Methodological Framework for Modelling and Empirical Approaches (Deliverable D1.1 in the H2020 MSCA ITN project SHAPE-IT)

Figalová, Nikol; Mbelekani, Naomi Yvonne; Zhang, Chi; Yang, Yue; Peng, Chen; Nasser, Mohamed; Yuan-Cheng, Liu ; Muhammad, Amna Pir ; Tabone, W.; Berge, S.H.

DOI
10.17196/shape-it/2021/02/D1.1

Publication date
2021

Document Version
Final published version

Citation (APA)
Figalová, N., Mbelekani, N. Y., Zhang, C., Yang, Y., Peng, C., Nasser, M., Yuan-Cheng, L., Muhammad, A. P., Tabone, W., Berge, S. H., Jokhio, S., He, X., Kalantari, A. H., Mohammadi, A., Yang, X., Bärgman, J., & Baumann, M. (2021). *Methodological Framework for Modelling and Empirical Approaches (Deliverable D1.1 in the H2020 MSCA ITN project SHAPE-IT)*. SHAPE-IT Consortium . https://doi.org/10.17196/shape-it/2021/02/D1.1

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Deliverable 1.1

Methodological Framework for Modelling and Empirical Approaches

Deliverable Status	**FINAL**
Type	**REPORT**
Dissemination level according to the Grant Agreement	**PUBLIC**

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Project Information

Project Name:	SHAPE-IT – Supporting the interaction of Humans and Automated vehicles: Preparing for the Environment of Tomorrow
Grant Agreement:	860410
Project Duration:	1 October 2019 – 30 September 2023
Coordinator	Chalmers University of Technology

Document Information

Due Date	31Mar 2021	**Due date in months**	M18
Date of submission	31Mar 2021	**Month of submission**	M18

Status	**Author**	**Date**
Prepared by	Nikol Figalova, Naomi Yvonne Mbelekani, Chi Zhang, Yue Yang, Chen Peng, Mohamed Nasser, Liu Yuan-Cheng, Amna Pir Muhammed, Wilbert Tabone, Siri Hegna Berge, Sarang Jokhio, Xiaolin He, Amir Hossein Kalantari, Ali Mohammadi, Xiaomi Yang	22 February 2021
Reviewed	Martin Baumann, Olga Pollatos	25 February 2021
Approved	Jonas Bärgman, Chalmers	25 March 2021

Cite this as:
Figalova, N., Jokhio, S., Nasser, M., Mbelekani, N. Y., Zhang, C., Yang, Y., Peng, C., Liu, Y. C., Muhammad, A. P., Tabone, W., Berge, S. H., He, X., Kalantari, A. H., Mohammadi, A., Yang, X., Pollatos, O., Bärgman, J. & Baumann, M. (2021). *Methodological Framework for Modelling and Empirical Approaches* (Deliverable D1.1 in the EC ITN project SHAPE-IT). SHAPE-IT Consortium. DOI: 10.17196/shape-it/2021/02/D1.1

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table of Contents

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

List of Tables

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

List of Figures

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

List of Abbreviations

ADAS	Advanced Driver Assistance Systems
AI	Artificial Intelligence
ANS	Autonomous Nervous System
AR	Augmented Reality
AV	Automated Vehicle
BGT	Behavioural Game Theory
CAN	Controller Area Network
CAVE	Cave Automatic Virtual Environment
CIT	Crossing Initiating Time
DAS	Data Acquisition System
DOF	Degree of Freedom
DL	Deep Learning
DST	Deceleration to Safety Time
EAM	Evidence Accumulation Model
ECG	Electrocardiography
EDA	Electrodermal Activity
EEG	Electroencephalography
EMG	Electromyography
ERP	Event-Related Potential
eHMI	External Human-Machine Interface
ESR	Early Stage Researcher
FOV	Field of View
FOR	Field of Regard
GT	Game Theory
HMD	Head-Mounted Display
HMI	Human-Machine Interface
HR	Heart Rate

HRV	Heart Rate Variability
MC	Method Champion
NDD	Naturalistic Driving Data
NDS	Naturalistic Driving Study
NDRT	Non-Driving Related Task
ODDs	Operational Design Domains)
RE	Requirements Engineering
SA	Situation Awareness
SCL	Skin Conductance Level
SCR	Skin Conductance Response
SMC	Supporting Method Champion
TOR	Take-Over Request
TOT	Take-Over Time
TTA	Time to Arrival
TTC	Time to Collision
VR	Virtual Reality
VRU	Vulnerable Road User
WoOZ	Wizard of Oz

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Executive Summary

The progress in technology development over the past decades, both with respect to software and hardware, offers the vision of automated vehicles as means of achieving zero fatalities in traffic. However, the promises of this new technology – an increase in road safety, traffic efficiency, and user comfort – can only be realized if this technology is smoothly introduced into the existing traffic system with all its complexities, constraints, and requirements. SHAPE-IT will contribute to this major undertaking by addressing research questions relevant for the development and introduction of automated vehicles in urban traffic scenarios. Previous research has pointed out several research areas that need more attention for a successful implementation and deployment of human-centred vehicle automation in urban environments.

In SHAPE-IT, for example, a better understanding of human behaviour and the underlying psychological mechanisms will lead to improved models of human behaviour that can help to predict the effects of automated systems on human behaviour already during system development. Such models can also be integrated into the algorithms of automated vehicles, enabling them to better understand the human interaction partners' behaviours.

Further, the development of vehicle automation is much about technology (software and hardware), but the users will be humans and they will interact with humans both inside and outside of the vehicle. To be successful in the development of automated vehicles functionalities, research must be performed on a variety of aspects. Actually, a highly interdisciplinary team of researchers, bringing together expertise and background from various scientific fields related to traffic safety, human factors, human-machine interaction design and evaluation, automation, computational modelling, and artificial intelligence, is likely needed to consider the human-technology aspects of vehicle automation.

Accordingly, SHAPE-IT has recruited fifteen PhD candidates (Early Stage Researchers – ESRs), that work together to facilitate this integration of automated vehicles into complex urban traffic by performing research to support the development of transparent, cooperative, accepted, trustworthy, and safe automated vehicles. With their (and their supervisors') different scientific background, the candidates bring different theoretical concepts and methodological approaches to the project. This interdisciplinarity of the project team offers the unique possibility for each PhD candidate to address research questions from a broad perspective – including theories and methodological approaches of other interrelated disciplines. This is the main reason why SHAPE-IT has been funded by the European Commission's Marie Skłodowska-Curie Innovative Training Network (ITN) program that is aimed to train early state researchers in multidisciplinary aspects of research including transferable skills. With the unique scope of SHAPE-IT, including the human-vehicle perspective, considering different road-users (inside and outside of the vehicle), addressing for example trust, transparency, and safety, and including a wide range of methodological approaches, the project members can substantially contribute to the development and deployment of safe and appreciated vehicle automation in the cities of the future.

To achieve the goal of interdisciplinary research, it is necessary to provide the individual PhD candidate with a starting point, especially on the different and diverse methodological approaches of the different disciplines. The empirical, user-centred approach for the development and evaluation of innovative automated vehicle concepts is central to SHAPE-IT. This deliverable (D1.1 "Methodological Framework for Modelling and Empirical Approaches") provides this starting point. That is, this document provides a broad overview of approaches and methodologies used and developed by the SHAPE-IT ESRs during their research. The SHAPE-IT PhD candidates, as well as other researchers and developers outside of SHAPE-IT, can use this document when searching for appropriate methodological approaches, or simply get a brief overview of research methodologies often employed in automated vehicle research.

The first chapter of the deliverable shortly describes the major methodological approaches to collect data relevant for investigating road user behaviour. Each subchapter describes one approach, ranging from naturalistic driving studies to controlled experiments in driving simulators, with the goal to provide the unfamiliar reader with a broad overview of the approach, including its scope, the type of data collected, and its limitations. Each subchapter ends with recommendations for further reading – literature that provide much more detail and examples.

The second chapter explains four different highly relevant tools for data collection, such as interviews, questionnaires, physiological measures, and as other current tools (the Wizard of Oz paradigm and Augmented and Virtual Reality). As in the first chapter this chapter provides the reader with information about advantages and disadvantages of the different tools and with proposed further readings.

The third chapter deals with computational models of human/agent interaction and presents in four subchapters different modelling approaches, ranging from models based on psychological mechanisms, rule-based and artificial intelligence models to simulation models of traffic interaction.

The fourth chapter is devoted to Requirements Engineering and the challenge of communicating knowledge (e.g., human factors) to developers of automated vehicles. When forming the SHAPE-IT proposal it was identified that there is a lack of communication of human factors knowledge about the highly technical development of automated vehicles. This is why it is highly important that the SHAPE-IT ESRs get training in requirement engineering. Regardless of the ESRs working in academia or industry after their studies it is important to learn how to communicate and disseminate the findings to engineers.

The deliverable ends with the chapter "Method Champions". Here the expertise and association of the different PhD candidates with the different topics are made explicit to facilitate and encourage networking between PhDs with special expertise and those seeking support, especially with regards to methodological questions.

Introduction

According to World Health Organization (2015), 1.25 million people die every year worldwide in traffic collisions, including over 25 000 in the European Union. Research shows that over 90% of traffic collisions are caused by human error. Car automation is considered to be a game-changer in the field of transportation. It promises not only safer, but also faster and more comfortable way of traveling. However, there is still a long journey ahead of us to reach full automation, i.e., when no human intervention is required under any circumstances. Furthermore, it is certainly not possible to have fully automated vehicles (AVs) replacing human driving vehicles overnight. There will be a long period of transition where AVs will coexist with human-driven vehicles. In case of urban roads and streets, the AVs will always coexist with vulnerable road users (VRUs) such as pedestrians or cyclists.

In SHAPE-IT, we aim to facilitate safe, acceptable, and desirable integration of user-centred and transparent AVs into the mixed urban traffic environments. To achieve this goal, an interdisciplinary approach is necessary. Fifteen PhD candidates are currently working on separate but related topics to promote safer traffic environment. We employ theories and methods from fields such as engineering, psychology, human factors, ergonomics, neuroscience, and computer science for the purpose of bridging links between humans and AVs.

However, along with many advantages, there are also challenges related to the interdisciplinary approach. Each discipline stands on a robust body of theoretical background and various methodological approaches, and each Early Stage Researcher (ESR) working within SHAPE-IT has a different background and knowledgebase. Therefore, in this document we try to outline an overview of different methodological approaches, to highlight their strengths and weaknesses, and to help the reader to get better oriented in the various research methodology employed in SHAPE-IT. Furthermore, we provide a broad knowledge about different empirical and modelling approaches, which guide researchers in knowing the suitable study design for their research question. Conclusively, for gaining an in-depth knowledge of the specific methods, we provide suggest readings at the end of each section.

The structure of this document is divided into four main chapters. Chapter one focuses on the empirical approaches employed in SHAPE-IT. It follows the structure of decreasing ecological validity and increasing scientific control. With each step forward in this chapter, the experimenter loses certain amount of ecological validity, but also reveals new research opportunities. First, we introduce the naturalistic driving studies, followed by on road studies, test track studies, and simulator studies. We briefly introduce driving simulators, pedestrian simulators, and cycling simulators. The chapter is closed by an overview of types of behavioural data we intend to collect in SHAPE-IT. Chapter two introduces several techniques and tools available to researchers in SHAPE-IT, which can be applied across the different empirical approaches mentioned in chapter one. Interviews, questionnaires, and psychophysiological techniques are introduced briefly. These allow researchers to control for covariates, collect qualitative data, understand better their experimental results, and also to broaden the pool of research questions that can be asked. The end of chapter two belongs to an introduction of the Wizard of Oz method and the virtual reality (VR) and augmented reality

(AR), which represent further tools that broaden the research possibilities in SHAPE-IT. Chapter three introduces the various modelling approaches that will be used in SHAPE-IT to understand behaviour of different road users (automated vehicles and vulnerable road users). Chapter four is an overview of the Requirements Engineering and its possible application in the AV research. Finally, Chapter five introduces the concept of Method Champions, and brings an overview of the ESRs and their domain of expertise.

