method for data collection.

Main advantages of quantitative research

- Larger samples, compared to qualitative research, often make the conclusions from quantitative studies generalizable (Rovai et al., 2013).
- Statistical methods are primarily used in quantitative data analysis. Those methods are precise and rigorous, which helps to establish a certain level of trust in quantitative methods among engineers (Rahman, 2016).
- Quantitative methods are also useful when a systematic, standardized measurement is needed.
- Quantitative research is independent of the researcher, and therefore, the evaluation process is less biased by the interviewer's viewpoint, his/her appearance or questions (Rahman, 2016).

Limitations of quantitative research

- Due to the reduced data feasibility, it is often not possible to measure the full complexity of human experience or perceptions. Therefore, the user experience can be divided into the measurable areas and studying them as parts (Rovai et al., 2013).
- Quantitative research allows seeing what things happened and how frequently they happened but cannot determine underlying explanations of why those things happened (Bouwer et al., 2014).
- The use of qualitative methods may give the wrong impression of homogeneity in a data-set. For example, the measured user experience of vehicle-owners might not be applicable to non-vehicle-owners. Therefore, some applications of qualitative methods may require clarification for homogeneity within the group.

2.3.3 Current trends of driver behavior assessment in the automotive industry

Despite the fact that both qualitative and quantitative research approaches are broadly applied in the automotive industry today, many studies are still conducted in isolation. Different evaluation groups of designers/engineers with diverse backgrounds are usually conducting studies that are based solely on qualitative or quantitative data, resulting in low cross contribution from one study to another.

The validity of results for these kinds of studies is always questionable, and therefore the need to combine different approaches is clearly recognized. Nevertheless, the results of different approaches are mainly used for the comparison or validation of their findings but do not aim to improve the quality of the studies. This could be explained by the low compatibility of qualitative and quantitative data, which often leads to the practice of prioritizing one of the approaches over another. A qualitative approach is mainly applied in user-related studies, due to long-term traditions within automotive OEMs. Quantitative research methods, in turn, are broadly used for the evaluation of a vehicle's mechanical parts and software, but are rarely in user-related studies.

However, the rapid development of objective data sensors and the variety of information generated by the automotive production platforms clearly indicates a need for a new methodology that considers both directions: extended quantitative data possibilities utilization and qualitative insights.

Since both quantitative and qualitative approaches have their strengths and drawbacks concerning user studies, an intelligent fusion of both approaches, implemented effectively, can improve the quality of user studies and increase the validity of the results.

While the mixed-method approach is widely described in the literature, the author's

understanding is similar to Johnson and Onwuegbuzie (2004) that defines this as a type of research where the research team combines qualitative and quantitative approaches to achieve in-depth understanding and validation of the results. Moreover, Greene (2007) states that effectively designed mixed-method research can "...offset inevitable method bias".

2.3.4   Naturalistic Driving Studies

The key approach adopted in this research can be characterized as a Naturalistic Driving (ND) study. A ND study usually refers to a study that is not constrained by a strict experimental design where the data is acquired for a relatively long-term period, in the natural driving context and under various driving conditions happening in a natural way. Data in ND studies is collected mainly from vehicle sensors data, GPS, vehicle's apps, and/or data from video cameras (Fridman et al., 2019). Vehicles are instrumented in the most unobtrusive way, allowing users to perform driving activities undisturbed. The sensors data is collected and processed with the help of wireless technologies. The data collection is systematic, within the timespan of several months and includes each single driving activity. The advantage of this approach is that the driver is not limited in his/her movements, time and frequency of driving. The driver uses the vehicle in his/her own way which is extremely important to create a natural environment for the ADAS user behavior evaluation.

The EuroFOT (European Field Operational Test) was one of the first large-scale projects focused on the investigation of possibilities to enhance safety, and reduce the environmental impact of vehicles instrumented with ADAS (Benmimoun et al., 2013). Another project, named 100-Car naturalistic driving study, was conducted on the US market with the aim to evaluate driver safety in crash and near-crash situations (Neale et al., 2005). Currently, on-going, the MIT Autonomous Vehicle Technology (MIT-AVT) study, which was launched in September 2015, seeks to understand how driver-vehicle interaction can be designed to be safe and enjoyable (Fridman et al., 2019).

The above described NDS have inspired a number of programs and organizations, such as SCOUT (Safe and COnnected AUtomation in Road Transport, 2019), CARTRE (Coordination of Automated Road Transport Deployment for Europe, 2019), SAFER (THE SAFER organization, 2019), SHRP2 (Strategic Highway Research Program 2, 2019), ADAS&ME (Adaptive ADAS to support incapacitated drivers Mitigate Effectively risks through tailor-made HMI under automation, 2019) and others. These initiatives aim to support the research field by exploring and developing the potential of ND studies further. The majority of the projects are supported by governmental organizations and focuses on the driver and traffic safety issues, investigating the driver behavior in crash and near-crash situations (Sander, 2017; Hatfield et al., 2017; Engström et al., 2018). The context-aware evaluation of driver behavior in the preceding moment is a critical factor in these studies, enabling investigation and explanation in detail of the driving behavior before the incident happens. Liang et al. (2016) underlined the importance of driving context analysis for detecting abnormal driver behavior, aiming to quantify the risks associated with various driver behaviors. Zhai et al. (2018) emphasize the importance of context-aware driver behavior evaluation, showing that the integrating of driving context provides reliable results regarding the driver behavior evaluation on the road. According to Papazikou et al. (2017), as well as Tivesten and Dozza (2014), the driving context is one of the most important factors for user behavior evaluation. Both conclude that the context might affect driver behavior, both in a positive and negative way. Further, Ahlström et al. (2018) emphasize the effect of the road environment on the development of driver sleepiness. Ahmed and Ghasemzadeh (2018) designed an automated method for heavy rain detection and measured the impact of heavy rain on driver behavior, discovering a correlation between the driver's age and the speed chosen under heavy rain conditions.

The mentioned research shows that the driving context needs to be synchronized with driver behavior and with the vehicle's state data, regarding the evaluated function. Moreover, the whole complexity of the driving context needs to be considered, since even one underestimated variable is able to alter all the results and our understanding regarding the collaboration between the driver and the evaluated system (Fridman et al., 2019).

Further, when performing a data-driven evaluation of a semi-automated system like ADAS, where the driver and system interrelate to each other, users can also be unaware of system functionalities or be influenced by factors that the developers are not aware of. The incorporation of vehicle data into driver behavior evaluation can help to reveal these misconceptions between the driver and the system (Brannen, 2005), which will enable an understanding of where problems in the interaction emerge, and if the problem requires attention, leading to an appropriate and sufficient validity evidence (Creswell and Miller, 2000).

However, it has to be acknowledged that sensors and the data acquisition systems are not sufficiently developed to assess the entire driving event. The driver remains an essential part of the driving event evaluation. In order to receive insights on driver perception and other human-related aspects, a combination of quantitative and qualitative results is needed to clarify the reasons why drivers are using ADAS the way they do.

3

RESEARCH APPROACH

The primary purpose of this research is to generate new knowledge that is valuable for both the academic society and for current engineering practices. This research is focused not only on providing insights for the people dealing with the investigated phenomena in practice but also contributes by the designing of methods for more effective application of newly generated knowledge.

 This chapter describes the research approach applied to the presented thesis. It provides the motivation for why the particular methodology was chosen and how it was adopted for the needs of this research. A specific focus is set to clarify the relations between studies, appended publications, and the research questions investigated.

3.1 DESIGN RESEARCH

Many definitions of the design research exist, depending on the application background. Research design, referred to the engineering context, is usually described as a set of purposeful activities that help to develop a product from a need to its complete realization. According to Blessing and Chakrabarti (2009), *"design is a complex, multifaceted phenomenon, involving people, a developing product, a process involving a multitude of activities and procedures; a wide variety of knowledge, tools and methods; an organization; as well as micro-economic and macro-economic context."* Hubka and Eder (1987), defined design science as *"the problem of determining and categorizing all regular phenomena of the systems to be designed, and of the design process. Design science is also concerned with deriving from the applied knowledge of the natural sciences appropriate information in a form suitable for the designer's use."*

 Research design can be considered to pass through three evolutionary phases: Experimental, Intellectual and Empirical (Wallace and Blessing, 2000). During the Experimental phase, which existed until the late 1950s, the activities and the experience of the senior designers were most valued. However, their observations in the design process were relevant to the specific domain they described and functioned within one technical domain, and therefore could not be applicable to the broader context. During the Intellectual phase stage, the emphasis was placed on the creation of a design basis using a variety of methodologies and principles of a design process. The Empirical phase started in the 1980s when the number of studies where empirical

data was gathered began to grow. The purpose was to understand how the designers performed the process of design. The Empirical phase investigated what impact new methods and tools had on these processes (Blessing and Chakrabarti, 2009).

3.2 AVAILABLE THEORETICAL FRAMEWORKS

Different theories and frameworks were introduced within the design research field, providing a theoretical basis to research in the product development domain. In particular, the following research approaches were introduced: Theory of Technical Systems (Hubka and Eder, 1987), Domain Theory (Andreasen, 1991), TRIZ (Altshuller et al., 1999), Axiomatic design (Suh, 2001), CK-Theory (Hatchuel and Weil, 2003), Function-Behavior-Structure framework (Gero and Kannengiesser, 2004), Design Research Methodology (Blessing and Chakrabarti, 2009), Mathematical Theory of Design (Braha and Maimon, 2013), and others. Despite the fact that different frameworks were introduced in the field of engineering design, there was no strict recommendation for the use of one method over the other. The above-described frameworks demonstrate a different level of applicability in the research projects, depending on the traditions of the university department and research group.

3.3 METHODOLOGY APPLIED IN THIS THESIS

The research presented in this thesis is based on the Design Research Methodology (DRM) framework proposed by Blessing and Chakrabarti (2009). DRM is intended to fulfill two purposes: first to understand the investigated phenomena and second to submit the tools, methods, or guidelines that can be introduced in practice. In this research, the author applied DRM, since this research is based on the research tradition of the university department and research group. DRM has strong relevance to the field of mechanical engineering and product development. Moreover, the DRM provides a context to position the research and encourage the reflection on the research approach and the choice of research methods, allowing the researcher to find new ways to deal with the investigated phenomena.

The framework consists of four main stages: *Research clarification, Descriptive study 1, Prescriptive study 1, Descriptive study 2* (see Figure 5). At the *Research Clarification* stage, the current understanding needs to be clarified, where the overall research aim is understood, the research questions are set, and the research plan that supports the work at consequent stages is provided. Consequently, *Descriptive study 1* aims to develop an understanding of the research phenomena and its influencing factors. At the next stage, *Prescriptive study 1*, takes into account the knowledge of the research phenomena generated in the Descriptive study 1 and aims to develop the new methods or tools that support the improvement of the existing model. *Descriptive study 2* aims to evaluate the applicability and effectiveness of the proposed design modification, focusing on the impact evaluation. To evaluate the contribution, the success criteria that directly reflect on the desired research goals need to be set. These criteria will be used to judge the research outcome against set goals.
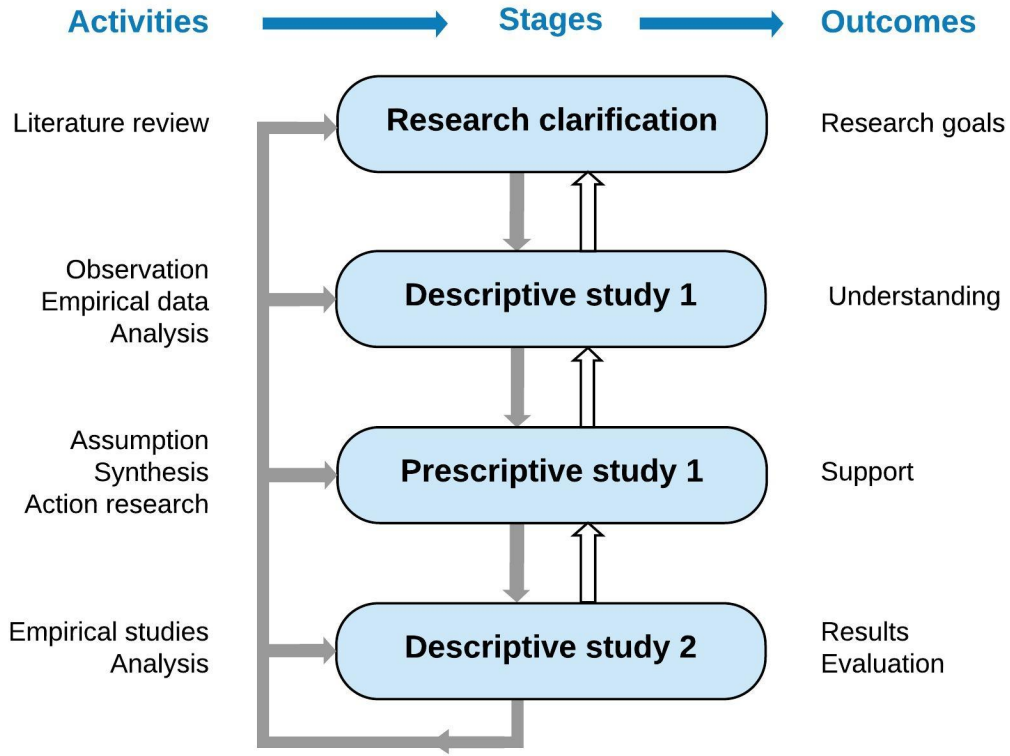
Figure 5: Overview of DRM Framework stages.