This document exists thanks to a joint effort of the 15 ESRs involved in SHAPE-IT. As each of the ESRs have a different background and field of expertise, the goal is to systematise the possibilities emerging from our interdisciplinary approach. We believe that this document will be useful not only for our colleagues within the project, but also for external readers who wish to get a better understanding of the empirical and modelling possibilities in the vehicle automation research.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

1 Empirical Strategies of Data Collection

This chapter serves as an overview of empirical approaches employed in SHAPE-IT. This text aims to highlight the inverse relationship between ecological validity and scientific control (Holleman, 2020) in empirical studies (see Figure 1). While naturalistic driving studies offer high ecological validity (data are collected from real vehicles in real traffic conditions), the experimenter has very limited options of how to control and/or manipulate the variables. Hence, the scientific control of naturalistic driving studies is low. Simulator studies are to be found on the other side of the spectrum. High scientific control of a simulator experiment allows the experimenter to minimize the effects of variables other than the independent variable. It increases reliability of the findings and allows for a replication of the experiment. However, it takes a toll on ecological validity of the results. Even in the best simulator, the participant is aware that the experiment is “just a simulation”, and it might be difficult to generalize the results for other settings (Kaptein, Theeuwes, & van der Horst, 1996). It is important to bear in mind that neither of these approaches is superior to the others, but rather it’s a matter of matching the right empirical approach to the question that is being asked by the researcher. The current state of scientific knowledge and the available technological tools can also impact what approach a researcher chooses (e.g., certain problem can be first studied in a simulator, then on a test track and/or on road, and further evaluated using naturalistic driving data).

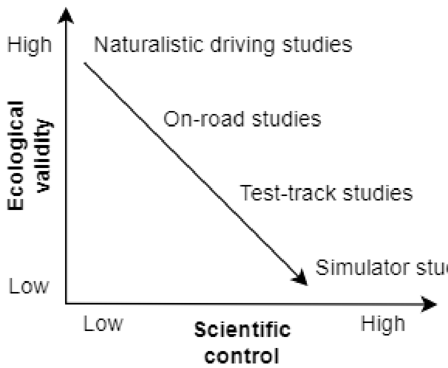

Figure 1 The inverse relationship between ecological validity and scientific control of selected empirical approaches

1.1 Naturalistic Driving Studies

Traffic is a highly complex social system, where different road users interact in shared space. Wilde (1976), argues that it is probably difficult to find instances of road user behaviour that are completely free from any form of social influence. Although there has been a considerable research to study the road user behaviour, mainly through conventional research methods (e.g., driving simulator studies or test track studies). However, these methods are inadequate to capture choices that drivers make in everyday situations (Van Nes et al., 2019). In a

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

naturalistic driving study (NDS), driver behaviour is studied during everyday trips by collecting details of driver, vehicle and surrounding traffic unobtrusively and without any experimental control (van Schagen et al., 2011). This way of collecting data ensures the highest possible ecological validity. Generally, drivers use their own vehicle and drives in normal manner. The vehicle is equipped with a dedicated Data Acquisition System (DAS) which is typically comprised of video cameras and other specialised data collection sensors (e.g., CAN, Radar, Lidar). A typical DAS is shown the Figure 2 Two extensive projects, involving collection of naturalistic driving data (NDD), are the UDRIVE (Van Nes et al., 2019) and SHRP2 (Hankey et al., 2016).

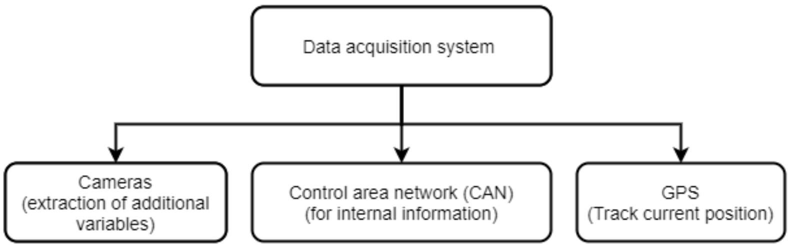

Figure 2 Typical DAS for naturalistic driving studies

First and foremost, NDD provide deeper insights into driver behaviour during everyday traffic situations. The advantage of NDS stems from its highly global perspective, and from there, guiding the researchers to focus on the critical aspects. For example, using NDD we can categorize driving behaviour into different driving styles or identify risk factors in everyday driving (Bärgman, 2016). In addition to that, the NDD is also used to investigate factors that contributes to crashes. Using actual crash data from SHRP2, Dingus et al., (2016) found that, driver-related factors (e.g., distraction) were present in almost 90% of the cases. In addition to that, the NDD is used to understand how driver behaviour affect the safety. The NDD also provide meaningful insights into observable impairment. Dingus et al. (2016) categorized impairment as (a) drug or alcohol impairment, (b) drowsiness and fatigue and (c) impairment caused by emotion (anger, sadness, crying). These statistics offer deeper insight to the fundamental issues that can lead to calamities. Moreover, the collection of such data acts as guidelines for researchers when designing novel strategies to reduce errors.

1.1.1 Data

Video data is the most common, unique, and important component of any typical NDS. The video data is collected unobtrusively over a longer time period. Video is useful to understand how individual drivers behave in certain situation and how do they interact with other road users. Depending on the scale of the study at least one camera facing forward towards the road is required. However, large scale studies often have multiple cameras (as shown in Figure 3) , both inside and outside of the vehicle. For example, the UDRIVE study installed 5 to 8 cameras depending on the type of vehicle. The cameras inside the cars are used to collect data on, facial expressions, view of the steering, view of drivers´ feet to capture use of brake and acceleration pedals etc.

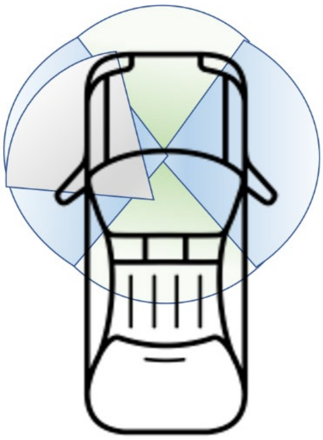

Figure 3 Typical camera placement for an NDS

The Controller Area Network (CAN) data is the second most important data type after the video data. Most of CAN data from a vehicle is available only to its manufacturer, supplier of the device, and/or trusted research organizations (Victor et al., 2010). However, the high-quality CAN data, such as – vehicle speed, 3-axis acceleration, use of brake and clutch, use of headlights windshield wiper, ambient light status etc. – is enough for typical ND research study. Speed, acceleration and brake pedal data is useful to understand pre-crash behaviour of the driver (Bärgman, 2016). While, other forms of data can be useful to understand factors that contributes to crashes. Furthermore, additional sensors (e.g., radar) can also be used to get some additional information about key features.

Questionnaires are also part of NDS, which are used to compliment in-vehicle sensor data. For example, in UDRIVE project, Van Nes et al. (2019), collected driver demographics, driver attitude (Parker et al., 1996) driving style (French et al., 1993), locus of control (Ozkan and Lajunen, 2005). In addition to that basic information about driver, Fridman et al. (2019), also collected questionnaires regarding initial impressions, and reported trust in select vehicle technologies. They also collected questionnaires to asses reported trust in vehicle technology, perceptions of safety and detailed understanding of the system after a month of naturalistic driving.

Although most of the data in NDS are collected using internal sources, some of the data can also be collected from external sources. The data from external sources mainly include road geometry data (e.g., number of lanes) and the type of road (rural or urban).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

1.1.2 Limitations

The large-scale NDS are preferred to investigate human behaviour in everyday traffic situations. However, the associated efforts and resources makes it very difficult to conduct such studies frequently. Below are the main limitations of NDS.

One of the main limitations of a NDS is the associated cost and time. NDS require vehicles to be equipped with dedicated Data Acquisition System (DAS) comprised of multiple cameras, GPS, and other sensors. Since NDS are aimed at collecting data over larger durations, there are also costs associated with the maintenance of the equipment. Furthermore, each person is also paid incentive for the participation in the study. For example, European NDS UDRIVE, paid 800 euros to each participant (Castermans, 2017). The total overall cost of the UDRIVE project was over 10.6 million euros.

Data retrieval and storage is another large limitation of NDS. The first challenge is faced during the installation of data acquisition system DAS. The installation and maintenance of DAS requires comprehensive training. For example, UDRIVE project developed and manufactured dedicated DAS along with comprehensive installation manual (Castermans, 2017). Another challenge is faced during data retrieval and storage. Since the onboard DAS has limited storage capacity, the collected data is downloaded from the onboard DAS and transferred to secure storage facility at regular intervals. For a more detailed perspective on challenges in data collection please refer to Castermans, (2017), section 7.

Suggested Reading

- Bärgman, J. (2016). Methods for analysis of naturalistic driving data in driver behavior research. Chalmers University of Technology.
- van Nes, N., Bärgman, J., Christoph, M., & van Schagen, I. (2019). The potential of naturalistic driving for in-depth understanding of driver behavior: UDRIVE results and beyond. Safety Science, 119, 11-20.
- van Schagen, I., & Sagberg, F. (2012). The potential benefits of naturalistic driving for road safety research: Theoretical and empirical considerations and challenges for the future. Procedia-social and behavioral sciences, 48, 692-701.

1.2 On-Road Studies

An on-road study – also a field study – is a broad concept in the field of vehicles and transportation. The on-road study approach can be defined as a study under controlled operating conditions by using experimental methods to evaluate a function or functions, or to investigate how the user/driver reacts to the vehicle and environment, or to research the impacts on transportation, environment and society (Barnard & Carsten, 2010). It is similar to NDS, but have some significant differences. Naturalistic Driving is a recently developed research method, observing road users' everyday driving behaviour. In an NDS, the subject will drive as he or she is accustomed to, without specific instructions or interventions, once the equipment has been installed. Participating drivers use the vehicles during their daily routines. Data are automatically recorded and the drivers do not receive special instructions about how and where to drive. In most of the cases, there is no experimenter or instructor in the vehicle, and on general the naturalistic driving (van Schagen & Sagberg, 2012). An on-road study, on the other hand, often comes with specific instructions or interventions. An important characteristic of on-road study is that it is a slightly more controlled experiment designed by using experimental design method to address specific research questions or hypothesizes. Compared to naturalistic driving study, on road studies are often aiming at evaluating certain functions such as an assistance system or a driving automation system.

Over the last decades, a large number of new technologies in vehicles have been successfully developed. For many applications, in particular in the field of advanced driver assistance systems (ADAS) and AVs, potentials in bring several benefits such as improved safety, reduced congestions and emissions, and enhanced mobility are expected (ERTRAC, 2017). However, what is not considered in this field is, under real traffic conditions, whether drivers will accept these new technologies and react to these new systems in the intended way. This might significantly affect the potentials of ADAS and AVs. Therefore, these systems and the assumptions on the potentials need to be confirmed in real life. On road study is one of the powerful tools to solve this problem.

In the light of limitations of the NDS, on-road studies offer the possibility to manipulate variables. The research questions are defined and specified prior the data collection, contrary to ND designs where the approach is mainly based on observations. When collecting the data, researchers do not need to obtain excessive amounts of data and can collect only research-relevant data. Conclusively, the data processing is typically easier compared to NDS. The parameters of on-road study design therefore enable us to ask specific questions regarding comfort, settings, design, interface, and so forth. Depending on the question, the focus of observation can be user-, vehicle-, and/or context cantered.

1.2.1 Scope

On-road studies can be categorized in many ways, but one way is to divide them into three types of foci:

- **User-centred tests:** addressing questions about user/driver reactions to automation, the control transition, drivers' situation awareness, the interaction between automated vehicles, their drivers and other road users, user acceptance/perceived safety and trust, etc. (Bazilinskyy et al., 2018; Portouli et al., 2007; Stapel et al., 2017)
- **Vehicle-centred tests:** addressing the question of how the automated vehicle behaves in different traffic conditions, questions about the interaction of the automated vehicle with the infrastructure need to be answered (Cafiso & Pappalardo, 2020).
- **Context-cantered tests:** addressing questions of how mobility changes, how this affects transportation, what ethical choices might be involved, and what would be the impacts on the environment and society (Karjalainen et al., 2014).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

1.2.2 Common Procedure

As for the test procedure of on-road studies, to what extent the standard test procedures should be defined is still under discussion. However, there are several steps of on road studies, which can be summarized as follows (Barnard et al., 2016):

- **Defining the study:** defining functions, use cases, research questions and hypotheses
- **Defining the target group(s):** in terms of age, gender, general driving experience, experience with specific vehicle functions like automated driving, racing experience, being professional drivers
- **Preparing the study:** determining performance indicators, study design, measures and sensors, vehicle instrumentation, and recruiting participants
- **Piloting the setup and procedures:** is everything working out as intended?
- **Conducting the study**: collecting data
- **Analysing the data**: storing and processing the data, analysing the data, testing hypotheses, answering research questions
- **Determining the impact:** impact assessment and deployment scenarios, socio-economic cost benefits analysis

1.2.3 Data

Trajectory data: path, velocity and acceleration of ego vehicle can be easily collected. Presence and motion of surrounding traffic can be gathered when access is obtained to the perception data of the vehicle or when such functionality is retro-fitted using commercially available or in-house designed sensor kits.

Driver's behaviour data: operation signal (steering wheel position, gas/braking pedal position) may be extracted from the vehicle, eye tracking data, hands/feet position, and physiological data require additional instrumentation and processing/annotation efforts.

The most important advantage is that on-road studies can provide fidelity and realism, compared to for example driving simulator studies, which makes the transfer of the results to actual traffic easier. Although the simulators become more and more advanced and realistic, the validity of it is still criticized. On-road experiments can directly evaluate certain functions such as an assistance system or a driving automation system in the actual environment in which it is intended to be used as a final product. It is therefore concerned with real-life manoeuvres and observations, which means that the ecological validity is satisfactory high as compared to simulator studies (see Figure 1), and allows researchers better scientific control than NDS.

1.2.4 Limitations

On-road studies involve observation in real-life environment, though slightly more controlled than in NDS. Though the validity is high, there still remain questions that are not easily answered. Using vehicles instrumented with technologies that are not proven on the road will not only be costly in terms of time and money, but also risky. For instance, the researchers cannot simply install a camera and wait for a collision to happen in order to understand the cognitive workload and the outcome behaviour of the driver. In this regard, on-road studies lack flexibility and the pool of questions to be asked is fairly small. The limitations are mainly due to the absolute condition of making sure the participant is safe. Other non-suitable questions for on-road study design implementations are precursors to crashes, such as fatigue or mind-altering substances (drugs, alcohol), and evaluation studies of new human-machine interface (HMI) designs that may alter safety.

Suggested Reading

- Bärgman, J. (2015). On the analysis of naturalistic driving data (Doctoral dissertation, Chalmers Tekniska Hogskola (Sweden)).
- Kiefer, R. J., LeBlanc, D. J., & Flannagan, C. A. (2005). Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. Accident Analysis & Prevention, 37(2), 295-303.
- Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation–Comparing the effects of different automation levels in an on-road experiment. Transportation research part F: traffic psychology and behaviour, 38, 67-82.
- Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. Ergonomics, 54(10), 932-942.
- Rendon-Velez, E., van Leeuwen, P. M., Happee, R., Horváth, I., van der Vegte, W. F., & de Winter, J. C. F. (2016). The effects of time pressure on driver performance and physiological activity: A driving simulator study. Transportation Research Part F: Traffic Psychology and Behaviour, 41, 150–169.
- Stapel, J., Mullakkal-Babu, F. A., & Happee, R. (2017, October). Driver behavior and workload in an on-road automated vehicle. In Proceedings of the RSS2017 Conference.
- Tenhundfeld, N. L., de Visser, E. J., Ries, A. J., Finomore, V. S., & Tossell, C. C. (2020). Trust and distrust of automated parking in a Tesla Model X. Human factors, 62(2), 194-210.
- Wilson, K. M., Yang, S., Roady, T., Kuo, J., & Lenné, M. G. (2020). Driver trust & mode confusion in an on-road study of level-2 automated vehicle technology. Safety Science, 130, 104845.

1.3 Test-Track Studies

From the last section we learned that, compared to NDS, an on-road study offers the possibility to manipulate variables to collect driver behaviours or human-AV interaction. But when there are surrounding traffic, it is still far from being a controlled experiment. Thus, the level of manipulation remains constrained. To solve this problem, test-track studies are often

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

considered an appealing candidate. Ranging from an intersection to a small town, test-tracks are highly controlled environments that could be used to collect data about driver behaviour, human-AV interactions, or the road infrastructure. Since all the traffic participants involving in the experiments are controllable, risky situations could be easily avoided and consistency on independent variables maintained throughout the experiment. Therefore, a great variety and combinations of traffic scenarios and participants can be set up in test track experiments in a safer way. Participants can be ranging from pedestrians, cyclists, drivers, or other road users in interaction with all sorts of vehicles (e.g., manually or automated vehicles trucks, trams). In a way, test-tracks could be said as being a simulation of a fraction of a traffic scenario, that is conducted in a safer manner.

Additionally, all sorts of traffic scenarios can in principle be realized, e.g., different types of intersections, highways, rural and urban traffic scenarios (Albert et al., 2015; Boda et al., 2018; Kundinger et al., 2020; Lotz et al., 2020; Szalay et al., 2018). Of course, this depends on the characteristics and the topography of the test-track area. Compared to real traffic situations, test-tracks are in more confined areas, with all interaction between traffic participants controlled and safety assured. Despite the lack of complex interactions existing in the real traffic, the better repeatability and data collection still make test-track study a compelling method in studying human-AV interaction. In addition, kinematic cues and risks are still perceivable to participants in test-track studies, which in turn enable researchers to acquire results more comparable with real-traffic. Conclusively, test-track studies offer minimum risks and at the same time offer a higher realm of manipulation in variables, compared to for example driving simulator studies, which enable us to answer a variety of questions. That is, depending on the specific question, different types of studies can be applied.

1.3.1 Types of Test-Track Studies

In AV research, the studies can be divided in two major categories: (a) studies using a functional prototype of a realistic automated vehicle (e.g., Albert et al., 2015; Kundinger et al., 2020), and (b) studies using a regular vehicle operated by a “wizard” who controls the vehicle (e.g., Habibovic et al., 2018; Poisson et al., 2020). Compared to on-road studies, test-track broadens the possibilities (e.g., by studying situations that could be dangerous in real traffic) while maintaining low risk ratio. For example, it would be contra-intuitive to allow participants to drive in a real-life traffic after spending a wakeful night as it clearly compromises safety. Kundinger and colleagues (2020) measured the effects of drowsiness in relationship to manual and automation driving modes. The test-vehicle was equipped to serve as a Level 2 automation. While measuring subjective KSS (Karolinska Sleepiness Scale) levels of drowsiness, researchers found that participants were drowsier in automation compared to manual driving. This could be explained by monotony of observation (Körber et al., 2015). On the other hand, while moving higher in automation level it seems to be that participants rather prefer to put as many tasks as possible to the AV, which suggests a raise in comfort level and trust in the system (Albert et al., 2015). Studying different level of automation and evaluating driver experience might therefore be very insightful when assessing trust and pre-conditions of trust in AV. In another study, Habibovic et al. (2018) examined whether the concepts of automated vehicle interaction principle (AVIP) showed by the external human-machine interface (eHMI) on the windshield could be correctly interpreted by pedestrians that were

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

about to cross the road. In the experiments, the AVIP was installed on the oncoming vehicle that was driven by a driving wizard acting as a passenger on the passenger seat, while on the driver seat was a fake driver with a steering wheel that was not functionable. The results suggested that after some trainings, most of participants were able to successfully interpret the signal of AVIP, and also felt more calmly when interacting with automated vehicles equipped with AVIP.

1.3.2 Data Collection

Most research regarding AVs focus on the behaviours of drivers, pedestrians or passenger. That is, all sorts of interactions between humans and AVs. In order to gather experiment results that is comparable to real traffic, on-road experiments are usually performed to serve this purpose. However, due to the lack of repeatability and multiple uncontrollable factors in real traffic, on-road studies are challenging to perform. As for the driving simulator, which is easy to set up and to control all parameters, it is argued that the lack of perceived risks and motion cues have impact on driver behaviour data (Boda et al., 2018; Poisson et al., 2020). What lies in between is the test-track studies. The collected data in test-track studies provide better ecological validity than in driving simulator with the risks and motion cues perceived, and is collected in a more controlled environment than on-road studies. If the data-analysis results from a test-track study meets the hypotheses and requirements, then the same research questions and procedures can be implemented in a real traffic environment for better ecological validity. If, however, the hypothesis is nullified, then a revision can be made and possibly even refined in a lower-level reality, i.e., simulator-based environment. Test-track studies are therefore appreciated when the research question resolves around obtaining valid subjective measures. With regards to objective measures, such as measuring response times and studying operating manoeuvres, test track design is typically good, but it is often not possible to create the truly naturalistic responses (e.g., response times) which would occur in real world settings. Conclusively, data from test-track studies may differ from data-collection in a naturalistic study as we expect different driving behaviour when the agent is safe contra not safe.

1.3.3 Limitations

Test-track studies would fall in-between the two ends of the spectrum; they are relatively high in both ecological validity and scientific controllability, compared to on-road studies. On the one hand, these are the advantages of test-track studies; on the other hand, these could also be a blindfold, making researcher oversee the fact that it is still some distance away from real implementation. The test-tracks used in research are often simplified: most of them are either a simple unsignalized intersection or a continuous circular track (Habibovic et al., 2018; Albert et al., 2015). Also, interactions between traffic participants in real world are often more complicated, with multiple participants interacting with each other. Despite being a method with high ecological validity, further research is needed that compare test-track studies to real traffic (Poisson et al., 2020; Habibovic et al., 2018; Feng et al., 2020). Another important factor to acknowledge is the safety aspect. It is very likely that participants behave differently in a safer environment than in the real-world (e.g., when encountering a dummy or a balloon car).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

While test-track studies offer real driving experience, it is also the case that the environment is controlled enough, i.e., that all measures for safety are accounted for (as far as possible). In one experiment, Weinbeer and colleagues (2017) demonstrated that participants, while drowsy, where still able to respond accurately to a take-over scenario. The study was conducted on a real highway in Germany which might suggest that humans are more alert when the safety is compromised. However, it should be noted that the request to intervene was rather a simple. Though test-track studies offer real life experience, it is only from a narrow perspective. The reason is because test-tracks are not designed to simulate complex driving scenarios with the wide range of any complex urban environment. This seem to limit the types of questions that can be answered through this type of design.

Suggested Reading

- Boda, C. N., Dozza, M., Bohman, K., Thalya, P., Larsson, A., & Lubbe, N. (2018). Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare? *Accident Analysis & Prevention*, *111*, 238-250.
- Habibovic, A., Lundgren, V. M., Andersson, J., Klingegård, M., Lagström, T., Sirkka, A., ... & Larsson, P. (2018). Communicating intent of automated vehicles to pedestrians. *Frontiers in psychology*, *9*, 1336.
- Kundinger, T., Riener, A., Sofra, N., & Weigl, K. (2020, March). Driver drowsiness in automated and manual driving: insights from a test track study. In *Proceedings of the 25th International Conference on Intelligent User Interfaces* (pp. 369-379).

1.4 Studies Conducted in a Simulator

Fundamentally, a simulation is a representative model of a real-life scenario – as far as possible within the constraints of the simulator used. It is used to investigate issues that are otherwise impossible to address under their respective natural circumstances Chang (2015). Simulator studies are exceptionally beneficial for studying critical situations, from which, new HMI or eHMI strategies arises.

Studies are conducted in highly controlled environments in order to answer questions that are limited by previous mentioned methods. Human Factors research is, for example, very interested in the behaviour of the driver or vulnerable road users (VRUs) during various conditions, such as dense/complex traffic environment, zebra crossing, lane change, intersections, severe weather condition and altered states of consciousness (i.e., alcohol, drugs) (Wang, Li, & Lu, 2014). These conditions, amongst other, can be studied interrelatedly or separately and create a large variety of complexities. Research has repeatedly shown how cognitive factors are hyperlinked and this creates complexities on many levels of analysis which ultimately give rise to further questions regarding the interaction with automated systems. Collecting massive data on driver performance and VRU behaviour is therefore key to understand traffic psychology and allows for some prediction models.

To investigate the relationship between various cognitive aspects and driving, and/or VRUs, researchers in this field have introduced terms such as *situation awareness (SA)*, *take-over request (TOR)*, *take-over-time (TOT)*, *response time*, *time to collision (TTC)*, *cueing*, *feedback-loop*, *non-driving related task (NDRT)*, *out-of-the-loop problem* and so forth. Studying these key-terms enable researchers to assess the necessary information needed for the driver/VRU to make the right judgements. Ultimately, researchers in this field aim to reach a consensus regarding these technicalities in order for the industry and policy makers to move in the right direction (providing larges positive impact on, e.g., safety). Situation awareness, in particular, has been used since the very beginning of research in this field and has influenced many guidelines with respect to TOR (for the driver) and better judgement (for a VRU) (Dey et al., 2020).