One of the advantages of the DRM Framework is that the implementation of the stages is not necessarily sequential or linear. Multiple iterations are possible between stages and within every research stage, providing the flexibility of the DRM Framework to fit any specific research project. This flexibility allows the researchers to look for a variety of new ways for the phenomena investigation and not directly follow the prescribed form. In this research, the DRM Framework was adopted due to its strong connection to the engineering field providing a robust methodology basis aiming, through the understanding of investigated phenomena, to propose an improvement model that can be applied and verified in practice.

To verify that the goals of the research are achieved, it is necessary to identify the success criteria. Success criteria, according to Blessing and Chakrabarti (2009), relate "to the ultimate goal to which the research project intends to contribute and usually reveal the purpose of the research." This research aims to design a reliable method for vehicle data-driven evaluation to assess the driver behavior of ADAS. The ultimate goal of this research is to integrate a vehicle data-based assessment into existing practices to enhance the quality of the obtained results.

The following sections describe how the research methodology was applied in the course of this thesis. The author will explain which research questions were investigated in the appended papers, what research methods were used, what type of results were achieved, and how this work was distributed among DRM main stages.

3.3.1 Research questions and DRM phases

The main goal of this research is to achieve an understanding and generate new methods for effective vehicle data utilization in driver behavior evaluation. To achieve this goal, a number of research questions were specified.

RQ1 *(How can vehicle data be used for data-driven user behavior evaluation?)* focuses on the understanding of main interdisciplinary concepts for this research project: Big Data and User

Behavior. Thus, **Paper A** focuses on how Big Data is explored and what the main challenges for its utilization in the automotive industry are. As a next step, **Paper B** investigates the applicability of Big Data to the Usability framework. The explorative research approach is deliberately chosen for both studies, including interviews with industrial partners, work with the OEM's internal documentation, focus group discussion, and several workshops. While **Paper A** primarily focuses on the feasibility of the vehicle data, **Paper B** investigates the possibility of merging two concepts: Big Data and Usability.

RQ2 (How can the data-driven approach be incorporated into existing methods for driver behavior evaluation?) is about how to position the objective data-driven analysis in the evaluation process, so that it is accepted in practice. **Paper A** and **B** perform the initial attempt to position the objective data analysis in the overall process of driver-systems evaluation in the driving context. However, **Paper C** contributes with the design of the mixed-method approach, for the complete incorporation of the quantitative approach into the qualitative assessment of in-vehicle systems.

RQ3 *(How can the validity of the data-based results be achieved?)* refers to the quality of data signals, data acquisition, data processing, and data analysis processes, as well as the assumptions made based on vehicle data analysis. In other words, can we extract reliable knowledge and create assumptions about driver behavior based on vehicle data analysis? **Paper C** focuses on the design of the mixed-method approach which controls the validity of results with the cross-validation approach. **Paper D**, with the use of the mixed-method design, provide comprehensive results on driver behavior evaluation regarding the use of ADAS functions and the main factors influencing this behavior. These papers made the first step to validate the findings of the quantitative approach by comparing the assumptions made based on vehicle data analysis with the interview data of the same participants. Nevertheless, more studies need to be done before the validity of the quantitative methods can be confirmed.

RQ4 *(How can vehicle data be used for the data-driven design of vehicle systems?)* is considered as the next step for this research. Therefore, this question has not yet been addressed in the corresponded papers.

According to Blessing and Chakrabarti (2009), a DRM Framework should not be interpreted as a strict and linear research process. The allocation of the research questions in the DRM framework has particular reasoning. The exploratory studies performed in this research were focused on defining the relevant data set for user-behavior evaluation. The results showed some technical and conceptual limitations that restricted vehicle data utilization. Thus, the descriptive study was focused on developing a method for comprehensive evaluation of the support systems, where the qualitative and quantitative approaches could contribute to insights of each other, allowing a complete understanding of how in-vehicle systems are used to be achieved. The distribution of the appended papers in the context of DRM phases is depicted in Figure 6.
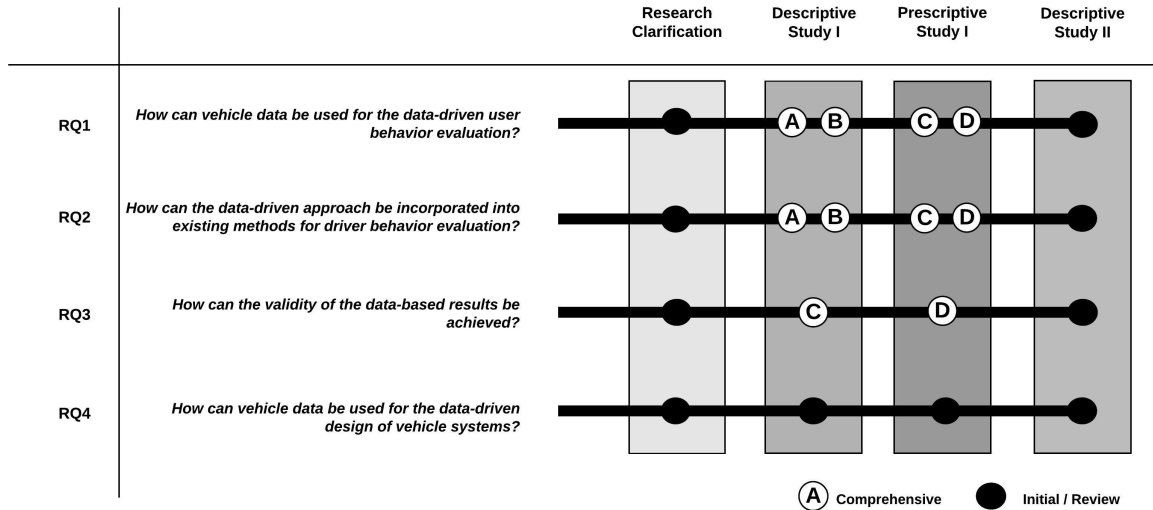
Figure 6: Distribution of papers A-D in the context of the DRM Framework.

3.3.2 Type of results

Several types of results were achieved in the presented research. Descriptive study results provide empirical and statistical data that lead to a better understanding of how vehicle data analysis contributes to user behavior evaluation. As a result of **Paper A**, the case study for ADAS assessment was designed, and the first prototyping outcomes, illustrating the ability of data to answer specified usability questions, were presented. **Paper B** resulted in the identification of the area within the usability attributes range where vehicle data can contribute. The list of attributes was presented together with the number of limitations identified for the usability assessment. Consequently, **Paper C**, based on the results of **Paper A** and **B**, presented the design of the mixed-method approach for comprehensive ADAS evaluation. Prescriptive results, based on the mixed-method design, were proposed in **Paper C**, and the comprehensive ADAS evaluation was applied at the OEM. **Papers D** confirms the applicability of the mixed-method and provide reliable results regarding the complex ADAS evaluation.

3.3.3 Methods used

There are numerous approaches for collecting data within design research, such as samplings, interviews, group interviews, surveys and observations, and others. The methods used in the course of this research are presented in this section.

A systematic **literature study** was performed in this research. The main goal was to understand the knowledge foundation, to be able to map the proposed methods and definitions, as well as to identify any existing gaps in the knowledge related to data-driver user behavior evaluation. Moreover, an extensive study of the OEM's internal documentation studies was performed. Among others, the documents mainly consisted of adopted attribute structure and its detailed descriptions, lists of functional and technical requirements, and methods for these requirements evaluation, technical reports, and plan of operations. The author participated in internal follow-up meetings, organized weekly, observing the practical approaches to the investigated phenomena.

Interview studies are typically classified as structured, unstructured and semi-structured interviews. According to Yin (2013), interview is a widely used method in qualitative research, aiming to collect respondents' subjective opinions on the investigated issues. In this research, only semi-structured interviews were performed. Semi-structured interviews include elements from both structured and unstructured interviews and "a fixed set of sequential questions is used

as an interview guide, but additional questions can be introduced to facilitate further exploration of issues brought up by the interviewee, thus almost taking a form of a managed conversation." (Cachia and Millward, 2011). Consequently, all interviews were transcribed verbatim, then were coded and analyzed with the help of qualitative data analysis software NVivo 12 (NVivo, 2019). Two independent coders examined the first transcript to identify different themes or nodes. In the next step, the themes were reviewed and discussed in order to determine coherence and minimize subjective discrepancy. After that, the interviews were coded by each researcher separately and a final session was held, where the open questions and themes were discussed, to review the quality of the coding.

A **field study** is a universal method for collecting data about users, user needs, and product requirements that involves direct or indirect observation and interviews. Normally the data is collected about task flows, user performance, detected issues and any types of inefficiencies in the user environment (Rosenbaum, 2002). Studying driving behavior in the dynamic context is a fundamental characteristic of Field Operational Tests (FOT) and ND studies. The ND study usually refers to the study not constrained by a strict experimental design, where the data is acquired for a relatively long term period, in the natural driving context and under various driving conditions happening in a natural way.

A ND study was designed and performed in the course of this research. Vehicles were instrumented in the most unobtrusive way, allowing drivers to perform driving activities undisturbed. The vehicle sensors' data was collected and processed with the help of wireless technologies. The data collection was systematic, within the time span of seven months and included each single drive cycle. The ND data included a combination of CAN bus data, GPS data, cloud data, external data that is provided through additional applications (e.g., the navigation data) that affect driver behavior or system performance. The data analysis was conducted with Power BI software for statistical analysis (Power BI Microsoft, 2019). The data was analyzed in four different layers of abstraction: single drive cycle (DC) evaluation layer (if something indicated unusual or interesting user behavior that needed in-depth investigation), one-driver evaluation layer (focused on in-depth user behavior evaluation of the same driver), groups comparison layer (based on the comparison of user behavior between different user groups), and overall assessment layer (based on the average calculation for all users). A detailed description of these abstraction levels can be found in the Methods chapter of Paper D. In general, the ND data analysis was based on the use of statistical methods that provide statistical significance and reliability to the obtained results. A more detailed description of how the above-described methods were integrated into the research design in the appended papers is described in   Chapter 4.

3.3.4   Validating the results in applied research

To establish the quality of the research, validation and verification of the results and methods are required. In the context of engineering design, verification refers to internal and external completeness and consistency, whereas validation refers to the justification of knowledge claims (Barlas & Carpenter, 1990). Nanda et al. (2000), stated the need for an interdisciplinary approach to address complex problems in the research.