Investigating the domain of human cognition in an applied interdisciplinary approach is a challenge and must be done with caution. As simulator studies offer the possibility to control the environment to very high degree and to measure various aspects of behaviour in a very detailed way, simulators allow to investigate causal relationships between theoretical concepts and external factors. This means simulator experiments may provide a high internal validity of their results compared to more realistic settings that are less controlled. The downside of this characteristic is that it also makes the experiment very contextual and narrowed, as is true for any highly controlled experimental study. This results in specific answers to specific questions, but they do offer us the opportunity to study cognition and behaviour under circumstances that are otherwise impossible to conduct in real-life. In a simulator, we can generate a near-crash, or even crash scenarios to get insight on humans' ability to cognitively process information and how decisions are made accordingly. This in turn provides us with opportunities to make hypotheses about, for example, danger-zones when the VRU/driver is facing a potentially fatal traffic situation. Critical situations are thus only possible in simulator-based studies (and, possibly, test-track studies with, for example, non-human VRUS, such as dummies) as these are safer for the participant and offer great deal of understanding of the mind's eye in all possible driving scenarios (but typically a very specific scenario for each individual study). By gathering data from simulator studies, researchers are able to make predictions about human behaviour in real-life traffic. Nonetheless, the journey is yet ongoing and we need more collaboration between engineers and human factors specialist in order to offer safe and trustable automation in traffic (Chang, 2015).

1.4.1 Driving Simulators

Driving simulators are sophisticated tools that are used frequently today as we are in the pre-phase of self-driving car era. An instrument can be defined as a driving simulator based on two aspects; one is that the study should reflect participants' perception and behaviour toward how the instrument operates, and the other one is that the instrument should be able to simulate the concerned driving situation in a similar way in terms of input and output (Fox, 1960).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Figure 4 One of the most technically advanced motion-base high-fidelity simulator in the world today, available at the University of Leeds

There are several benefits of employing a driving simulator, for example: (1) a driving simulator is regarded as a safe place for human drivers to experience simulated automated driving; (2) it typically requires less resources (e.g., time, money and researchers' effort) but is more achievable and efficient compared with on-road experiments or test-track studies, which require technical realization or prototyping of the autonomous driving (with an exception when employing the Wizard of Oz method, which is discussed in another section); (3) it allows more experimental control from the experimenters so that results obtained can be targeting proposed research questions and hypotheses (Bellem et al, 2017; Fox, 1960). In sum, driving simulators offers great potential in simulating traffic scenarios that would be dangerous to implement in real-life. In this a highly controlled environment, novel ideas and domain-specific research questions are easily assessed.

Figure 5 Fixed-base driving simulator with 190° field of vision, available at the Ulm University

1.4.1.1 Tools

A variety of driving simulators with different complexities or fidelity exist (see the review of Václav et al., 2017). Some simple simulators are only used to elicit human responses to specific driving situations, for example, driving simulation presented on one display, or participants playing video-game-like driving tasks (e.g., Morton & White, 2013). With medium fidelity, the driving simulator may be equipped with a car-like cabin and large displays providing wider field of view (e.g., Eriksson et al., 2013). A driving simulator with high fidelity is regarded as more realistic with larger film screens and has at least 180° field of view (FOV). Rear viewing can be achieved with simulated rear-mirrors (e.g., Antonson et al., 2009; Louw et al., 2019). When the driving simulator is built closer to the real settings, the ecological validity of the simulation will increase. Yet, the high-fidelity driving simulator is very expensive and may not be necessary depending on the research purpose (Fox, 1960).

1.4.1.2 Data

Drivers' manual driving behaviour

Data about how drivers operate the vehicle during manual driving when they are totally in control can be obtained. Examples include the number of errors the driver made, subjective feedback, detailed driving operations relevant to accelerator, brake pedal, steering, etc., detailed vehicle kinematic data like speeding, turning, stopping etc. One example in the AVs context is the study of Louw et al. (2019); They studied drivers' speed choice when they were exposed to different road contexts leading to various risk levels. Road environments (e.g., presence of oncoming vehicles, roadside furniture, persistence of roadside risky elements etc.) were simulated in the high-fidelity University of Leeds Driving Simulator. These manual behavioural data could be used to build more human-like autonomous vehicle controllers.

Drivers' reaction to automated driving

Several studies using driving simulators have been conducted to understand human drivers' perceptual and behavioural reactions to automated driving. Partially automated vehicles need the driver to take over control in some case while the driver's attention might not be in the driving task. Research questions, like how quickly the driver can accomplish the transition from the vehicle being in control to manual driving, how attention of the driver is allocated, and what the influence of non-driving activities are and so on, have been raised and examined in driving simulators (e.g., Melcher et al., 2015).

The driving styles of highly automated vehicles have been reported to influence human occupants' comfort, perceived safety and trust (Bellem et al., 2018; Paddeu et al., 2020; Rossner & Bullinger, 2020). Different automated driving styles have been simulated. For example, participants were exposed to the different automated driving algorithms, whereafter they reported their feelings toward the autonomous driving styles (e.g., assertive versus defensive in Yusof et al., 2016). Besides subjective measurement, other physiological data of the driver can be collected in driving simulators, for instance, head movements, brain activity, electrodermal skin responses etc (e.g., Radhakrishnan et al., 2020).

User test for internal HMI design

HMI is a key factor in this field of research as it is critical to have a clear and straightforward communication between the driver and the AV. New strategies are tested frequently in order to find the most suitable two-way (driver-vehicle) communication technique. Some novel internal Human-Machine Interface designs have been tested in driving simulators, for example, head-up display (Liu, 2003). HMI research is safely evaluated in simulator studies before implementing it in real-life.

1.4.2 Other Simulators

1.4.2.1 Pedestrian Simulators

While pedestrians are vulnerable in urban traffic, the application of future AV can help monitor and assist their safe interactions with the potential to reduce the road injures. However, to which extent the AV is designed to understand and communicate with the pedestrian in a safe manner is unknown and how to improve this capability is still a major concern. Pedestrian simulators are therefore widely applied in this pre-stage of AV development and application through tracking the pedestrians' performances in virtual tasks experiencing various interaction configurations with controlled variables regarding the AV design. It evaluates the efficacy of AV and pedestrian communication in a safe and low-cost space with more complex, controlled and flexible scenarios. Meanwhile, it assesses the feasibility of external interfaces and perception of pedestrian towards AV, regarding their trust, acceptance and user experience. When pedestrians cannot reply on the explicit communication such as gestures, waving and eye contacts from the drivers, the inability of AV communicating its intents timely and efficiently could give rise to the risky consequences. By applying the pedestrian simulators, the imitated real-world interactions address these safety concerns thus assuring the pedestrian's safe interactions with AV.

Apart from the benefits mentioned above, controllability is the main advantages of the pedestrian simulators, when the virtual scenes can be built up and modified according to the variables in high experimental control. Data richness and quality is also a highlighted benefit when simulator enables the easier and quicker manner to gather timely data regardless the complicated and hazardous situations it could be in real world.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

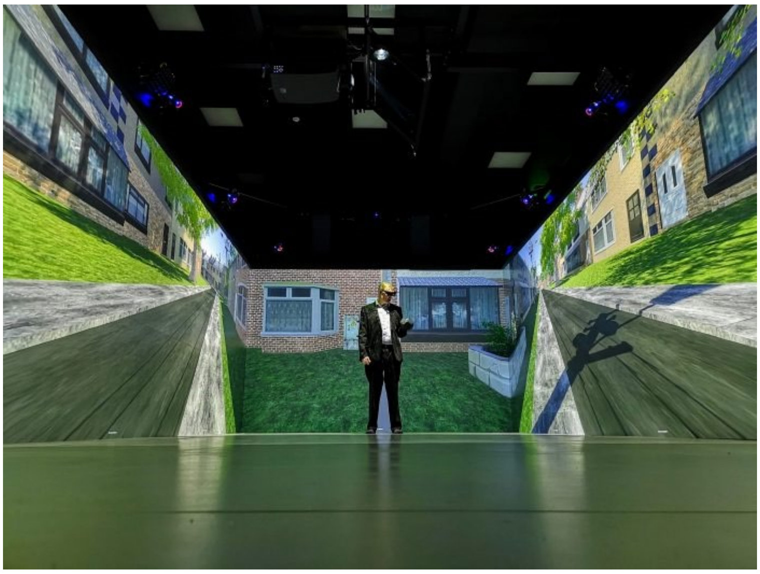

Figure 6 The Highly Immersive Kinematic Experimental Research (HIKER) lab, which is the largest 'CAVE-based' pedestrian simulation environment of its type in the world, available at the University of Leeds

1.4.2.1.1 Data

Behavioural data

In a pedestrian simulator, pedestrian crossing is the most common use case for studying the safe AV/VRU interactions. To better understand their decision-making process in the interactions, the classical behaviours data including average waiting time, average crossing time, average distance to collision, average time to stop, average jaywalking time, average crossing speed are collected. Further, the behaviour reflecting different decisions such as movement trajectories and route choices, are recorded. Additionally, the behaviour data combining the time series analysis and safety measurements assessment is gathered to present a more holistic reflection of pedestrian behaviours, such as time to arrival (TTA), time to collision (TTC), deceleration to safety time (DST), and post-encroachment-time. Learning from driver simulators, the eye tracking and the glances data are often recorded to assess the interactions. Moreover, head rotation, crossing initiating time (CIT) and hesitancy are also taken into consideration when conducting pedestrian simulations based on the research questions and variables.

Use test for external HMI design

Use test is the process of investigating the functions of an interface via testing with the users who will be asked to perform tasks in upfront designed scenarios. In order to evaluate the efficacy of external AV interfaces, use tests among pedestrians are conducted through techniques and followed by a user experience questionnaire. For example, Walker et al. (2019) designed a slider for pedestrians to show their willingness to cross over different external interfaces design. Additionally, the acceptance, trust and user experience of pedestrian can also be measured through the use test for an optimal AV external interface design.

Psychological and self-report data

In addition to the behaviour measurements, psychological data is also often, in AV research, collected through the self-report questionnaires in relation to the perceived safety, trust, acceptance and overall experience. Researchers often use, for instance, trust in automated systems questionnaire (Jian, 1998), perceived intelligence scale (Bartneck et al., 2009), usefulness and satisfaction scales from system acceptance scale (Van der Laan et al., 1997).

1.4.2.1.2 Tools

In the early time of investigation pedestrian behaviours, the researchers have been limited to the field studies and accident analysis since the experiments in real world are rare for the potential risks (Dietrich et al., 2018). Nowadays, new tools are emerging making pedestrian safe simulations accessible under the development of the technology. And this section will introduce these tools in a progression order.

VR technology

Pedestrian simulations are to date mainly based on VR technologies with motion tracking, monitoring pedestrian's intentions to cross. In VR simulations, pedestrians face a screen (typically a screen surrounding the pedestrian) showing the road scenarios in relation to which the pedestrian is to conduct the task of the study, such as pressing a button or verbally indicating the willingness to cross (Oxley et al., 2005). Further, Velde et al. (2005) conducted a crossing experiment by directing participants to cross the virtual road physically for the first time, introducing the exploration of physical interactions using VR simulators.

Head-mounted display (HMD)

The head-mounted display (HMD) simulator is a monoscopic display device worn on head. An HMD setup provides an unlimited field of regard (FOR) but a restricted field of view (FOV) (Dietrich et al., 2018). For example, the first HMD setup built by Simpson et al. (2003) can only provide a horizontal field of view (FOV) of 48°. However, the application of HUD enables the interactions to happen in the simulations by enabling the illusion of stereoscopy thus collecting more sorts of data like the head rotation and eye tracking.

CAVE-like device

The CAVE-like device (cave automatic virtual environment) was proposed when the research institute IFSTTAR enhanced their simulator. They used ten rear projections screens surrounding the pedestrians on the edge of virtual street with 180° FOV and 300° FOV for those in the middle.

Motion tracking

As pedestrian simulation laboratories have evolved, and new simulations have been created, more up-to-date tools have been used, including a control centre, motion capture cameras and system, high resolution HMD, self-developed motion suit, etc. Meanwhile, pedestrian avatars in the virtual world have been enhanced, using motion capture data with functional software and algorithms in the simulation.

1.4.2.2 Cycling Simulators

As shares of cyclists in urban areas have increased in the recent years, the use of cycling simulators to study cyclist behaviour has become more relevant (Herpers et al., 2008; Lee et al., 2017; Nazemi et al., 2019; O'Hern et al., 2017). Using a cycling simulator to study behaviour ensures a safe laboratory environment to test and evaluate data from AV-cyclist interaction, eHMI designs and road infrastructure design along with cyclists' reaction to future scenarios with AVs.