Thus, to achieve the validity of the results in this research, a cross-validation approach was used. *Cross-validation* refers to the procedure by which sets of scientific data generated using two or more methods are critically assessed. Cross-validation can have two dimensions: analytical data validation and method validation. Analytical data validation in this research is supported by the sequential mixed-method approach (used in **Papers C** and **D**), which helps to cross-validate the findings comparing the results of qualitative and quantitative evaluation.

The method validation can be performed by *Validation by acceptance* that focuses on having new scientific contributions accepted by scientific and industrial experts within the field.

Adoption of the method in the industry, and publishing the method in scientific journals are the first indicators of validation by acceptance.

Verification of the results can be ensured by *Logical verification*, which provides the analysis of coherency, completeness of results, and consistency of internal/external elements. Validation and verification of the results from this research will be further discussed in section 5.2.

4

RESULTS

This chapter presents the core findings from the papers appended to this thesis. The main focus however, is given to the achieved results. For more detailed information, please refer to the full-texts of the papers at the back of this thesis.

4.1 PAPER A

4.1.1 Purpose

This research project was initiated from the hypothesis that big data can be utilized for driver behavior evaluation providing a better understanding of user behavior and enhancing the quality of human-machine interaction evaluation.

The purpose of the first study was to investigate the real-time data availability, its applicability for user-related studies, and limitations in the automotive sector. The author also looked for user behavioral data that could be used to contribute to better user understanding and evaluation in an automated way. The overall goal was to confirm with engineers the ability of vehicle data to provide useful insights regarding driver behavior evaluation and to position the objective assessment into the overall ADAS evaluation, contributing to better design of the evaluating methods.

4.1.2 Method

This study was designed as a combination of different methods for qualitative and quantitative research. Within the framework of qualitative research, semi-structured interviews were conducted, followed by organized focus group discussions and workshops. The qualitative study in this paper aimed to clarify the use scenarios, identify the relevant data, and verify the data feasibility for the user behavior evaluation of ADAS in the real-driving environment. As for quantitative research, a pre-study to verify the approach for the user behavior evaluation of ADAS was conducted. Preliminary data analysis was performed with the focus to identify potential usability issues. The overview of the methods used and the main research structure can be seen in Figure 7.
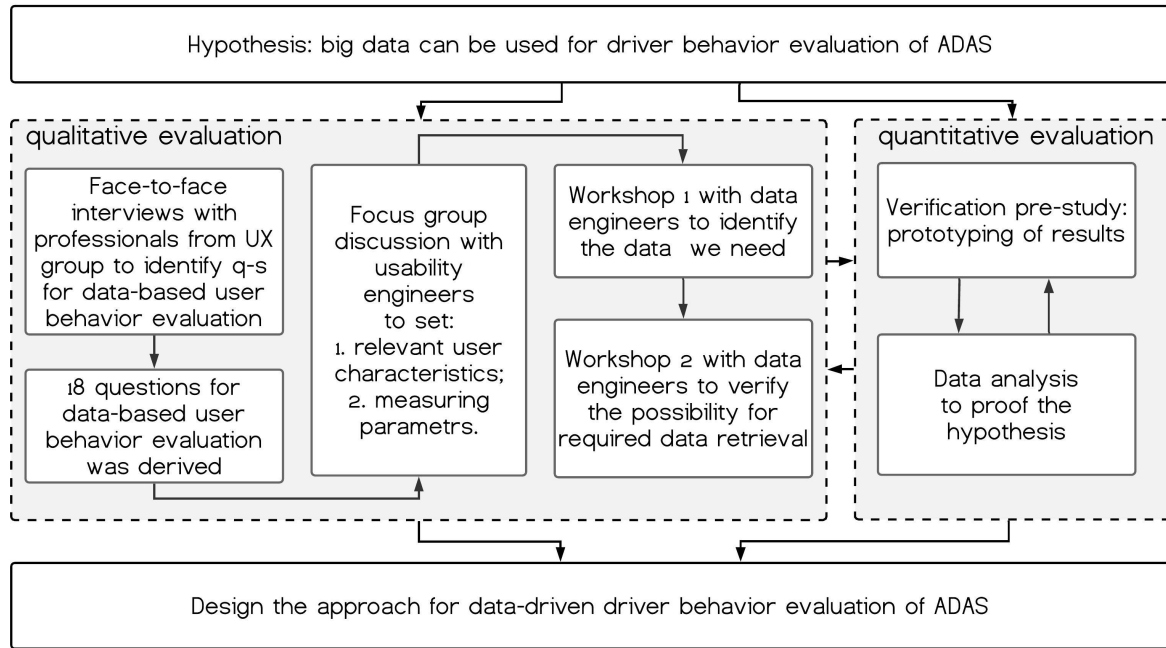
Figure 7. Paper A: method design.

4.1.3 Main results

This paper concluded that big data and technological advances have an undeniably great value for understanding user behavior in the future. However, big data is not knowledge. The need for the development of methods regarding classification and sophisticated extraction of the relevant information for successful product design is clearly identified. Moreover, this transformation cannot be done throughout one discipline alone. Specialists from several disciplines must be involved (e.g., software/hardware developers, data analysts, UX experts, and others), where the applied approach has to be combined with the theoretical. In this study the presented design of the method for data-based user behavior evaluation of ADAS was verified through the pre-study results and was validated by industry professionals.

According to the designed approach for data-driven user behavior evaluation, big data analysis can serve as a basis for qualitative assessment. Figure 8 illustrates the main steps of the data-driven approach for driver behavior evaluation: (i) define the evaluation questions/inquiries; (ii) identify the required data; (iii) determine the dataset, considering the current limitation for data acquisition, and define the measuring parameters: i.e., the size of the study, the number of trials, the time frame for measuring, specific user parameters, if any; (iv) collect the required data according to measuring parameters; (v) analyze collected data to answer the evaluation questions; (vi) perform an evaluation to determine if any hidden knowledge can be extracted (or confirm the known hypotheses); (vii) cluster users regarding their behavior and measure the magnitude of detected usability issue within clustered groups to understand if the detected issue is essential for a particular evaluation question; (viii) evaluate the ability to include and consider as many related data signals as identified to learn more about the detected issues until the level of uncertainty is reached; (ix) conduct a subjective assessment, which can be based on methods of customer polling or data-based expert evaluation.

Specifically, the quantitative loop shown in Figure 8 is based on the utilization of purely vehicle data and aims to detect potential usability issues and measure the magnitude of the issues by analyzing the number of affected vehicles and the frequency of the detected problems. Such objective results can inform the evaluation team where to focus their attention during the following interviews with the car drivers.
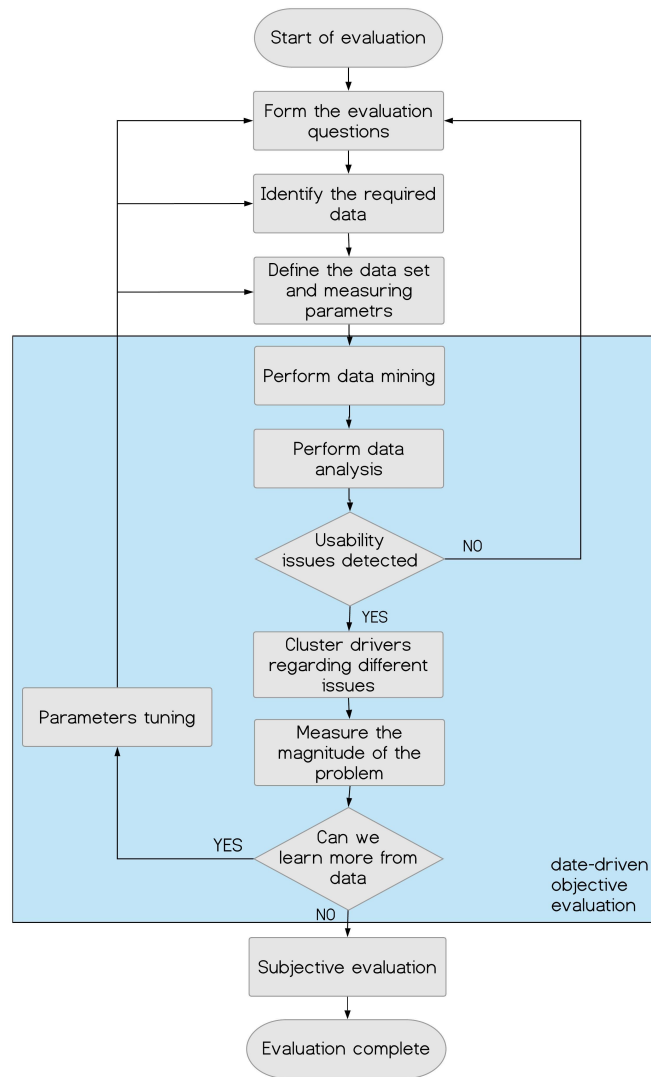
Figure 8. Paper A: Data-driven approach for driver behavior evaluation of ADAS.

The first prototyping results demonstrated that big data use can contribute to the detection of usability issues, allowing their magnitude to be measured by clustering users with similar behavior. For example, the preliminary results indicated that the PA function was used significantly less compared to the ACC. Since the only difference between the two ADAS functions is the steering capabilities that added to the PA function, it becomes evident that most of the users cannot accept the PA function. This behavior can indicate a potential usability issue. Two assumptions, derived solely from the quantitative data, were further investigated by qualitative methods: (1) drivers do not trust the system enough to delegate the steering function; (2) PA, as a function, does not provide sufficient quality. Utilizing this type of analysis the OEM can obtain an understanding of how ADAS can be improved: the trust in the function needs to be developed, or the ADAS function itself needs to be more robust under the different driving conditions.

The benefits of big data usage for the OEM were also evaluated. The most important benefits are the potential of getting more accurate feedback on user behavior, compared to the user feedback obtained with "classic" qualitative only methods, and the possibility to evaluate the user behavior with a reflection on the system performance. Analyzing the user and system performance together can help to identify the hidden user mistakes or detect the system imperfection that needs to be improved.

27

This paper, with the title: *"Big data usage can be a solution for user behavior evaluation: an automotive industry example"* was presented as a podium presentation at the *CIRP-CMS 2018 conference in Stockholm and published in CIRP Procedia 2018.*

4.2 PAPER B

4.2.1 Purpose

As a logical continuation of the research described in Paper A, where the feasibility and applicability of big data was investigated, in this paper the methods for Usability Evaluation of real-driving environments were evaluated. The primary purpose of this research was to identify the applicability of vehicle data for usability attributes assessment.

Usability evaluation focuses on how well users can understand and use the product to achieve their goals. Usability also measures the level of user satisfaction and user perception of the product. To gather this information, a variety of usability evaluation methods (UEMs) were used. However, current UEMs are not always suitable for usability attributes assessment in real driving activities due to the complexity and dynamic environment that changes all the time. Drivers are often unable to provide sufficient information on the matter.

On the other hand, engineers need more precise and rigorous data than qualitative insights to base their decisions on. They are used to working with numbers, and it is often challenging to translate subjective input into the terms and measurements that engineers can work with. Therefore, in this research, the applicability of the quantitative user data for usability attributes assessment was investigated. Since the quantitative evaluation is based on the use of statistical methods, the delivered results have a high level of precision and can be better translated into the engineering language. Therefore, the results of quantitative data analysis can be clearer to the engineering teams involved in ADAS development and evaluation.