The, to date, most common approaches of research into human factors aspects of cycling have employed either subjective or objective measures to study the effect of different individual and environmental factors on cyclists' perception. The subjective measures were mainly obtained from surveys such as self-report data (Lawson et al., 2013; Abadi and Hurwitz, 2018). While objective measures are obtained from exiting spatial data or field audits (Schepers et al., 2011). However, the advancement in technology has opened the new opportunities to study cyclist's behaviour in a virtual environment using cyclists' simulators. For example, Nazemi et al., (2018), used cycling simulator to study the effect of road infrastructure design and environment properties on cyclist's perceived safety.

The ease of designing new scenario and controllability in cycling simulators allows us to collect data which was impossible otherwise. For example, using cycling simulator, it is possible to study how the presence of autonomous vehicles will affect the perceived safety of cyclists. Furthermore, by using physiological measures such as EEG, ECG and eye tracking we could study cyclists' internal states (e.g., cognitive workload, emotions, attention, and situational awareness).

1.4.3 Limitations of Simulator Studies

So far, this report has introduced various methods to acquire empirical data, all of which offer solutions to the pre-mentioned section, but poses further limitations. Simulator studies are no exception to this rule. When designing simulation-based studies, it is important to understand that the conclusions are bound to the experimental design and cannot be taken for granted (i.e., to assume that a certain behaviour will be the same even in real-life scenario). The risk-free simulation environment may deprive the subconscious reactions in hazardous situations thus making them less comprehensive and reliable. In simulator studies, participants may act in much different ways as compared to real-life situations. This of course is one major problem with simulator studies. This gap, that sets us out of reality, can be reduced to a certain degree by improving the feeling of reality.

Another general issue with simulator studies is the simulator sickness. This syndrome, similar to motion sickness, can be experienced during or after exposure to simulated environments. It often leads to symptoms such as dizziness, fatigue, difficulty concentrating, fullness of head, anxiety, and nausea. The symptomatology and severity of the malaise depends on many variables (e.g., age, gender, stress, anxiety), however, the severity of the symptoms usually increases with increased exposure time (Dużmańska, Strojny, & Strojny, 2018). Several theories explaining simulator sickness were formulated. The most prevalent theories in literature are the Sensory Conflict Theory (Reason and Brand, 1975) and the Postural Instability Theory (Riccio and Stoffregen, 1991).

Driving simulator studies come with several drawbacks. The main concern is whether or not the simulator study is able to provide as reliable and valid results as on-road studies (Shechtman, Classen, Awadzi & Mann, 2009). This issue is related to validity concerns of the driving simulator including physical validity and behavioural validity. Physical validity describes how capable a driving simulator is in terms of reproducing the physical environment (Blana, 1996), and can be compared to ecological validity. Behavioural validity means how driving behaviours responding to the simulated world match with that in the real world (Mullen, Charlton, Devlin & Bedard, 2001). As settings (e.g., parameters) of simulators employed in different studies vary, the specific research questions and corresponding designed driving tasks influence behavioural validity (Bellem et al., 2017; Mullen et al., 2001). On one hand, based on studies conducted in the driving simulator, some resulted insights have been proved to have satisfactory behavioural validity. For example, driving comfort, an important subjective experience to acceptance and trust of AVs, has been investigated if it can be studied in the driving simulator. Bellem et al. (2017) compared participants' comfort responses collected from the simulator and a test-track study. They have confirmed consistence of these comfort evaluations under specific simulator configurations (i.e., with proper lateral and longitudinal motion scaling factors). On the other hand, a study conducted by Boda and colleagues (2018) showed behavioural differences between participants performing the experimental task in the driving simulator contra those who performed it in a test-track. While approaching an intersection, a bicycle would cross. It was observed that participants in the driving simulator released the gas pedal much quicker after spotting the bicycle, than participants in the test-rack group. This might though be explained by lack of sensory cues (e.g., engine, kinematics (Hoffman et al., 2002), Researchers further analysed behaviour in terms of proactive (releasing gas pedal *before* spotting bicycle) and reactive (releasing gas pedal *after* spotting the bicycle) attitudes; and saw that drivers were 4 times more prone to be proactive in the simulator than in the test-track (Boda et al., 2018). One behavioural response however, seemed to be rather consistent between simulator and test-track environment. Boda and colleagues (2018) claims that brake onset was very similar between the two groups. This is useful as it suggests that behavioural data are not always merely bound to its environment and that it can indeed be some overlaps between different methodological approaches.

As for pedestrian simulators, the size of the simulator screen will limit the field of regard (FOR) in a CAVE-like simulator, and the HMD will limit the FOV as the limitations of the pedestrian simulator setting up (Dietrich et al., 2018). Meanwhile, the pedestrian behaviour and perception differ from physical world from the virtual environments, when they make crossing decisions based on temporal distances (real world) other than spatial distances (virtual test) of approaching vehicles (Feldstein & Dyszak, 2020), the pedestrian simulator is therefore critical in its behaviour validity.

Apart from motion sickness due to the lack of physical forces, physical resistance is a challenge for implementation of bicycle simulators. Technical limitations of bicycle simulator hardware and software may cause the bike to react too slow when adjusting pedalling resistance, and acceleration of the wheels, pedalling or braking may not match biking in the virtual environment (Schramka et al., 2017). Paired with technical limitations of VR hardware, the lack of realism is a limitation of simulator studies in general.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Suggested Reading

Driving simulators:

- Bellem, H., Klüver, M., Schrauf, M., Schöner, H. P., Hecht, H., & Krems, J. F. (2017). Can We Study Autonomous Driving Comfort in Moving-Base Driving Simulators? A Validation Study. *Human Factors*, *59*(3), 442–456. https://doi.org/10.1177/0018720816682647
- Fox, B. H. (1960). Engineering and psychological uses of a driving simulator. *Highway Research Board Bulletin*, *261*, 14–37. http://dx.doi.org/
- Václav, L., Petr, Z., & Aleš, Z. (2017). The Use of Driving Simulators in Psychological Research. *Sibirskiy Psikhologicheskiy Zhurnal*, *64*, 65–75. https://doi.org/10.17223/17267080/64/4

Pedestrian simulators:

- Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3
- Dietrich, A., Bengler, K., Feldstein, I. T., & Lehsing, C. (2018). Pedestrian simulators for traffic research: state of the art and future of a motion lab. *International Journal of Human Factors Modelling and Simulation*, *6*(4), 250. https://doi.org/10.1504/ijhfms.2018.10017411
- Feldstein, I. T., & Dyszak, G. N. (2020). Road crossing decisions in real and virtual environments: A comparative study on simulator validity. *Accident Analysis and Prevention*, *137*(February), 105356. https://doi.org/10.1016/j.aap.2019.105356
- Oxley, J. A., Ihsen, E., Fildes, B. N., Charlton, J. L., & Day, R. H. (2005). Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. Accident Analysis and Prevention, 37(5), 962–971. https://doi.org/10.1016/j.aap.2005.04.017
- Simpson, G., Johnston, L., & Richardson, M. (2003). An investigation of road crossing in a virtual environment. Accident Analysis and Prevention, 35(5), 787–796. https://doi.org/10.1016/S0001-4575(02)00081-7
- Velde, A.F., van der Kamp, J., Barela, J.A. and Savelsbergh, G.J.P. (2005) 'Visual timing and adaptive behavior in a road-crossing simulation study', Accident Analysis and Prevention, Vol. 37, No. 3, pp.399–406.
- Walker, F., Martens, M., Dey, D., Pfleging, B., Eggen, B., & Terken, J. (2019). Feeling-of-safety slider: Measuring pedestrian willingness to cross roads in field interactions with vehicles. Conference on Human Factors in Computing Systems - Proceedings, 1–6. https://doi.org/10.1145/3290607.3312880

Cycling simulators:

- O'Hern, S., Oxley, J., & Stevenson, M. (2017). Validation of a bicycle simulator for road safety research. *Accident Analysis & Prevention, 100,* 53-58.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

1.5 Types of Empirical Data Collected in SHAPE-IT

As SHAPE-IT research focus on human interaction with AV, the behavioural data we will collect in the SHAPE-IT project include the behavioural data of drivers inside AV and of VRUs. We plan to have several simulator studies, test track experiments and on-road experiments to collect different types of behavioural data. The summary of the behavioural data is classified into drivers inside the AV and VRU interacting with AV.

A short description of the different empirical approaches, simulator, test-track, and on road experiments will be explored. We aim to answer the following questions: what is the data to be collected, how it should be collected (equipment, for example), and why it should be collected.

1.5.1 (Automated) Vehicle-Driver Interaction Data

In the past years, Human-Machine Interaction has become a well-established field, which has fostered insightful research approaches and parameters when collecting relevant and useful data for the progression of the field. As a result, the prevalence of the science of human factors in user centred data collection and behaviour interaction techniques have given rise to a new but related approach: the use of behavioural data to understand and make assumptions about drivers/users' interaction patterns with automated vehicles. We argue that behavioural data give rise to new opportunities for human-machine interaction designers and analysts, as we move towards automated vehicles in urban environments. With the advent of massive new tools and techniques with which to collect, manipulate and analyse behaviour data, there is a need to scrutinise these tools or equipment in more detail, especially within SHAPE-IT.

In this project, we explore behaviour data collection as simply a collection of humans' driving behaviour information when interacting with an automated vehicle, but when it is combined with external variables – such as VRU behaviour (see section 1.5.2); it takes a holistic approach to understand how humans interact with automation, internally and externally. The interaction between these three concepts (driver – AV – VRU) gives rise to the demand for the collection of human behaviour data in human-machine interactions using simulator experiments and/or real-life experiments, among others. As a result, this type of data is quite important as it helps to improve the interaction between humans and automated vehicles, enhance safety assurance, and generate a comfortable experience, as well as foster trust and acceptance between humans and machines.

Accordingly, Table 1 presents examples of types of behaviour data that will play a major role in SHAPE-IT. Note, however that this list does not include all data to be collected in the duration of the project, not the least since research is a dynamic process were, for example, findings early in a project lead to modified study plans. Thus, it should be taken as a mere starting point (examples) to provide a rough overview for the SHAPE-IT ESRs planning their first studies.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table 1 SHAPE-IT behavioural data examples

Type of data	Description	Research procedure	How it should be collected (e.g., equipment)	Why it should be collected (e.g., its' importance)
Gas/braking pedal position	This involves information about how the driver brakes or pumps for gas while driving	The research process may include the collection of data on driver's performance in a driving simulator.	This information can be collected using a camera setup inside the vehicle in order to monitor drivers' braking-performance, reaction time and response time, and eye-movement.	This information is important as it helps in understanding how drivers can reach efficiency with AVs.
Braking acceleration,	This involves the drivers' tendency to brake acceleration when driving	The research process may include the collection of data on driver's performance and perceived safety ratings in a driving simulator, and possibly on-road experiment in order to evaluate these factors.	This information can be collected using a camera inside the vehicle connected on the brakes. This may also include eye-tracking techniques or equipment.	This information is important as it helps in understanding take-over effects when driving.
Steering angle and lateral acceleration	This involves information about the steering styles of the driver while driving	The research process may include the use of an output of the simulator (braking intensity and times, steering wheel angles, perhaps lateral acceleration, etc.), as well as evaluate drivers' behaviour in the form of observations and checklists.	This information may be collected using a camera inside the vehicle. An eye-tracker can be used to observe where drivers look (e.g., did he look in the rear-view mirror before changing lanes).	This information is important as it helps in understanding the effects that influence steering pattern and acceleration during driving.
Lane changing response	This involves information about how the driver changes their lane based on time dimensional factors	The research process may include the collection of behaviour data in the form of simulator experiments, and possibly on-road experiments.	This information may be collected using a camera setup inside the vehicle. The use of an eye-tracker to observe which information topics drives are looking at while reacting to over-take requests.	This information is important as it helps in understanding safety parameters in collaborative driving between a human and AVs.
Eye-glance movement	Usually, a device is mounted and tracks where in space the subject is looking and can record at what point in time the subject is looking, which can then be measured in conjunction with other data collections e.g., reacting to an obstacle etc.	Eye-movements are tracked through the device in any given scenario. Eye tracking can be measured in all sorts of experiments.	Data is collected by eye-tracking device. There are numerous versions of such devices.	Tracking eye-movements is very helpful in many ways, knowing for instance what type of object catches the attention of the driver. It's also helpful in deducting precise data of the subjects' reaction time to any given stimuli.

When collecting behaviour data of participants, it is important to pay attention to the following

- How the driver controls the driving tasks or process (for studies that have explored this aspect see Kolekar, S., de Winter, J. & Abbink, 2020; Tenhundfeld, N, L., et al., 2019; Beggiato, Hartwich, & Krems, 2019),
- The drivers' reaction time when initiation braking time (for studies that have explored this aspect see Tenhundfeld, N, L., et al., 2019; Cleij, Venrooij, Pretto, Pool, Mulder, & Bülthoff, 2018, Markkula et al., 2016) and transition frequency (for studies that have explored this aspect see Hecht, Kratzert & Bengler, 2020; Beggiato, Hartwich, & Krems, 2019).
- The time before activation (TB) and time before activation plus personal approach (TBP) (for studies that have explored this aspect see Danner, Pfromm, & Bengler, 2020)
- Information needs and visual attention during driving – eye tracking behaviour (see for studies that have explored this aspect Feierle, Danner, Steininger & Bengler, 2020; Hecht, Kratzert & Bengler, 2020), as this information helps in understanding where the driver looks when initiation a specific move or driving task.

Furthermore, it is important to note that most studies may encompass one or more of the behaviour data mentioned above. The structure of the environment and the road in which the driver moves are important, as it tends to influence the drivers' control signals, reaction time, steering, accelerating and braking, and eye movement when driving. Another influential factor to consider is that of the automated vehicle's level of automation, as this has the potential to substantially influence driver behaviour, perceived trust and acceptance.

1.5.2 Automated Vehicle-Vulnerable Road User Interaction Data

The interaction between VRU with AV is different from manual driving vehicles, as some interaction in manual driving, like gestures and eye contact between human driver and VRU, are removed or to be replaced (Löcken, Golling, and Riener 2019; Lundgren et al. 2017). Therefore, to understand the interaction between VRU and AV and further to evaluate safety of AV by VRU behaviour modelling, data collection of VRU behaviour is necessary.

The data planned to be collected include dynamics information of VRU, like crossing speed, crossing acceleration, crossing angle and trajectories, also VRU's reaction or waiting time when VRUs interact with AV's manoeuvres (for example, stop, start or other eHMIs). The posture information is also considered in the data collection as it is not obvious how gesture interaction will be changed in the AV driving scenarios. Eye movement may also be recorded in some studies by equipping VRUs with one eye tracking equipment.

The data mentioned above will be collected by combination of different equipment. Virtual reality (VR) and augmented reality (AR) environment are commonly used in simulator studies for VRU behavioural data collection (see section 2.4.2 for more information related to AR). Besides, camera placed inside the AV (front camera) and outside the AV, some signal sensors inside the AV like LIDAR, and sensors on the VRU are also considered in the data collection plan.

Suggested Reading

AV-driver interaction:

- Beggiato, M, Hartwich, F. and Krems, J. (2019). Physiological correlates of discomfort in automated driving. *Transportation Research Part F* 66. 445–458. doi.org/10.1016/j.trf.2019.09.018
- Cleij, D., Venrooij, J., Pretto, P., Pool, D.M., Mulder, M. and Bülthoff, H.H. (2018, February). Continuous Subjective Rating of Perceived Motion Incongruence During Driving Simulation. *IEEE Transactions on Human-Machine Systems,* Vol 48, No. 1
- Danner, S.; Pfromm, M.; Bengler, K. (2020). Does Information on Automated Driving Functions and the Way of Presenting It before Activation Influence Users' Behavior and Perception of the System? *Information*, *11*, 54.
- Feierle, A.; Danner, S.; Steininger, S.; Bengler, K. (2020). Information Needs and Visual Attention during Urban, Highly Automated Driving—An Investigation of Potential Influencing Factors. *Information*, *11*, 62.
- Hecht, T.; Kratzert, S.; Bengler, K. (2020). The Effects of a Predictive HMI and Different Transition Frequencies on Acceptance, Workload, Usability, and Gaze Behavior during Urban Automated Driving. *Information*, *11*, 73.
- Kolekar, S., de Winter, J. and Abbink, D. (2020). Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Applied Ergonomics* 89. doi.org/10.1016/j.apergo.2020.103196
- Tenhundfeld, N, L., et al. (2019). Trust and Distrust of Automated Parking in a Tesla Model X. *Human Factors* Vol. XX, No. X, Month XXXX, pp. 1–17. DOI: 10.1177/0018720819865412

AV-VRU interaction:

- Löcken, A., Golling, C., & Riener, A. (2019, September). How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 262-274).

1.5.3 Remarks for Ethical Implication in Collecting Drivers' Data

There are several key ethical principles that have to be followed when collecting empirical data within the SHAPE-IT project. For the purpose of this document, we refer to it as the "LANE model".

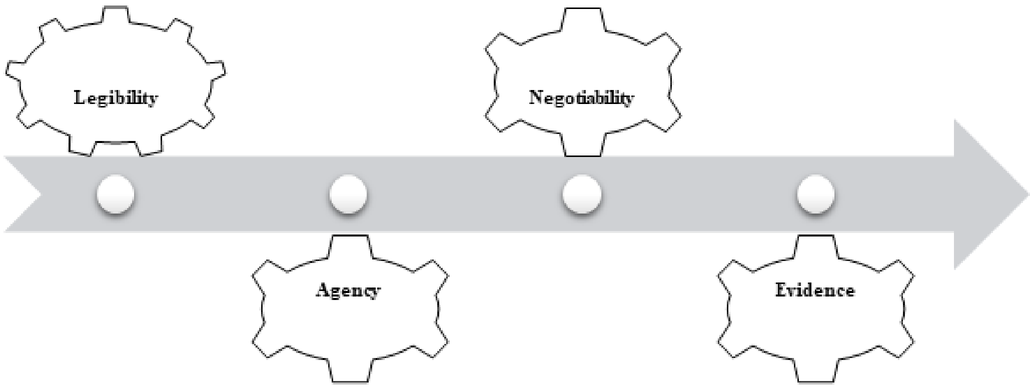

Figure 7 The LANE model - the errors represent a movement from one aspect of importance to another (including reversibility if need so) in collecting behaviour data.

When collecting empirical data in research studies, the ESRs need to make sure that the following are given careful consideration.

Legibility to ethical obligations – this is the process of making sure that the empirical data collected falls within the framework of ethical guidelines, and that it does not infringe on the human user's rights within the research process. Thus, the use of the data to be collected should be made clear to the participants, though without influencing the research results. This process is imperative from a scientific research ethical standpoint, and should aim to protect the intellectual property.

Agency for data consent – this process involves the participant's right to withdraw or withhold their consent to the use of their data during the data gathering process. This is an important factor and is aligned with the above notion or principle. As a result, this should be made clear before the research study commences.

Negotiability between participants and researcher – this is concerned with the societal and scientist contract surrounding the use of empirical data. After the data have been collected, its use should be aligned with research storage and the duration of its use. As a result, within the SHAPE-IT project, the aim is to keep to the ethical frameworks of scientific data use and storage of the empirical data that will be collected.

Evidence diffusion on publications – This is the process of making sure that the collected data are analysed in an ethical manner, and that all possible misrepresentation have been removed before publishing the results. As a result, proper techniques should be employed when analysing it. For example, the following questions should be addressed and clarified: What driver behaviour data is used from amongst the collected ones? What ways are inferences deduced from the drivers' behaviour data set?

2 Additional Tools for Empirical Studies

Collecting relevant, valid and reliable data is one of the most important factors in any study design. SHAPE-IT is an extensive project – studying road users both inside and outside the vehicle, along with attitudes and trust – which is why it's very important to account for all the relevant data that underlies the interaction between AVs and humans. Depending on the research question and the assessment of data, SHAPE-IT ensures that researchers are equipped with tools that allows for various angles in approaching a certain phenomenon, whether it is of qualitative and or quantitative data. These include physiological (including neurophysiological) data, behavioural data, interviews, self-reporting scales and observations. This section will therefore provide deeper insights on various methods, complementary to those presented in Chapter 1.

2.1 Interviews

Interviews are designed to gather in-depth information, and allow researchers to explore, explain and synthesize people's opinions, behaviour, experiences and attitudes towards a phenomenon. In SHAPE-IT, we are using a mix of different interviewing techniques, mentioned below. Depending on the precise research question, one technique will be more favourable than the other. The three most relevant types of interviews are semi-structured interviews, focus groups and structured interviews. To enable a full analysis of interview data, semi-structured and focus group interviews should be audio- or video-recorded and transcribed. However, recording an interview may impact the participants and the data collected. Whether you choose to record audio or video, or both, may also have ethical implications. The data recorded in interviews must be collected, stored and processed in accordance with GDPR and human research ethical guidelines. These guidelines allow for each researcher to derive valuable data from subjects that can be used to sharpen research questions and at the same time inspire for innovation. Table 2 brings an overview of main advantages and limitations of the semi-structured interviews, structured interviews, and focus groups. These are further discussed in the following sections.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table 2 Advantages and disadvantages of semi-structured, structured and focus group interviews

Type of interview	Advantages	Disadvantages
Semi-structured	Encourages reflection and exploration of more in-depth information	Questions need development, review and testing before implementation
	Flexibility: The interviewer can tailor follow-up questions to fit the topic and situation on the fly	Prone to social desirability bias, especially if the topic is sensitive
	Encourages two-way communication	Time-consuming
	Possibility to compare previous and future qualitative studies	Relies on the interviewer: Training is required
	Opportunity to learn the reasoning behind the answers	
	Suitable for interviewing about sensitive topics	
Structured	Standardised questions	Rigorous testing and piloting of questions
	Reliable	Rigid
	Replicable	Limited scope
	Cost effective: No training required	No in-depth information
	Interviewer can clear up misunderstandings	Interviewer cannot alter questions
	Reduced social desirability bias	
Focus groups	Easily measure reaction to a topic or evaluate a design prototype	Social desirability bias
	Timesaving by interviewing several persons at once	Requires moderator training
	Understand met and unmet needs	May miss out on important aspects due to group dynamic
	Receive feedback on a design strategy or product	Moderator bias
	Uncover new concept ideas	Sample may not be representative
	Explore decision making processes	Relies on group dynamic
	Encourages dialogue and interaction	Depends on moderator and group sample

2.1.1 Semi-structured Interviews

Semi-structured interviewing is a flexible and widely used method of data collection in qualitative research (Willig, 2008). Semi-structured interviews adhere to a pre-prepared interview guide, often phrased with open-ended questions with emphasis on narrative and experience to encourage reflection on the topic. What differs from structured interviews is the flexibility; semi-structured interviews allow the interviewer to ask follow-up questions or circle back to a topic brought up by the participant and elaborate their responses

Semi-structured interviews are suitable for exploring and synthesizing novel topics as they promote two-way communication allowing comprehensive discussions in the topic of investigation. Liu et al. (2020) explored the diversity of user acceptance of connected AVs by interviewing 36 experts in the field of AV cyber security and privacy. Based on the analysis, the authors provided recommendations for design strategies, how to mitigate the risk of cyber security and privacy dangers regarding connected AVs, and guidelines for developing trust among end users. In a driving simulator study of older people interacting with an HMI, Li et al.

(2029) utilized semi-structured interviews as well. Topics discussed involved driving behaviour, expectations on highly automated vehicles, opinions and propositions for manufactures, NDRT and driving style of AV. Qualitative data on elderly people provided some insight for researchers on preferences. In-depth data can guide researchers in design processes to fulfil traits such as comfort and acceptance (Li et al., 2019). For example, it was noted that the elderly preferred to hands-on opportunities, a good first-time experience with the AV as it goes a long way in developing trust (Eisma et al., 2003). These interviews are concentrated and researcher can be agile with follow-up questions which means that interviews might, up to a point, differ from one participant to another. This constrains the degree to which answers are comparable and reliable (as opposed to structured interviews). Another reported downside is that spontaneous questions might be biased.