4.2.2 Method

This study was designed as a combination of different methods for qualitative research. ACC and PA systems were chosen as examples for the support systems evaluation, due to their novelty and ambiguity of use. An overview of the applied methods for the study design presented in this paper can be seen in Figure 9.
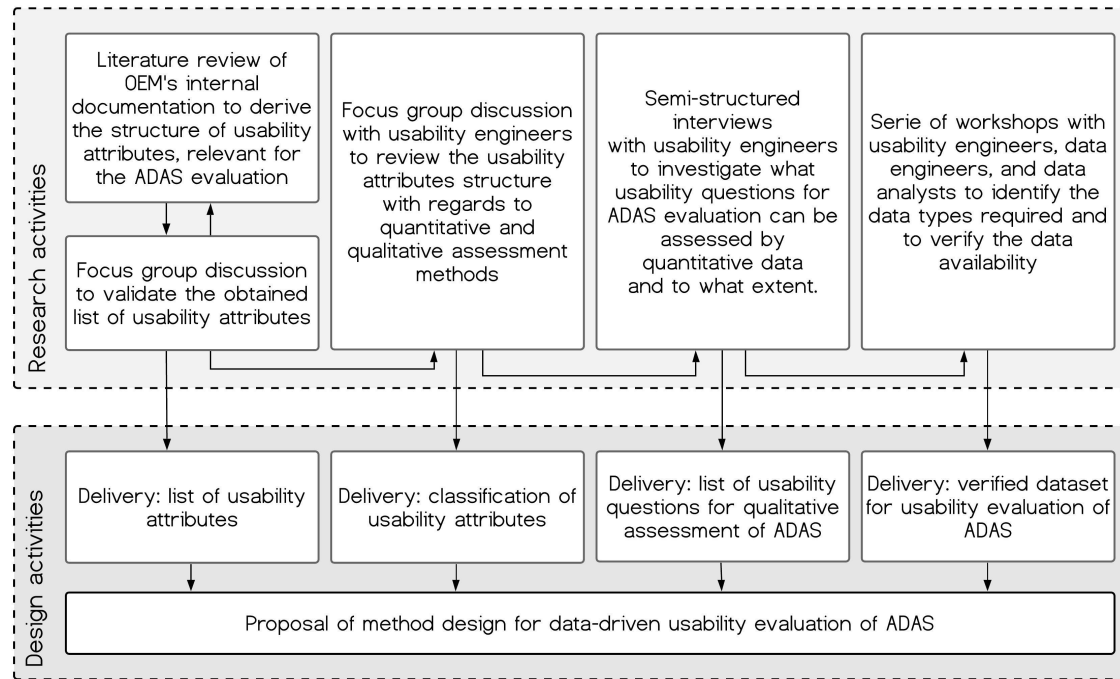
Figure 9. Paper B: an overview of research activities and deliveries.

The research project was based on the internal study of OEM documentation regarding the usability attributes relevant to the assessment of the evaluated system. As a result step, 16 usability attributes relevant to the usability evaluation of ACC and PA systems were identified together with their definitions accepted by the OEM. In the next step, these attributes were validated through a discussion session with the usability engineers involved in the ADAS evaluation process. Subsequently, the focus group discussion was organized to restructure usability attributes considering the ability of vehicle data to support the assessment of driver behavior. As a result of this step a new classification of usability attributes was proposed (see Figure 10).

Figure 10. Categorization of usability attributes regarding user performance, system performance and user perception.

Furthermore, semi-structured interviews with nine usability engineers at the OEM were conducted with the primary purpose of identifying the measurable area for quantitative usability attributes assessment. As a result, a list of 18 usability questions for the quantitative evaluation was designed.
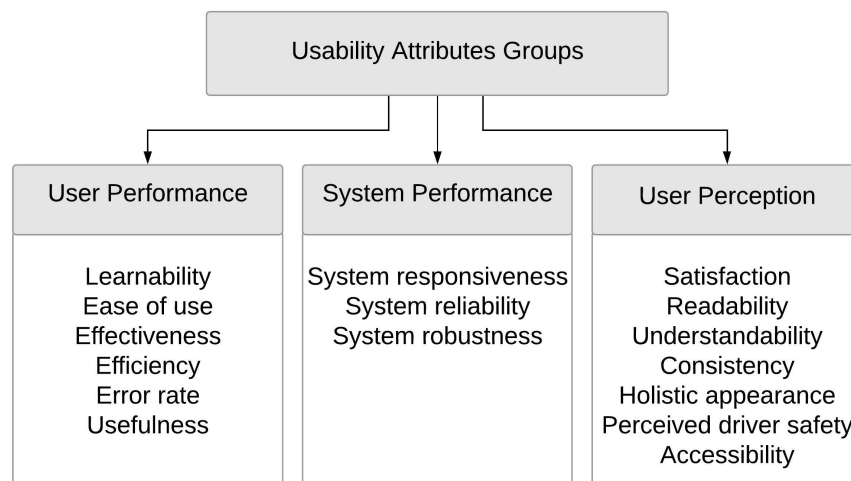
In the final stage, a few workshops with usability engineers, data engineers, and data analysts were organized. The focus was set on the verification of the ability of quantitative data to answer usability questions set in the previous step.

During the study design activities, the deliveries from all research activities were considered. The essential user characteristics for the study were added and different measuring parameters were defined.

4.2.3 Main results

The study was performed with the primary purpose of eliciting the identified usability attributes list (see Paper B, table 1) and to evaluate possible pitfalls of usability attributes assessment for the released product. The study disclosed the evidence that subjective usability attributes assessment combined with data measured in the real driving environment, can significantly improve current usability attributes evaluation practices. It was found that vehicle data can contribute to the User Performance and System Performance usability attributes assessment. The proposed design for the comprehensive evaluation of usability attributes is depicted in Figure 11.

Quantitative data analysis, performed with statistical methods for data analysis, helps to minimize the human effect on the achieved results and, therefore, this approach can increase the quality of usability evaluation.
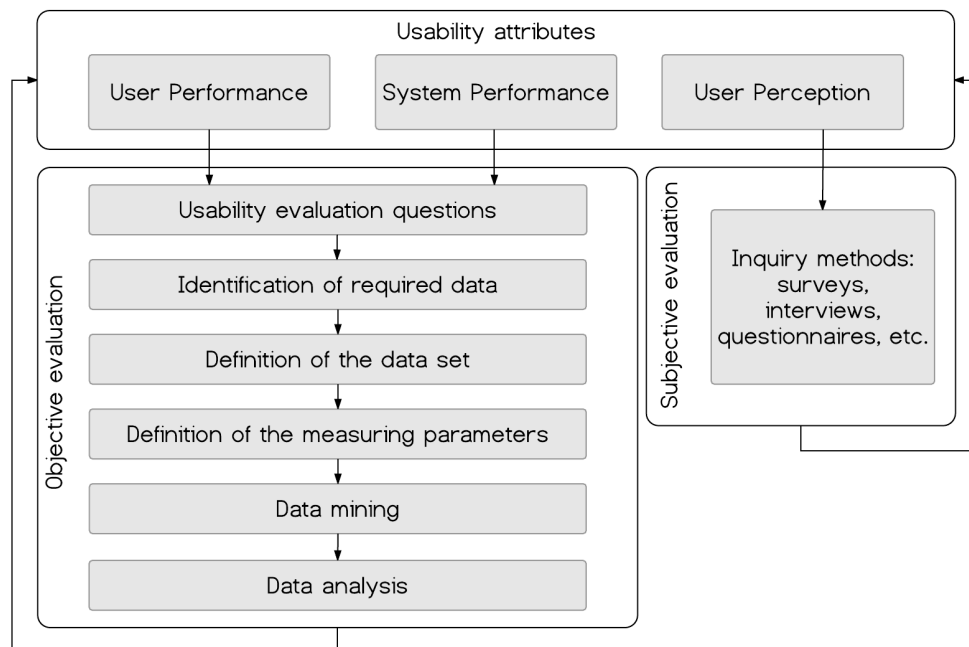

Figure 11. Case study design for usability evaluation of ADAS.

However, the study revealed that quantitative methods could not be used to assess User Perception attributes such as user satisfaction, holistic appearance, perceived driver safety, and others. Since vehicle data in this project presumed to measure data from users in the real-driving

environment, no kind of advanced technologies measuring psychosomatic parameters that could be used at the usability LAB was available for use in the real driving environment.

In this paper, the applicability of big data analysis for usability attributes assessment was investigated. The ability of vehicle data to feed the existing usability attributes assessment by vehicle data was proven by presenting the prototyped results. Thus, the proposed approach can increase the usability issues detection, allowing measurement of their magnitude by clustering users with similar behavior. The industry professionals validated the case study design for ADAS assessment.

A conference paper with the title *"Big data analysis as a new approach for usability attributes evaluation of user interfaces: an automotive industry context,"* was presented as a podium presentation at the *Design 2018 Conference* in Dubrovnik, Croatia and was published in *Proceedings of the DESIGN 2018.*

4.3 PAPER C

4.3.1 Purpose

In this research, the author continuously stresses the point that both qualitative and quantitative approaches can effectively support the user behavior evaluation process. Different types of user data used in these approaches contribute to different types of knowledge with regard to the understanding of user behavior. Traditionally qualitative methods are more often used for user behavior evaluation. However, continuously increasing in-vehicle connectivity opens new capabilities for obtaining new objective usability data. Therefore, the combination of qualitative and quantitative methods can lead to a better research approach than the use of each of them in isolation. In our next study, we propose a new, improved methodology design for user behavior evaluation. We take ADAS as an example of the vehicle systems for the assessment and design the method based on intelligent quantitative and qualitative approaches consolidation, aiming at substantial improvement of methodology design for usability behavior evaluation, as well as improving the validity of the results.

4.3.2 Method

In this paper, an imperative study was performed. Based on the findings of papers A and B, where the vehicle data abilities and limitations to support the usability evaluation process were investigated, the basis of the current design was formed. Analysis of data feasibility and current practices at the OEM helped to reflect on the effectiveness of the methods in current practice and led to the proposal of the explanatory sequential design for vehicle support systems evaluation. Thus, the explanatory sequential mixed-method design became a logical reflection on previous studies.

4.3.3 Main results

The proposed explanatory sequential mixed-method design for vehicle support systems includes four major phases: (i) initial setup of the study; (ii) quantitative evaluation; (iii) qualitative evaluation; (iv) feedback loop (see Figure 12).
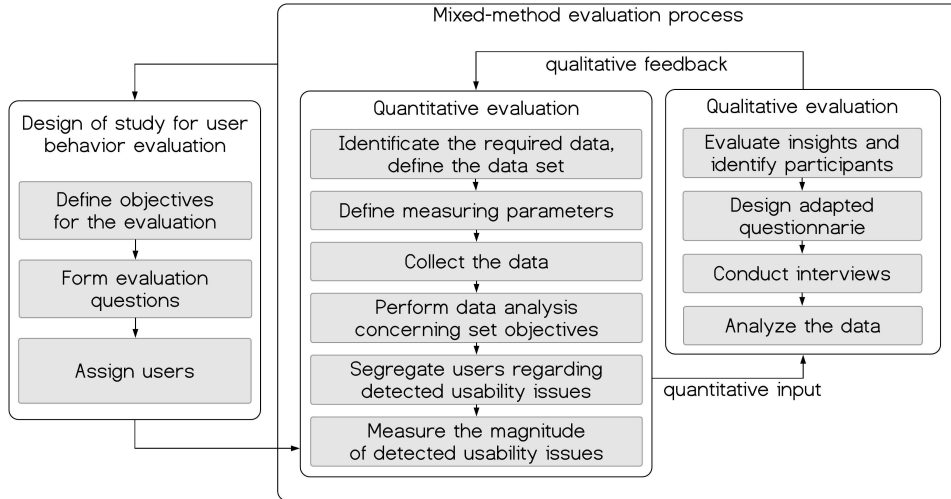
Figure 12. Design for mixed-method user behavior evaluation of ADAS.

According to the proposed design, the design of the study for user behavior evaluation requires initial setup. The main objectives and the focus of the study need to be set. This helps the usability team to design evaluation questions. The study can be focused on identification of any trends in user or system behavior, or can look for any correlations in user behavior to reveal possible usability problems. Further, when the focus of the study is set, relevant drivers for the study need to be assigned. Drivers' background, previous experience regarding the evaluated system, gender, age, work responsibilities, and other parameters can bias the results if drivers are not chosen correctly.

During the quantitative evaluation phase, vehicle data collection and analysis is performed. To design the dataset, every evaluation question needs to be linked to the dataset that supports the answer to the particular question. In addition, the measured parameters, such as the frequency of data collection, the length of the study, the number of participants, number of context parameters, etc. need to be defined. Data collection designed as part of the ND study was proposed as the most natural way of unobtrusive driver-system evaluation in a real environment. During the ND study, performance data for both the driver and system needs to be measured together with contextual information affecting the system performance positively or negatively.

The subsequent application of the qualitative survey was built on the results of the quantitative study aiming to explain the identified phenomena. Semi-structured in-depth interviews with drivers from the quantitative phase were chosen as an appropriate method, aiming to explain and uncover detected issues. The qualitative study design, therefore, focused on the clarification of the subjective reasoning of the drivers inside the detected target groups to understand the specific user behavior.

In a final step, the authors propose to feedback the qualitative findings to the quantitative level for further verification. To achieve a complete understanding, it is helpful to examine if a particular user explanation, received during the qualitative study, applies every time in the same context, and how other drivers behave under the same conditions. This type of analysis helps to understand if the qualitative explanations can be generalized. The mixed-method feedback-loop can also help to identify other relevant data that can be useful in the next round of the quantitative assessment. For example, if a specific interrelation between user and system was detected during the qualitative data analysis, the evaluating team can examine the possibility to include additional vehicle data into the quantitative evaluation for better support of the identified phenomena. Thus, this approach contributes to the further development of a

quantitative dataset.

Additionally, this paper presents preliminary results of entirely quantitative ADAS assessment, confirming the feasibility of the proposed method design. The data analysis was carried out with a focus on the defined objectives and questions formulated beforehand. The contribution of quantitative evaluation for the ADAS functions (namely ACC and PA) usage was measured. The qualitative study helped to (i) measure the usage level for ADAS functions; (ii) differentiate patterns/trends in user behavior by clustering drivers who behave similarly under the same conditions; (iii) evaluate and consider the system performance in the user study; (iv) understand that System Reliability varies for different users performing under the same driving conditions; (v) detect specific usability issues and measure their magnitude; (vi) set a number of hypotheses regarding driver behavior based solely on the vehicle data analysis.

Thus, the inclusion of quantitative evaluation into an existing methodology contributes to more efficient and effective product development. Moreover, the authors believe that this sequential use of quantitative and qualitative approaches and the feedback of the results into the process can support designers and engineers within research and development to create synergies in the development process.

The conference paper outlining these results had the title *"Mixed methods design for the usability studies as a reflection on changing reality in the automotive industry context,"* and was presented as a podium presentation at *ICED conference 2019 in Delft, Netherland*s, published in *Proceedings of the Design Society: International Conference on Engineering Design*.

4.4 PAPER D

4.4.1 Purpose

A full-scale study for ADAS evaluation was conducted with respect to the driving context as one of the major factors influencing driver behavior. The effect of stand-alone contextual variables on driver behavior of ADAS has been assessed by other researchers. However, a complete investigation of this topic is lacking. Therefore, this paper aims to investigate and understand how the driving context affects the use of ADAS. An additional goal of this paper was to apply in practice, and validate, the mixed-method design proposed in Paper C.

4.4.2 Method

The explanatory sequential mixed-method approach proposed in Paper C was adopted and modified for the needs of the current research. The sequential use of quantitative and qualitative methods (see Figure 13) aims to facilitate an integrated interpretation regarding the effect of the driving context on ADAS usage.
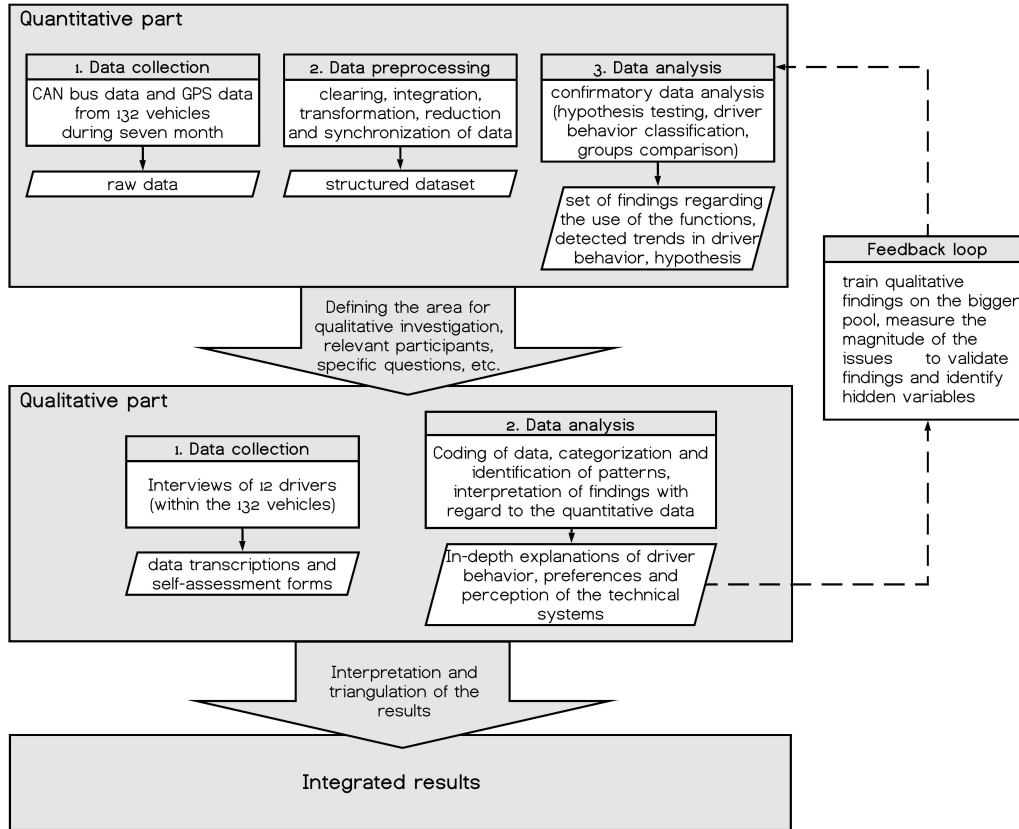
Figure 13. Explanatory sequential mixed-method design.

The following approach proposes two distinct phases: a quantitative and qualitative evaluation. In the course of quantitative study, ND data for both the driver and system performance was collected over a period of seven months. Data variables that enabled the understanding of the driving context for ADAS were also included in the assessment. In total, data from 132 vehicles was collected. Consequently, in the data pre-processing step, different methods were used to clear up, integrate and transform the raw data into the structured data set. All corrupt and inaccurate records were removed from the dataset. The data was synchronized in time, providing order and structure for the initial dataset. Finally, statistical data analysis of collected data was made with the help of software for statistical analysis (Power BI Microsoft, 2019). The data was analyzed in four different layers of abstraction: single DC evaluation layer (if anything indicated unusual or interesting user behavior that needed in-depth investigation), one-driver evaluation layer (focused on in-depth user behavior evaluation of the same driver), groups comparison layer (based on the comparison of user behavior between different user groups), and overall assessment layer (based on the average calculation for all users).

Subsequently the qualitative phase was designed with consideration of the quantitative study results and explained emerging phenomena. In the course of the qualitative study, semi-structured interviews were held with 12 respondents participating in the quantitative study. The aim of the interviews was to explain and uncover the human perception of the driving context and its effect on system usage. The interview data was subsequently transcribed and analyzed by two independent coders using NVivo Software.

The purpose of the triangulation design was to revise the completeness of the quantitative dataset by identifying relevant data-variables from the qualitative study and to verify these in the complete vehicle pool. Moreover, the feedback loop for the qualitative findings was utilized for further investigations at the quantitative level. The qualitative insights were tested on a wide

range of users, aiming to cross validate the hypothesis based on quantitative evaluation and the qualitative explanations.

4.4.3 Main results

This study revealed the effect of the driving context on ADAS performance and driver behavior. The authors advocate the ADAS evaluation approach where the driving context will be considered as one of the major factors for the evaluation of support systems, since the interrelation between driver, ADAS, and the driving context is very high.

The quantitative data analysis in this study enabled the assessment of driver and system performance, as well as the driving context variables indicating the weather, road and traffic conditions. Based on quantitative data analysis, the authors measured the average of ADAS usage for the complete vehicle fleet, as well as the individual grade of ADAS usage. This knowledge helped the authors in drivers categorization based on different use levels of ADAS functions. Further analysis revealed that the driving context, especially the road and traffic conditions, can have a significant effect on the use scenarios that two groups chose for ADAS usage. Therefore, the authors compared the groups' behavior and investigated what the differences were in the way the groups handled the different driving conditions.

The consequent qualitative study confirmed the quantitatively detected differences in drivers behaviors and contributed to the holistic interpretation of the results. The interview data analysis revealed that the driving context had a dual effect on driver behavior: (i) a direct effect, because the driver has to consider the driving situation every time he/she wants to activate the ADAS function; (ii) an indirect effect - the system performance, which also depends on the driving context due to defined limitations, has a different impact on driver perception, affecting the usage of ADAS.

4.5 SUMMARY OF THE RESULTS

The results of the appended papers can be summarized as follows:

- The author investigated the ability to gain relevant vehicle data for user-related studies. As a result, the requirements for the data were specified, and the initial dataset was established.
- The author investigated the methods for usability evaluation of real-driving environments. The primary purpose of this research was to identify the applicability of vehicle data for usability attributes assessment. The borderlines where vehicle data can contribute to driver behavior evaluation were identified.
- The effectiveness of qualitative and quantitative methods for usability was investigated. This led to the conclusion that in the areas of system performance, driver performance, and the driving context assessments, the use of vehicle data is more efficient. Vehicle data is considered more reliable since it records all interactions with the system and driving environment, compared to the user, who tends to "forget" or generalize certain issues.
- The limitations regarding vehicle data acquisition and utilization were identified. This included the technical limitations that affect the data feasibility and applicability negatively (e.g., the driver identification in the vehicle) and the barriers associated with sensitive data processing, preventing us from an extensive evaluation of real users in the natural driving environment.
- The author also examined the methods used for a driver behavior assessment. The lack of integration of qualitative and quantitative approaches leads to the practice, where studies are conducted in isolation making the synthesis of the results difficult.

At the same time, the combination of both quantitative and qualitative methods can help to benefit from different data sources.

- A mixed-method approach for user behavior evaluation of ADAS was proposed. The method helps to integrate quantitative assessment into the existing qualitative methods, resulting in more precise and comprehensive results.
- Practical implementation of the method confirmed the effectiveness of this approach, resulting in a more comprehensive context-aware ADAS evaluation, where both driver and system performance are considered. The quantitative and qualitative approaches in this method complement and validate the results of each other. Thus, this method, according to the author's understanding, can result in better trust in the results from industry professionals.

5

DISCUSSION

This chapter is dedicated to the discussion of the results in connection to the research questions. Additionally, this chapter aims to discuss the quality of the results in relation to the research approach.

5.1 ANSWERING THE RESEARCH QUESTIONS

RQ 1) How can vehicle data be used for data-driven user behavior evaluation?

According to SAE International (2018), there are three primary actors in driving with ADAS: the driver, the ADAS, and "the vehicle." "The vehicle" includes vehicle systems and components influencing ADAS performance, since the ADAS is communicating with a number of vehicle systems that support its performance. Moreover, all interactions between the driver and the system are happening in the dynamic driving context, which is equally important for ADAS evaluation. The driving context is the summary of external factors that affect driver behavior while using the evaluated system (Zhai et al., 2018), and this highly depends on the evaluated objectives. For the ADAS evaluation, the driving context is defined as the aggregation of traffic, road, and weather conditions that, in association, can encourage or discourage the usage of ADAS.

Since ADAS do not perform equally well in all driving contexts, to be able to evaluate driver behavior, all context parameters affecting ADAS performance need to be considered (see the ADAS interrelating factors in Figure 14). According to this, ADAS performance depends on the additional in-vehicle systems' performance, contributing to ADAS performance and driver behavior. The driver behavior, in turn, depends on the cumulative abilities of technical solutions supporting the ADAS performance in the current driving situation. Furthermore, the driving context affects all ADAS interrelations, including the driver, the system, and the performance of additional systems contributing to the ADAS functionality.