2.1.2 Structured Interviews

Structured interviews are standardized, and in essence, researcher-administered surveys. This approach ensures the participants are presented with exactly the same questions in the same order. The strength of this approach is responses with high reliability that can be statistically analysed and compared between groups. The main difference from a self-administered questionnaire is that the researcher is present, asking the questions and collecting data. However, this opens up for social desirability bias, where participants tend to answer questions in a manner that can be viewed favourably by the researcher. On the other hand, while administering structured interviews, the researcher can clear up any confusing phrases or questions the participant may have regarding the data collection in person. As mentioned earlier, eHMI designs are not standardized yet, e.g., no consensus on where the interface should be placed (Bengler et al., 2020). Inviting participants to a study and gather large sample of data fairly quick (compared to semi-structured interviews) can shed light on preference and usability. Faas, Mathis & Baumann (2020) applied structured interview after a study exploring eHMI. The data suggested that pedestrians are more trusting and perceived the AV as intelligent if the eHMI at least provide information on its status. Researchers also discovered that information about what the AV is perceiving, did not yield any positive outcomes, rather, it was reported that the information disrupt traffic flow. Structured interviews are in this sense useful while allowing for reliable statistics. On the downside, it appears that there is no room for the participant to elaborate on any question which could mean that the response is not precise/intentional. This can occur if the questions are not thought through and in questionnaires with poor dimensional scales

2.1.3 Focus Groups

Focus groups, or group interviews, are suitable when the interaction between the participants is a vital source of data. As an alternative to semi-structured interviews, focus groups' advantage is the ability to challenge, develop and extend the participants statements through dialogue and interaction. The approach can be used to spark new insight or ideas, or to evaluate a design or product by plenary discussion. Focus groups must be moderated by an interviewer, whose task is to introduce the topic at hand and steer the discussion, prompting group members to respond and elaborate their statements. A focus group should not consist of more than six members (Willig, 2008).

2.1.4 Methods of Analysis of Interview Data

Method of analysis depends on the research question and is chosen accordingly. In transport research, thematic analysis has been applied as the method of analysis for several interview studies (see Alyavina et al., 2020; Gössling et al., 2016; Hafner et al., 2017; Liu et al., n.d.). Thematic analysis is a flexible approach for identifying, analysing and interpreting patterns or themes within qualitative interview data (Braun & Clarke, 2006). The methodology is systematic and rigorous enough for determining the credibility and validity of the process (Nowell et al., 2017).

Other methods of analysis that may be suitable for interview data in SHAPE-IT are discourse analysis, grounded theory and interpretative phenomenology (Willig, 2008). In instances where the objective of the research is so congregate and summarize expert opinions, less vigorous methods may be applied. In the SHAPE-IT position paper, Tabone et al. (2020) utilized a semi-structured interview approach by interviewing 16 Human Factors' researchers. Their responses were summarized by the main and second author and sent to the interviewees for approval. The aim of the paper was to gather expert opinions from independent Human Factors' researchers regarding their perspectives on automated vehicles and interactions with vulnerable road users in future urban environments. Similar approaches have been used in other studies in the transport sector, such as Kyriakidis et al. (2019), along with other expert consensus papers from working groups and workshops (e.g. International Transport Forum, 2019). In these types of papers and reports, the findings are summarized and discussed without the use of systematic, qualitative methodology.

Suggested Reading

Paper comparing different types of pattern-based, qualitative methods of analysis:

- Braun, V., & Clarke, V. (2020). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, capr.12360. https://doi.org/10.1002/capr.12360
- Saldaña, J. (2009). *Coding Manual for Qualitative Researchers*. SAGE Publications Inc. Saldana, J. (2015). *The Coding Manual for Qualitative Researchers* (3rd ed.). SAGE Publications Ltd.

Practical guide to cognitive interviewing techniques:

- Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design*. sage publications. Retrieved from https://www.dea.univr.it/documenti/OccorrenzaIns/matdid/matdid823948.pdf

2.2 Questionnaires

Questionnaires and surveys are useful tools for collecting data such as personality characteristics, opinions, attitudes and experiences from a large sample of participants, as well as participant demographics (Langdridge & Hagger-Johnson, 2009). There are some general design principles to mind in order to collect a reliable and valid set of data. The following principles are adapted from Langdridge & Hagger-Johnson (2009):

- Keep the survey as brief as possible. The response rate goes down with increased length. Adjust the language to your target sample. Avoid technical jargon and terminology.
- Ask one question at a time.
- Phrase your questions as unambiguous as possible.
- Start with the easiest questions and gradually build up to more difficult ones.
- Phrase the questions in a neutral manner.
- Pilot your questionnaire before administering it to your study sample.

Measuring psychometrics such as attitudes or opinions are often done with Likert scales (Likert, 1932). Likert scales are constructed in a five- or seven-point scale where the participant expresses their degree of agreement with a statement (from disagree to agree). To assess reliability of the measurement scale, tests for Cronbach's alpha coefficient or split-half reliability may be applied (Langdridge & Hagger-Johnson, 2009).

As an alternative to Likert scales, semantic differential scales (Osgood et al., 1957) are also well-suited for measuring attitudes. This approach is more indirect than Likert scales. Semantic differential scales build on humans' ability to think in metaphors and to draw parallels between different experiences. With these types of questions, the participants are asked to indicate thoughts or feelings based on a scale of antonyms, for example on a scale from good to bad, or from interesting to uninteresting. Semantic differential scales have shown to be reliable and correlate well with other scales measuring attitudes (Langdridge & Hagger-Johnson, 2009).

However, it is important to discuss the reliability and validity of scales and questionnaires here. When developing their own measures, scholars often apply measurement building procedures that are inconsistent with best practices. Test/scale development is an extensive scientific discipline, that involves numerous theoretical, methodological, and statistical competencies (Carpenter, 2018). Simply put, items have to be designed in a certain manner (e.g., linguistically, logically, following certain theoretical background) and a pilot study should be run. An item analysis should be performed (regrading, for example, internal consistency, inter-item and item-total correlation, or attenuation effects). Once the quality of the items seems satisfactory, data from a larger sample should be collected in order to perform further statistical and psychometric analyses to evaluate, to name a few, the internal consistency, stability in time, construct validity, or convergent validity of the measure (for a brief overview of the scale development process, see Carpenter, 2018). This is the only way to assure that the questionnaire is reliable (i.e., the amount of random error in the measure is minimised) and valid (i.e., the questionnaire measures what it claims to measure). If done correctly, this process is extremely time-consuming, very expensive, and requires researcher's expertise in

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

the discipline of test development. Therefore, we strongly recommend researchers to use already validated questionnaires instead of constructing new ones or using unvalidated questionnaires. For interested readers, we recommend reviewing keywords such as “classical test theory” and “item-response theory”.

2.2.1 Questionnaires in Automated Vehicle Research

In AV-human interaction research, questionnaires can be administered to measure behavioural data on various topics and aspects. For instance, Hagenzieker et al., (2020) measured behavioural intent and expectations of cyclists when interacting with AVs, while other studies have collected survey data on crossing intentions and decisions of pedestrians (De Clercq et al., 2019; Nuñez Velasco et al., 2019), and trust in AVs before and after take-over situations (Gold et al., 2018). Questionnaires can also be used to evaluate HMI and eHMI design strategies (Bazilinskyy et al., 2019; Fridman et al., 2017), as well as for psychometric testing, like measuring how personality traits affects human-technology interaction (Attig et al., 2017). Table 3 presents some of the questionnaires that will be used in the SHAPE-IT research projects, and may serve as an inspiration for others when planning their study.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table 3 Questionnaires considered in SHAPE-IT research activities

Domain	Method	Reference
Relationship to technology		
	Trust in automation LETRAS-G (German questionnaire)	*Kraus, J. M. (2020, August). Psychological processes in the formation and calibration of trust in automation. https://doi.org/10.18725/OPARU-32583*
	Inter-cultural Scale to Measure Trust in Automation	Chien, S. Y., Semnani-Azad, Z., Lewis, M., & Sycara, K. (2014, June). Towards the development of an inter-cultural scale to measure trust in automation. In *International conference on cross-cultural design* (pp. 35-46). Springer, Cham.
	Propensity to trust in automation	Merritt, S. M., Heimbaugh, H., LaChapell, J., & Lee, D. (2013). I trust it, but I don't know why: Effects of implicit attitudes toward automation on trust in an automated system. *Human factors*, *55*(3), 520-534.
	Trust in technology	McKnight, H., Carter, M., & Clay, P. (2009). Trust in technology: development of a set of constructs and measures. *Digit 2009 proceedings*, 10.
	Trust in specific technology	Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on management information systems (TMIS)*, *2*(2), 1-25. https://doi-org.tudelft.idm.oclc.org/10.1145/1985347.1985353
	Technophobia scale	Sinkovics, R. (2014). Technophobia scale. *Compilation of social science items and scales (ZIS)* . https://doi.org/10.6102/zis62
Immersiveness of virtual environment		
	Presence Questionnaire	Singer, M. J., & Witmer, B. G. (1999). On selecting the right yardstick. Presence: *Teleoperators and Virtual Environments, 8*(5), 566–573. https://doi.org/ 10.1162/105474699566486.
Personality		
Locus of Control	Internal-External Control, IE-4 (German version)	*Kovaleva, A., Beierlein, C., Kemper, C. J. & Rammstedt, B. (2014). Internale-Externale-Kontrollüberzeugung-4 (IE-4). Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). https://doi.org/10.6102/zis184*
	Scale of Internal and External Control (German Version)	Rost-Schaude, E., Kumpf, M., & Frey, D. (2014). Interne-Externe Kontrolle. *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).* https://doi.otg/10.6102/zis128
Self esteem	Rosenberg Self-Esteem Scale (German version)	*Ferring, D. & Filipp, S.-H. (1996). Messung des Selbstwertgefühls: Befunde zu Reliabilität, Validität und Stabilität der Rosenberg-Skala. Diagnostica, 42(3), 284-292.*
Self-Efficacy	Self-Efficacy Scale - Short Form (German version)	Beierlein, C., Kovaleva, A., Kemper, C. J. & Rammstedt, B. (2014). Allgemeine Selbstwirksamkeit Kurzskala (ASKU). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).* https://doi.org/10.6102/zis35
Impulsivity	Impulsive Behaviour Scale-8 (German version)	Kovaleva, A., Beierlein, C., Kemper, C. J. & Rammstedt, B. (2014). Die Skala Impulsives-Verhalten-8 (I-8). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).* https://doi.org/10.6102/zis183

Sensation seeking	Arnett Inventory of Sensation Seeking (AISS) (German version)	Roth, M. & Mayerhofer, D. (2014). Deutsche Version des Arnett Inventory of Sensation Seeking (AISS-d). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)* https://doi.org/10.6102/zis73
	Arnett Inventory of Sensation Seeking (AISS) (English version)	Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences, 16*(2), 289-296. doi: 10.1016/0191-8869(94)90165-1
Sensation seeking	Brief Sensation Seeking Scale (BSSS)	Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and individual differences*, *32*(3), 401-414.
Big five	Big Five Inventory-10 (German Version)	Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2014). Big Five Inventory (BFI-10). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).* https://doi.org/10.6102/zis76
Susceptibility to motion sickness	Motion Sickness susceptibility Questionnaire (MSSQ)	Van Emmerik, M. L., De Vries, S. C., & Bos, J. E. (2011). Internal and external fields of view affect cybersickness. *Displays, 32*(4), 169–174. https://doi.org/ 10.1016/j.displa.2010.11.003
Mood		
	Positive and negative Affect Schedule (PANAS) (German version)	Janke, S. & Glöckner-Rist, A. (2014). Deutsche Version der Positive and Negative Affect Schedule (PANAS). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS*). https://doi.org/10.6102/zis146
Motion/Simulator sickness		
	Misery Scale (MISC),	Van Emmerik, M. L., De Vries, S. C., & Bos, J. E. (2011). Internal and external fields of view affect cybersickness. *Displays, 32*(4), 169–174. https://doi.org/ 10.1016/j.displa.2010.11.003
	Simulator Sickness Questionnaire (SSQ)	Van Emmerik, M. L., De Vries, S. C., & Bos, J. E. (2011). *Internal and external fields of view affect cybersickness. Displays, 32*(4), 169–174. https://doi.org/ 10.1016/j.displa.2010.11.003
Workload		
	NASA-TLX	Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research*. In Advances in psychology* (Vol. 52, pp. 139-183). North-Holland.
	The driving activity load index (DALI)	Pauzié, A. (2008). A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems, 2*(4), 315. https://doi.org/10.1049/iet-its:20080023
Situation Awareness		
	Situation Awareness Rating Technique (SART)	Taylor, R. M. (1990). Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design. *Situational Awareness in Aerospace Operations (AGARD-CP-478).* pp. 3/1 – 3/17, Neuilly Sur Seine, France: NATOAGARD.