As it is revealed (see Figure 14), there is a strong interrelation between ADAS factors. This indicates the strong need for the approach that is able to assess the complexity of these interrelations.
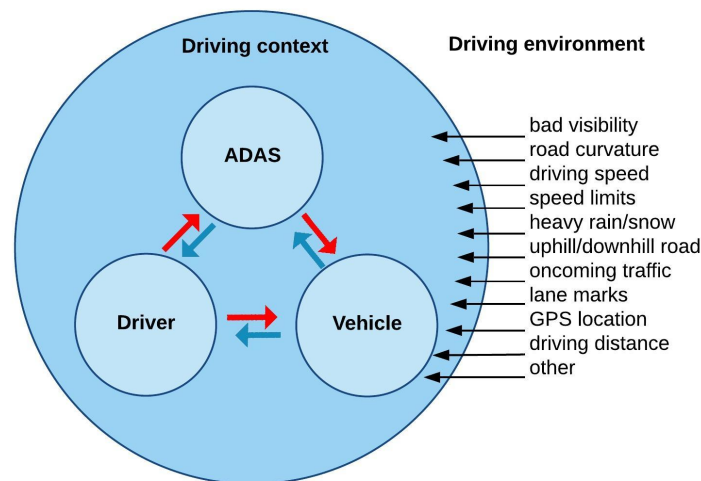
Figure 14. ADAS interrelating factors

The vehicle data, as it is demonstrated in this research, can support the complex assessment of driver-system interactions, considering the spectrum of contextual factors affecting these interactions. In particular, vehicle data can support the measurement of both the driver and system(s) performance, as well as contextual information such as the weather conditions, the road conditions, and the data indicating the traffic conditions on the roads. The summary of the measured variables for ADAS evaluation is presented in Table 1.

Table 1. Summary of variables used for the ADAS evaluation

Driving context variables	Description
Wiping status	to detect heavy rain or snow
Fog illumination	to control visibility on the road
Ambient temperature	to exclude slippery road conditions
Lane marks reading	a precondition for ADAS performance
Speed limits	to identify the road type
Driving speed	to see the deviation from speed limits
Driving distance	to determine the distance between changes
Braking/Acceleration	to identify condensed traffic
Vehicle systems variables	Description
ACC performance	on/off/standby mode - contributes to PA performance
PA performance	on/off/standby mode
Radar On/Off	the signal from the radar ensure the ADAS performance
Camera On/Off	the signal from the cameras ensure the ADAS performance
Driver-related variables	Description
Number of DCs	per day/week/month to understand the level of activity
Frequency of ADAS usage	number of activations/deactivations
Time of activ./deactivations	to calculate the time ADAS is in use
DC length	long/medium/short DC - affects the use of ADAS
DC type	commute/within city/between cities/to other countries
GPS location	to map driver behavior to the driving context in the zoom-in analysis
Error rate	measuring the mismatch between user request and system activation

Vehicle data offers the possibility to determine individual user behavior, describe, categorize, and compare this to the average within a group. Furthermore, it allows identification of specific use errors or a change in driver's ADAS use strategy. Vehicle data analysis enables the understanding of the severity of detected issues by checking the number of vehicles or the

amount of DCs that accounted for the same problem. All of the above mentioned provide the ability for the effective application of quantitative research methods focusing on detection and investigation of driver behavior patterns.

Moreover, vehicle data acquired from the ND study is the only way of unobtrusively logging the interrelations between the system and the human in a real driving environment. In general, the ND data analysis allows precise and reliable results to be obtained, since the outcomes are based on the use of statistical methods and can always be assessed with regard to their statistical significance.

However, it is important to acknowledge that the vehicle data is not perfect and need to be further developed. A number of limitations that can prevent efficient use of vehicle data have been identified. The summary of the main limitations identified in this research is presented in Table 2.

Table 2. Summary of data restrictions identified in this research.

Missing data variables	Comments
Driver Identification	The lack of driver identification unit restricts the ability of recording the "clear" driver behavior, separating the data of drivers that occasionally share the vehicle with another driver
Data from the driver profile	The driver profile can provide better insight into the driver type and can be used for driver categorization. Today driver profile uses optionally.
Secondary tasks' performance	Other activities, except driving (e.g., use of the phone, use of the infotainment or navigation system, etc.)
Driver's interactions with the interfaces	Today only the interactions that lead to the vehicle status change can be measured.

Moreover, since the availability of data coming from the in-vehicle sensors' network is continually increasing, the dataset representing the contextual variables can be further improved. For example, the effect of the oncoming traffic or the in-vehicle context (the use of a mobile phone, distraction from passengers, etc.) were not assessed yet, due to the technical feasibility of establishing these signals. Thus, vehicle sensors and means of driving context assessment need to be continuously improved to be able to provide a more detailed understanding of driver-system interaction.

RQ1 was considered to be a primary research question since a majority of limitations and technical constraints were met at the beginning of this research. **Papers A-D** are focused on building up the data set required for user-related studies and documenting the limitations.

RQ 2) How can the data-driven approach be incorporated into existing methods for driver behavior evaluation?

Today, there is no systematic approach regarding the utilization of vehicle data. Vehicle data is extensively used for system performance verification, but is less used in driver behavior evaluation and driving context assessments. Therefore, the author believes that the ND data can be more effectively used when the number of vehicles with built-in telematics systems increases. The inclusions of on-board diagnostics into the complex assessment of interrelations between the driver, the system and the driving context seems to be the most promising approach for complex systems like ADAS.

So far, there is no single method that is able to support the whole complexity of ADAS interrelations. The traditional way of driver behavior evaluation solely based on subjective user data is not able to consider all interrelating factors between the driver and the system in a dynamic driving context. **Papers A** and **B** concluded that driver behavior evaluation based

solely on vehicle data is not possible. As a reflection on these findings, **Paper C** proposed the explanatory sequential mixed-method design, aiming to effectively utilize both quantitative and qualitative data types for the comprehensive assessment of complex systems akin to ADAS. **Papers D** focused on the implementation of the mixed-method design and validation of the design method through practical study.

Therefore, the explanatory sequential mixed-method design proposed by the author aims to improve the quality of ADAS assessment by combining quantitative and qualitative methods in the most effective way, where it is possible to overcome the limitations of quantitative data by the perks of qualitative data and vice versa. However, a simple consolidation of the results does not always lead to the achievement of a comprehensive understanding of the investigated phenomena. Therefore, the sequential use of both methods allows us to build the in-depth qualitative investigation, using the insights of quantitative study. In this approach, the qualitative study design is a way to explain quantitatively detected issues. Thus, the focus of the qualitative investigation, the choice of participants, and the design of the questionnaire should be made after the results from the quantitative study are obtained. Such an approach contributes to high compatibility of the results between studies and allows the optimization of the data flow and resources for data utilization. Moreover, the sequential mixed-method approach helps to cross-validate the results of both studies, and helps to evaluate the completeness of the datasets of both studies, by reflecting over the missing knowledge in the overall assessment.

RQ 3) How can the validity of the data-based results be achieved?

The validity of the insights based on vehicle data needs to be proven. Developers need to know if they can trust the data and if the conclusions based solely on vehicle data analysis are valid. The sequential mixed-method approach helps to cross-validate the results of both quantitative and qualitative studies. For this purpose, quantitative and qualitative data for the same participants were matched and compared. A high correlation in findings was found between the results of both studies.

However, this is only the first step in the validation chain. The results achieved in **Papers D** shows that the mixed-method loop helps to reflect over the completeness of the dataset for ADAS evaluation, identifying other relevant data that can be included to enhance the quality of quantitative results further. For example, if some specific factor that affects user behavior was detected during the qualitative study but was not measured at the quantitative level, the dataset, and the possibility to include new sensor data that will enable tracking of this specific context, has to be revised. This type of analysis needs to be repeated every time after the evaluation is completed. It will help to increase the precision of data analysis gradually and therefore achieve better validity of the results based on quantitative data analysis.

The author believes that today, when we are still learning how vehicle data can be handled, cross-validating of the results with other methods is one of the ways to ensure the validity of the results. Another way is to ensure the quality of the collected data. What data do we collect? What is the optimal number of signals that should describe the event? How do we pre-process data? What criteria do we apply for data-driven analysis? The author plans a subsequent study where these questions will be addressed, contributing to a better quality of vehicle data and higher validity of data-driven results.

RQ 4) How can vehicle data be used for the data-driven design of vehicle systems?

This research question was not addressed at the current stage of the research. However, it is discussed in the next chapter under the Future work section.

40

5.2 CLARIFICATION OF RESULTS AND SUCCESS CRITERIA

Many factors influence research success. According to the DRM, there are no established metrics to measure success. It is suggested to set the measurable success criteria that are linked to the research goals. The term "measurable" refers to the possibility of evaluation criteria during the research project, i.e., mixed methods can be used in this case (Blessing & Chakrabarti, 2009). In this research, the success criteria related to the research questions were set as follows:

- Possibility to acquire the data needed for data-driven evaluation.
- Ability to handle the data and perform data analysis, achieving reliable results.
- Design a methodology capable of assessing driver behavior for the system under evaluation based on vehicle data.
- Implement the developed methodology in industrial practice.
- Decrease the time and improve the quality of the evaluating processes.

For most of the criteria, it is possible to acknowledge their fulfillment, even with the limitations mainly connected to the vehicle data feasibility. Indeed, the evaluation of the efficiency of vehicle data-driven assessment needs to be further tested in several industrial studies. Since it is a novel approach for the automotive industry, the author expects better availability of vehicle data in the near future. This will increase the possibilities of vehicle data use and can modify the proposed methodology towards higher applicability. The mixed-method approach applied in this research helps to continually evaluate the quality of the vehicle data collection, and reflect over the method and its further modification. The validity of results is confirmed by cross-validation of subjective and objective data. A more detailed discussion of how the verification and validation of the results were achieved is presented in the following section.

5.3 VERIFICATION AND VALIDATION

In order to establish a good quality of research it is important to verify and validate the results. The verification of the results can be ensured by *Logical verification,* which entails the analysis of *coherency*, *completeness* of the results, and *consistency* of internal and external elements. *Validation by acceptance* focuses on the acceptance of new scientific contributions by the scientific community and industry experts within the field.

5.3.1   Logical verification

Coherency is understood as the agreement between established methods and theories. In this research, the author ensured the *coherency* by constructing the method's elements from previously applied research. The achieved results and findings demonstrate *completeness* if they fit into the established theories. The *completeness* of this research is verified by following the steps and guidance of the applied research methodology. *Consistency* is achieved if there are no conflicts in terminology or between different research theories. The current research is based on the combination of established research approaches. The results were always compared to the research publications within and outside the field, which ensured the *external consistency* of this research. Regarding the *internal consistency*, no conflict elements were observed in this research.

5.3.2   Validation by acceptance

A mixed-method approach has been applied in this research. The validity of this type of research needs to be discussed from two different perspectives: quantitative and qualitative.

Validity in quantitative research

Validity in a quantitative study is defined as the extent to which a concept is accurately measured. According to Heale and Twycross (2015), there are three major types of validity: *content validity*, *construct validity*, and *criterion validity*.

Content validity concerns the correctness and accuracy of measurements determined for the assessment of the phenomena. In this research, all datasets designed for the studies were validated by industrial professionals, when the signal descriptions and the correctness of signals' outputs were discussed before the measurements started and after achieving the first prototyping results.

Construct validity is the extent to which a research instrument (data acquisition system in this research) measures the intended construct. The *construct validity* refers to whether one can draw conclusions about test scores related to the concept being studied. The data acquisition system used in the performed studies belongs to the industrial partner. The system is under continuous development, which includes the number of tests and verifications that support this process. However, the data acquisition system used in the studies still has a number of limitations, which often leads to the restriction of required data. This means that sometimes indirect parameters that can only indicate specific issues were considered during the analysis. In these cases, the author was careful in drawing conclusions, and always described the limitations. A study to ensure the *construct validity* by comparing it to solutions developed by other OEMs is planned as a future step.

The final measure of validity is *criterion validity*. *Criterion validity* is the extent to which a research instrument is related to other instruments that measure the same variables. The *criterion validity* in this research was assessed based on literature review within the same field. ND studies and the description of variables they measure, as well as the results they achieve, correlates with the data used in this research and the results achieved.

Validity in a qualitative research

To ensure the validity of research elements in qualitative research, three main aspects need to be discussed: *internal validity*, *external validity* and *construct validity* (Winter, 2000).

Internal validity ensures the validity of the results within the study. This internal validity aspect was considered by designing a number of pre-studies where the prototyping results were delivered and analyzed together with the industrial partners responsible for the quality of data delivery.

External validity concerns the generalizability of the results. This aspect was approached by the deliberate choice of measurement parameters, which are quite broad (e.g., different types of vehicle, the extensive range of users, a variety of vehicle models, different markets) This approach helps to achieve a broader understanding of the ADAS functions and contributes to the generalizability of the results. However, comparison of results across the OEMs was performed based on literature study only. Thus, the author is aware that more studies to validate and compare numerous findings are required. One study aiming to ensure *external validity* is planned as the next step. Similar cases at Daimler, Germany, Volvo cars and Volvo trucks, Sweden will be designed to compare current practices of data-driven driver-behavior evaluation at the different OEMs.

Construct validity establishes correct operational measures for the concept being studied. The subject of analysis in this research was related to both the OEM and the drivers. The use of structured coding techniques correlates with the presented descriptive information associated with the collected data.

Cross-validation

As was previously mentioned (see section 3.3.4), a cross-validation approach was used. The drivers' interview insights were cross-validated with quantitative findings, and the correlation of the results was confirmed (**Paper D**).

Moreover, both quantitative and qualitative insights were cross-validated by additional discussions with the OEM's professionals. The intermediate and final results were presented numerous times at the company. All published papers had been granted permission from the OEM to be published. The papers included in this thesis have undergone the peer-review process, resulting in one conditional acceptance in a journal and three podium presentations at international conferences.

5.4 RESEARCH CONTRIBUTION

5.4.1 Scientific contribution

The scientific goal of this research is to design methodology for effective user behavior evaluation utilizing vehicle data and to understand how these methods can be incorporated into the existing practices of user behavior evaluation. According to these goals, the following steps were carried out:

- The author designed a novel methodology for vehicle data utilization, defining the area where vehicle data can be used, identifying the influencing factors for the evaluated objects, defining the relevant data for the data-driven driver behavior evaluation and investigating ways to improve the feasibility of vehicle data. This is a reasonably new area within the automotive industry research with the ongoing development of methods for logging, processing, analysis, and visualization of interaction data. Therefore, the research in this area is ongoing.
- Furthermore, a proposed mixed-method approach was applied in practice. As long as there is no single method that can help to capture the complexity of user behavior, this research contributes to the new design for user behavior evaluation.

5.4.2 Industrial contribution

The results contribute to the industrial practice by enhancing the quality of the ADAS evaluation. The main industrial goal of this research project was to learn how to utilize vehicle signals in user-related studies and to transfer this knowledge to the engineers dealing with these types of tasks in practice. The results of this research were transferred to the OEM. All possible implications, advantages, and limitations of the data-driven evaluation were discussed. During this research, the technical feasibility of vehicle data was significantly improved. Moreover, the methods for data acquisition, data pre-processing, and data analysis were also improved.

The designed approach was successfully tested at the company and improved the quality of the driver behavior evaluation by the effective combination of different types of data and data analysis. The application of a mixed-method approach, where feedback of the results flows back into the evaluating process, can support synergies between product developers and UX designers. As a result, the findings can echo into more efficient and effective product development, providing an automated way of data collection and driver behavior evaluation that saves company resources and significantly decreases the time for this type of assessment.

6

CONCLUSIONS

This chapter presents the results and the research challenges identified. In addition, future research plans are discussed.

6.1 CONCLUSIONS

The conducted research revealed the great potential of vehicle data utilization for data-driven user behavior evaluation. In the course of this research, a number of activities that contributed to this research was carried out: (i) the feasibility of vehicle data from the Volvo ND study was investigated; (ii) the required data for ADAS driver and system evaluation were specified; (iii) all measuring parameters relevant for the ADAS evaluation (i.e., driving context, measuring period, specific user parameters, etc.) were investigated and defined; (iv) collected data was statistically analyzed on different levels of abstraction, starting from average comparisons between drivers or groups of drivers and becoming more rooted to the one driver evaluation level or even one single driving activity evaluation level.

The vehicle data analysis revealed that the objective assessment of driver and system performance, as well as the driving context variables such as the weather, road and traffic conditions are possible. Vehicle data offers the possibility to determine individual user behavior, and to describe, categorize, and compare this to the average within a group. Furthermore, it allows the identification of specific use errors or a change of driver's use strategy. Vehicle data analysis enables the understanding of the severity of detected issues by checking the number of vehicles or the amount of DCs that accounted for the same problem. All of the above mentioned makes the applicability of vehicle data for user-related studies meaningful.

However, despite the significant potential of ND data for ADAS evaluation and valuable results that can be achieved, there are still some limitations that need to be considered. One of the limitations is that although vehicle data allows context-aware driver and system performance evaluation, the underlying explanations for why objectively detected things happened cannot be determined through the vehicle data alone. Due to the restricted data collection procedures, it is often not possible to measure such human-related aspects as driver perception or driver subjective impression on the interaction with the system. Therefore, in the

course of this research, an explanatory sequential mixed-method was designed and tested at Volvo, as an industrial case for ADAS evaluation. The combination of qualitative and qualitative approaches contributed to more effective ADAS evaluation where the driver behavior and his/her human-related aspects are considered. The practical implementation of the method showed the ability for a comprehensive view of all factors affecting the ADAS usage: the driver, the system (including other subsystems affecting the driver), and the driving context that has a high impact on driver and system behavior.

Another limitation is the feasibility of vehicle data that often restricts the study design to more limited driver/system evaluation. Thus, the means and methods for driver/system behavior evaluation needs to be further improved.

In addition, the author acknowledges the need for conducting more studies where the proposed mixed-method and the results based on vehicle data analysis can be further validated.

6.2 FUTURE WORK

The knowledge about driver behavior has to date been obtained from historical datasets that are gathered from the ND study. As a result, the data can only be used for driver behavior evaluation, but cannot support the user while using ADAS. The processing of such historical data on driver behavior, along with the real-time data of the same driver, can potentially enable real-time analysis of driver behavior and provide the information regarding possible improvements of the ADAS use strategies.

Since vehicle data used in this research provides information about personalized driver behavior, this helps in the understanding of what types of use strategy each driver has, what kind of mistakes he/she makes, if any, and how the driver can be better supported in ADAS usage. The ultimate goal of further research is to utilize vehicle signals to understand the implications, advantages and limitations of the system from a user point of view.

Thus, the author's ambitions are to design and test a data-driven communication framework that can utilize historical and real-time vehicle data to provide real-time support to ADAS users. The data-driven communication framework can better communicate the ADAS capabilities and limitations and suggest effective use of the system in real-time driving situations. The author believes that this type of assistance can improve a driver's understanding of ADAS functionality, encourage its usage or the effectiveness of ADAS use strategies.

However, vehicle data utilization in the design of ADAS personalized support requires additional research in a number of areas. Besides the development of machine-learning algorithms for real-time driver behavior evaluation and driver response measured on the provided support, a number of other factors also need to be considered:

- Real-time data analysis needs to be acquired, not only to understand and categorize the driver behavior regarding the ADAS. There is a need for automated driving event recognition since, as previous research has revealed, the driving context has a high impact on driver behavior and ADAS performance.
- To identify the right moment for driver-system communication, the system awareness about the driver's workload needs to be improved. We need to evaluate the driver distraction caused by primary driving activities and the driving situation on the road. We also need to consider the performance of secondary tasks that the driver can be involved in. Driver-system communication must only take place when a driver's workload is medium-low. The safety of the driver must always be prioritized.
- We need to overcome one of the main limitations of data-driven user behavior evaluation - driver recognition, since the system must know who is behind the wheel to provide personalized support.
- Different communication strategies can be introduced; explaining the ADAS performance, stimulating the use of ADAS function, or warning the driver who uses the

system in critical driving conditions.

- The algorithms that read the driver's behavior and provide the communication need to be adjustable and depend on the driver's reaction to prior communication and consider all changes in driver's use strategies.

REFERENCES

Abrahão, S., Bourdeleau, F., Cheng, B., Kokaly, S., Paige, R., Stöerrle, H. and Whittle, J., 2017, September. User experience for model-driven engineering: Challenges and future directions. In 2017 ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems (MODELS) (pp. 229-236). IEEE.

ADAS&ME - Adaptive ADAS to support incapacitated drivers Mitigate Effectively risks through tailor-made HMI under automation, 2019. http://www.adasandme.com/about-adasme/objectives/

Ahlström, C., Anund, A., Fors, C. and Åkerstedt, T., 2018. Effects of the road environment on the development of driver sleepiness in young male drivers. Accident Analysis & Prevention, 112, pp.127-134.

Ahmed, M.M. and Ghasemzadeh, A., 2018. The impacts of heavy rain on speed and headway behaviors: an investigation using the SHRP2 naturalistic driving study data. Transportation research part C: emerging technologies, 91, pp.371-384.

Albert, W. and Tullis, T., 2013. Measuring the user experience: collecting, analyzing, and presenting usability metrics. Newnes.

Alpaydin, E., 2010. Introduction to Machine Learning, 2nd ed., The MIT Press, London England.

Altshuller, G., Shulyak, L., & Rodman, S., 1999. The innovation algorithm: TRIZ, systematic innovation and technical creativity.

Andreasen, M.M., 1991. Design methodology. *Journal of Engineering Design*, *2*(4), pp.321-335.

Angelini, M., Blasilli, G., Lenti, S. and Santucci, G., 2018. STEIN: Speeding up Evaluation Activities With a Seamless Testing Environment INtegrator. In EuroVis.

Atieno, O.P., 2009. An analysis of the strengths and limitation of qualitative and quantitative research paradigms. Problems of Education in the 21st Century, 13(1), pp.13-38.

Barlas, Y. and Carpenter, S., 1990. Philosophical roots of model validation: two paradigms. *System Dynamics Review*, *6*(2), pp.148-166.

Berman, J.J., 2013. Principles of big data: preparing, sharing, and analyzing complex information. Newnes.

Benmimoun, M., Pütz, A., Zlocki, A. and Eckstein, L., 2013. Eurofot: Field operational test and impact assessment of advanced driver assistance systems: Final results. In Proceedings of the FISITA 2012 World Automotive Congress (pp. 537-547). Springer, Berlin, Heidelberg.

Blessing, L.T. and Chakrabarti, A., 2009. *DRM: A design research methodology* (pp. 13-42). Springer London. http://doi.org/10.1007/978-1-84882-587-1

Bouwer, R. et al., 2014. Effect of genre on the generalizability of writing scores. Language Testing, 32(1), pp.83–100. Available at: http://dx.doi.org/10.1177/0265532214542994.

Braha, D. and Maimon, O., 2013. *A mathematical theory of design: foundations, algorithms and applications* (Vol. 17). Springer Science & Business Media.

Brannen, J., 2005. Mixing methods: The entry of qualitative and quantitative approaches into the research process. International journal of social research methodology, 8(3), pp.173-184.

Bruno, V. and Al-Qaimari, G., 2004. Usability attributes: An initial step toward effective user-centred development. The *Proceedings of Australian Computer Human Interaction Conference*.

Cachia, M. and Millward, L., 2011. The telephone medium and semi-structured interviews: A complementary fit. *Qualitative Research in Organizations and Management: An International Journal*, 6(3), pp.265-277.

Carta, T., Paternò, F. & de Santana, V.F., 2011. Web Usability Probe: A Tool for Supporting Remote Usability Evaluation of Web Sites. Lecture Notes in Computer Science, pp.349–357. Available at: http://dx.doi.org/10.1007/978-3-642-23768-3_29.

Chen, M., Mao, S. and Liu, Y., 2014. Big data: A survey. Mobile networks and applications, 19(2), pp.171-209.

CARTRE - Coordination of Automated Road Transport Deployment for Europe, 2019. https://connectedautomateddriving.eu/about-us/cartre/

Creswell, J.W., 2014. Research design: Qualitative, quantitative, and mixed methods approaches (4th ed.). Sage publications.

Creswell, J.W. and Miller, D.L., 2000. Determining validity in qualitative inquiry. Theory into practice, 39(3), pp.124-130.

Denzin, N.K. and Lincoln, Y.S., 2005. The Sage handbook of qualitative research. Sage Publications Ltd.

Denzin, N.K. and Lincoln, Y.S., 2008. Collecting and interpreting qualitative materials (Vol. 3). Sage.

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G.B. and Victor, T., 2018. Great expectations: a predictive processing account of automobile driving. Theoretical issues in ergonomics science, 19(2), pp.156-194.

Folmer, E. and Bosch, J., 2004. Architecting for usability: a survey. Journal of systems and software, 70(1-2), pp.61-78.

Forlizzi, J. and Battarbee, K., 2004, August. Understanding experience in interactive systems. In Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 261-268). ACM.

Fridman, L., Brown, D.E., Glazer, M., Angell, W., Dodd, S., Jenik, B., Terwilliger, J., Patsekin, A., Kindelsberger, J., Ding, L. and Seaman, S., 2019. MIT advanced vehicle technology study: Large-scale naturalistic driving study of driver behavior and interaction with automation. IEEE Access, 7, pp.102021-102038.

Gantz, J. and Reinsel, D., 2015. Extracting value from chaos. IDC iView (2011).

Gero, J.S. and Kannengiesser, U., 2004. The situated function–behaviour–structure framework. *Design studies*, 25(4), pp.373-391.

Greene, J.C., 2007. Mixed methods in social inquiry (Vol. 9). John Wiley & Sons.

Hatchuel, A. and Weil, B., 2003. A new approach of innovative Design: an introduction to CK theory. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm*.

Hatfield, J., Dozza, M., Patton, D.A., Maharaj, P., Boufous, S. and Eveston, T., 2017. On the use of naturalistic methods to examine safety-relevant behaviours amongst children and evaluate a cycling education program. Accident Analysis & Prevention, 108, pp.91-99.

Hubka, V. and Eder, W.E., 1987. A scientific approach to engineering design. *Design studies*, 8(3), pp.123-137.

ISO/IEC 9241-11:1998: Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability. Available at https://www.iso.org/standard/16883.html

ISO/IEC 9241-210:2010: Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. Available at: https://www.iso.org/standard/52075.html

Itoh, M., 2012. Toward overtrust-free advanced driver assistance systems. Cognition,

technology & work, 14(1), pp.51-60.

Ivory, M.Y. & Hearst, M.A., 2001. The state of the art in automating usability evaluation of user interfaces. ACM Computing Surveys, 33(4), pp.470–516. Available at: http://dx.doi.org/10.1145/503112.503114.

Jenness, J.W., Lerner, N.D., Mazor, S., Osberg, J.S. and Tefft, B.C., 2008. Use of advanced in-vehicle technology by young and older early adopters. Survey Results on Adaptive Cruise Control Systems. Report No. DOT HS, 810, p.917.

Johanson, M., 2017. WICE: Automotive telematics, fleet management, rapid prototyping and software download for test vehicles.[online] Alkit Communications AB.

Johnson, R.B. & Onwuegbuzie, A.J., 2004. Mixed Methods Research: A Research Paradigm Whose Time Has Come. Educational Researcher, 33(7), pp.14–26. Available at: http://dx.doi.org/10.3102/0013189x033007014.

Kazi, T., Stanton, N.A., Walker, G.H. and Young, M.S., 2007. Designer driving: drivers' conceptual models and level of trust in adaptive cruise control.

Kim, J.H., Gunn, D.V., Schuh, E., Phillips, B., Pagulayan, R.J. and Wixon, D., 2008, April. Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 443-452). ACM.

Köhler, M., Falk, B. and Schmitt, R., 2015. Applying Eye-Tracking in Kansei Engineering Method for Design Evaluations in Product Development. International Journal of Affective Engineering, 14(3), pp.241-251.

Liang, Y., McLaurin, E.J., Simmons, L.A., Verma, S.K., Horrey, W.J. and Lesch, M.F., 2016, October. Considering traffic and roadway context in driver behavior assessments: a preliminary analysis. In Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 87-92). ACM.

Llaneras, R.E., 2006. Exploratory study of early adopters, safety-related driving with advanced technologies. Draft final task 2 report: In-vehicle systems inventory, recruitment methods & approaches, and owner interview results (No. HS-809 972).

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H., 2011. Big data: the next frontier for innovation, competition, and productivity. McKinsey Global Institute, McKinsey & Company.

Merriam, S.B., 2009. Qualitative Research: a guide to design and interpretation. San Francisco: Jos-sey-Bass.

Molich, R., Thomsen, A.D., Karyukina, B., Schmidt, L., Ede, M., van Oel, W. and Arcuri, M., 1999, May. Comparative evaluation of usability tests. In Conference on Human Factors in Computing Systems: CHI'99 extended abstracts on Human factors in computing systems (Vol. 15, No. 20, pp. 83-84).

Nanda, S., Rivas, A., Trochim, W. and Deshler, J., 2000. Emphasis on validation in research: A meta-analysis. Scientometrics, 48(1), pp.45-64. http://doi.org/10.1023/A%3A1005628301541

Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J. and Goodman, M., 2005. An overview of the 100-car naturalistic study and findings. National Highway Traffic Safety Administration, Paper, 5, p.0400.

Nielsen, J., 1994. Usability engineering. Elsevier.

NN/g Conference, 2008. User Experience: Nielsen Norman Group Conference, Amsterdam. Available at: https://usabilitygeek.com/the-difference-between-usability-and-user-experience/

NVivo, 2019. Powerful research made easier. Available at: https://www.qsrinternational.com/nvivo/home

Orlovska, J., Wickman, C. and Söderberg, R., 2018. BIG DATA ANALYSIS AS A NEW APPROACH FOR USABILITY ATTRIBUTES EVALUATION OF USER INTERFACES: AN AUTOMOTIVE INDUSTRY CONTEXT. Proceedings of the DESIGN 2018 15th International Design Conference. Available at: http://dx.doi.org/10.21278/idc.2018.0243.

Orlovska, J., Novakazi, F., Wickman, C. & Söderberg, R., 2019. Mixed-Method Design for User Behavior Evaluation of Automated Driver Assistance Systems: An Automotive Industry Case. Submitted to 22nd International Conference on Engineering Design, 2019.

Ovaska, S., 1991. Usability as a Goal for the Design of Computer Systems. Scandinavian J. Inf. Systems, 3(1), p.4.

Papazikou, E., Quddus, M.A. and Thomas, P., 2017. Detecting deviation from normal driving using SHRP2 NDS data. Presented at the Transportation Research Board (TRB) 96th Annual Meeting, Washington D.C., US, 8th-12th January 2017.

Peham, M., Breitfuss, G. and Michalczuk, R., 2014, October. The ecogator app: gamification for enhanced energy efficiency in europe. In Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality (pp. 179-183). ACM.

Power BI Microsoft, 2019. Business intelligence like never before. Available at: https://powerbi.microsoft.com/en-us/

Rahman, M.S., 2016. The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language "Testing and Assessment" Research: A Literature Review. Journal of Education and Learning, 6(1), p.102. Available at: http://dx.doi.org/10.5539/jel.v6n1p102.

Rosenbaum, S., 2002. Usability in Practice: Field Studies. Field Studies–Evolution and Revolution CHI.

Roto, V., Law, E., Vermeeren, A.P.O.S. and Hoonhout, J., 2011, September. User experience white paper: Bringing clarity to the concept of user experience. In Dagstuhl Seminar on Demarcating User Experience (p. 12).

Rovai, A.P., Baker, J.D. and Ponton, M.K., 2013. Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS. Watertree Press LLC.

SAE International (2018): J3016_201806 - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles.

SCOUT - Safe and COnnected AUtomation in Road Transport, 2019. https://connectedautomateddriving.eu/about-us/scout/

Sallee, M.W. & Flood, J.T., 2012. Using Qualitative Research to Bridge Research, Policy, and Practice. Theory Into Practice, 51(2), pp.137–144. Available at: http://dx.doi.org/10.1080/00405841.2012.662873.

Sander, U., 2017. Opportunities and limitations for intersection collision intervention—A study of real world 'left turn across path' accidents. Accident Analysis & Prevention, 99, pp.342-355. http://dx.doi.org/10.1016/j.aap.2016.12.011

SHRP2 - Strategic Highway Research Program 2, 2019. https://insight.shrp2nds.us/

Suh, N.P., 2001. *Axiomatic design: advances and applications*. Mit-Pappalardo Series in Mecha.

THE SAFER organization, 2019. https://www.saferresearch.com/about#block-views-block-research-area-directors-block-1

Tivesten, E. and Dozza, M., 2014. Driving context and visual-manual phone tasks influence glance behavior in naturalistic driving. Transportation research part F: traffic

psychology and behaviour, 26, pp.258-272.

Tornell, S. et al., 2015. Simplifying the in-vehicle connectivity for ITS applications. Proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. Available at: http://dx.doi.org/10.4108/eai.22-7-2015.2260058.

Visuri, A., van Berkel, N., Luo, C., Goncalves, J., Ferreira, D. and Kostakos, V., 2017, July. Challenges of quantified-self: encouraging self-reported data logging during recurrent smartphone usage. In Proceedings of the 31st British Computer Society Human Computer Interaction Conference(p. 81). BCS Learning & Development Ltd.

Volvo Cars, 2019. Tips for using Pilot Assist. Available at: https://www.volvocars.com/uk/support/topics/use-your-car/car-functions/tips-for-using-pilot-assist.

Vuillemot, R., Boy, J., Tabard, A., Perin, C. and Fekete, J.D., 2016. Challenges in Logging Interactive Visualizations and Visualizing Interaction Logs. LIVVIL: Logging Interactive Visualizations & Visualizing Interaction Logs.

Wallace, K.M. and Blessing, L.T., 2000. Observations on some German contributions to engineering design in memory of Professor Wolfgang Beitz. *Research in Engineering Design*, *12*(1), pp.2-7.

Yin, R.K., 2013. Case Study Research Design and Methods, 5th Revise.

Zhai, Y., Wo, T., Lin, X., Huang, Z. and Chen, J., 2018, June. A Context-Aware Evaluation Method of Driving Behavior. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 462-474). Springer, Cham.

PAPER A

BIG DATA USAGE CAN BE A SOLUTION FOR USER BEHAVIOR EVALUATION: AN AUTOMOTIVE INDUSTRY EXAMPLE

Julia Orlovska, Casper Wickman, Rikard Söderberg

PAPER B

BIG DATA ANALYSIS AS A NEW APPROACH FOR USABILITY ATTRIBUTES
EVALUATION OF USER INTERFACES: AN AUTOMOTIVE INDUSTRY CONTEXT

Julia Orlovska, Casper Wickman, Rikard Söderberg

PAPER C

MIXED-METHOD DESIGN FOR USER BEHAVIOR EVALUATION OF AUTOMATED DRIVER ASSISTANCE SYSTEMS: AN AUTOMOTIVE INDUSTRY CASE

Julia Orlovska, Fjollë Novakazi, Casper Wickman, Rikard Söderberg

PAPER D

EFFECTS OF THE DRIVING CONTEXT ON THE USAGE OF AUTOMATED DRIVER
ASSISTANCE SYSTEMS (ADAS)
NATURALISTIC DRIVING STUDY FOR ADAS EVALUATION

Julia Orlovska, Fjollë Novakazi, Lars-Ola Bligård, MariAnne Karlsson, Casper Wickman,
Rikard Söderberg