Suggested Reading

- Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). Web survey methodology. Sage. http://www.websm.org/uploadi/editor/doc/1443689013razsirjenoKazalo.pdf
- Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, *12*(1), 25-44.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: the tailored design method. John Wiley & Sons.
- Kraus, J., Scholz, D., & Baumann, M. (2020). What's Driving Me? Exploration and Validation of a Hierarchical Personality Model for Trust in Automated Driving. *Human Factors*. https://doi.org/10.1177/0018720820922653
- Oppenheim, A. N. (1992). Questionnaire Design, Interviewing and Attitude Measurement. London: Cassell.
- Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). The science of web surveys. Oxford University Press.

2.3 Physiological Measures

Traditionally, performance-based data (e.g., take-over time, minimum time to collision) and self-assessment methods (e.g., questionnaires, interviews) are used to assess driver's cognitive states. With the development of low-cost, non-invasive, wearable sensors a new possibility to assess driver's cognitive states via psychophysiological signals emerges. Such data provide a broad picture of the internal states (e.g., cognitive workload, emotions, attention, and situational awareness) that drivers experience, and allow to reflect on how driver's cognitive states are affected by non-driving related tasks (activities that are not related to driving, e.g. operating comfort or infotainment systems, communicating with passengers or remote people, eating and drinking; Pfleging & Schmidt, 2015), vehicle configurations, or driving environments (Du, Yang, & Zhou, 2020).

Measuring the changes in central and peripheral nervous system functioning is a possible way to improve the assessment of cognitive states (e.g., attention, perception, decision making). Such assessment can be done continuously, in an unobtrusive way, without disturbing the real-time task. Psychophysiological measures complement and extend performance-based metrics (e.g., reaction times). Moreover, as humans may not always be accurate in making judgements about their cognitive states, psychophysiological measures improve assessments of motorists' state-level changes in cognition. Applying a multi-method approach – combining the subjective, performance-based, and physiological measures – provides an additional value in automated driving research, and for future applications in driver monitoring systems, adaptive alert systems, and neuroadaptive HMIs (Lohani, Payne, & Strayer, 2019).

Commonly used techniques in vehicle-related research, which are also considered to be implemented within SHAPE-IT, include electroencephalography (EEG) and event-related potentials (ERPs), pupillometry, electrocardiography (ECG), electrodermal activity (EDA), and Electromyography (EMG). In the following sections, these techniques are briefly introduced and discussed in the context of driving research. Main advantages and disadvantages of each

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

technique are introduced. However, these are very context-dependent, and should be critically evaluated for every research design. The goal of this chapter is to introduce and briefly discuss the opportunities brought to driving research by implementing psychophysiological measures. Many techniques (such as respiration and blood pressure measures, functional near infrared spectroscopy, or thermal imaging) were omitted, as the aim of this chapter is to focus on the techniques considered within SHAPE-IT rather than bringing a comprehensive overview of all psychophysiological techniques. At the end of the sections, advantages and limitations of the presented techniques are overviewed.

Figure 8 A 32-channel active electrode EEG system that will be used in driving simulator studies at Ulm University

2.3.1 Electroencephalography and Event-related Potentials

Electroencephalography (EEG) is a non-invasive neuroimaging technique which allows to record electrical activity of the brain. It has high temporal resolution (in milliseconds), is portable, and relatively low-cost. It allows to study brain functions such as memory, vision, motor imagery, emotion, or perception (Malik & Amin, 2017). EEG electrodes are placed on a pre-defined position of the scalp (usually using the 10–20 system) and continuously record the neural activity. This activity is produced mostly by post-synaptic potentials of cortical pyramidal cells, which play a critical role in advanced cognitive functions (Lohani, Payne, & Strayer, 2019). The activity of a single neuron is not measurable on the scalp, therefore most of the activity measured by EEG consists of summation of the activity of big populations of neurons with similar spatial orientation (Kamel & Malik, 2015).

Neurons can generate action potentials in a rhythmic pattern (oscillatory activity). Spectral information about the oscillatory activity may be decomposed from the EEG signal using the Fourier transform (Lohani et al., 2019). There are five major types of oscillatory activity based on their frequencies: gamma (above 30 Hz), beta (13-30 Hz), alpha (8-13 Hz), theta (4-8 Hz), and delta (0.1-4 Hz).

In driving research, the most commonly studied frequency bands are alpha and theta waves. Alpha and theta waves are related to each other in a reciprocal (opposite) way – with increase in theta power, we can observe decrease in alpha power, and vice versa (Klimesch, 1999).

Previous research suggests that increased theta and reduced alpha powers are closely associated with increasing mental workload (e.g., Brouwer et al, 2012, Diaz-Piedra, Sebastián, & Di Stasi, 2020, Missonnier et al, 2006), while fatigue increases alpha power (Käthner et al., 20014). Increased alpha power was also observed during the relaxed condition compared with the engaged condition in an autonomous driving setting (Zander et al., 2017).

The measurement of the brain's response to a specific stimulus is called event-related potential (ERP), and is the most widely used method in cognitive neuroscience research to study physiological correlation associated with information processing (Kamel & Malik, 2015). ERPs allow distinguishing perceptual, cognitive and motor processes implicated in complex situations (Paxion, Galy, & Berthelon, 2014). In human factors and ergonomics, ERPs are used to study aspects such as vigilance, mental workload, fatigue, adaptive aiding, stressor effect on cognition, and automation (Brookhuis & de Waard, 2010). One of the most widely studied ERP component is the P3 (or P300). The P3 (comprising the P3a and P3b components) is a positive deflection in voltage, which appears roughly 250 to 500 ms post-stimulus, and is a well-established parameter for analysing cognitive functions such as attention and memory (Protzak & Gramann, 2018). Reduced ERP amplitudes were previously linked to under-arousal states (such as fatigue, time on task, lower vigilance) as well as to over-arousal state (such as high workload) (Lohani et al., 2019).

2.3.2 Pupillometry

The pupil permits light to enter the eye and reach retina. Its diameter is controlled by two sets of smooth muscles in the iris, constrictor muscles decreasing its diameter and dilator muscles increasing it. The diameter changes are meant to optimise vision via modulating the amount of light that reaches retina. Since 1960s, researchers also study the changes in pupil diameter as an index of cognitive functioning (Sirois & Brisson, 2014). Previous research found that pupil diameter increases with task difficulty, mental workload, emotionality of stimuli, and information-processing demands (Beggiato, Hartwich, & Krems, 2018, Strauch et al., 2019).

Pupil diameter is a continuous variable. Researchers often evaluate fluctuations of pupil diameter as a function of time-locked events (such as onset of specific events at precise time points in video or real-world sequences). An eye-tracking device is used to collect the data. It is a non-invasive, low-cost method which can be used in various contexts (Sirois & Brisson, 2014).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Figure 9 Eye-tracking device, available at the Ulm University

To ensure that the changes in pupil diameter is caused by changes in emotional or cognitive states, and not purely by the light intensity, it is important that researchers control for the amount of light that the participant is exposed to. Palinko and Kun (2012) suggest that, in certain situations, it is possible to separate the effects of illumination and visual cognitive load on pupil diameter. Nonetheless, interpreting changes in pupil size outside of controlled laboratory settings becomes a major challenge due to its heavy dependence on ambient light (Beggiato et al., 2018).

In the context of driving, pupillometry has been often studied as a measure of cognitive workload (Dlugosch, Conti, & Bengler, 2013; Palinko & Kun, 2012; Schwalm, Keinath, & Zimmer, 2008, Strauch et al., 2019). Moreover, Beggiato et al. (2018) used pupillometry to study discomfort in automated driving. They observed that pupil diameter increased significantly during the discomfort interval and decreased steadily after reported discomfort.

2.3.3 Electrocardiography

Electrocardiography (ECG, sometimes also EKG) is a method to record the electrical activity of the heart muscles. ECG is a non-invasive technique using surface electrodes placed on the skin. The number of electrodes and their location vary. Three electrodes are commonly used in research settings, placed under the left and right clavicular line and above the hip of the participant. ECG produces an electrocardiogram (see Figure 10), which is a recording of voltage changes over time. It comprises the P wave, QRS complex, and T wave. The R component has large magnitude, hence is easily detectable (Lohani et al., 2019).

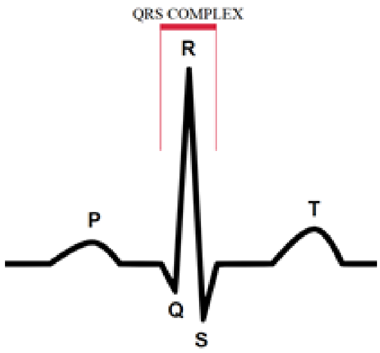

Figure 10 Example of one heart beat recorded on ECG

Heart rate (HR) and heart rate variability (HRV) are commonly used techniques of ECG analysis in the context of driving research. HR is the number of heartbeats in one minute. HR is generally derived by converting mean heart period (the time between two successive R spikes) to heart rate in beats per minute (Lohani et al., 2019). HRV is the fluctuation in the intervals between adjacent heartbeats, which are generated by heart-brain interactions and autonomic nervous system processes. There is a number of ways to quantify HRV (for an overview, see Shaffer & Ginsberg, 2017).

Previous research has shown that HR can be used as a physiological index of arousal induced by driving demands. In over-arousal cognitive states (such as high mental workload), the HR increases. On the contrary, a significant drop in HR is observed in under-arousal states (such as drowsiness). HRV also varies with workload experienced by drivers. HRV decreases with increasing task demands, while drowsiness led to higher HRV (for an overview, see Lohani et al., 2019).

2.3.4 Electrodermal Activity

Electrodermal activity (EDA, also known as a galvanic skin response, or skin conductance) refers to the variation of the electrical properties of the skin in response to sweat secretion. The activity of sweat glands is controlled exclusively by the sympathetic autonomous nervous system (ANS), and plays an important role not only in thermoregulation, but it is also a concomitant of emotional arousal. EDA is frequently used method in psychophysiological research. It is fairly easy to obtain a distinct electrodermal response, recording is possible with inexpensive equipment, is non-invasive, and can be done both in laboratory and field conditions (Benedek & Kaernbach, 2010).

EDA can be measured via exosomatic or endosomatic techniques. Exosomatic techniques apply a small current through a pair of electrodes and measure electrical resistance from the skin (either direct current or altering current). Endosomatic techniques use only potential differences originating in skin itself (Boucsein, 2011).

The time series of EDA comprise tonic activity (slowly varying, skin conductance level; SCL) and phasic activity (fast varying, skin conductance response; SCR). Phasic activity may reflect stimulus-specific responses. The event-related activity is commonly assessed by gauging the amplitude of the elicited SCR, which appears in a predefined response window, typically 1–3 s or 1–5 s after the stimulus (Benedek & Kaernbach, 2010).

Higher EDA is indicative of physiological arousal, caused by increased sympathetic ANS activity. It has shown to be related to many cognitive states, such as workload, stress, anxiety, or sleepiness. However, EDA is also sensitive to physiological reactivity and not all individuals have the expected SCR. EDA is also influenced by many other factors, such as respiration or mental effort. Therefore, interpreting EDA results in an applied and less-controlled setting should be done with caution, as it is sensitive to many psychological variables. It has also lower temporal resolution, as the response appears only 1-3 s after the stimulus onset (Lohani et al., 2019).

Cognitive workload in the driving context has been often investigated using EDA. SCL and SCR amplitude is higher during increased workload and stressful events. Higher SCL could be also an indicative of lower levels of trust in automation (for an overview, see Lohani et al., 2019; Mühl et al., 2019).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

2.3.5 Electromyography

Electromyography (EMG) is a discipline that deals with the detection, analysis, and use of the electrical signal that originates in neuromuscular activation associated with a contracting muscle. Muscle fibres are innervated in groups called motor units, which generate a motor unit action potential when activated. With increasing force in the muscle, more and more motor units are activated. Concurrently activated motor units superimpose to create the EMG signal. The EMG signal can be recorded using either surface electrodes placed on the skin, or inserted electrodes (wire or needle). In psychophysiological research, the unobtrusive, non-invasive surface electrodes are generally preferred (De Luca, 2006).

The surface EMG electrodes capture an activity of specific muscles of interest. Numerous features can be extracted from EMG signal (see De Luca, 2006). Commonly, the root mean square of the signal is reported, as well as peak spectral density, peak amplitude, and peak frequency. EMG can provide insights into emotional processes, as well as into mental processes such as stress. However, it might be challenging to tease apart muscular activity due to other confounding reasons (such as posture changes, scratching skin) from activity relevant changes in cognitive states (Lohani et al., 2019).

In the context of driving, EMG was studied in relation to stress and fatigue. Increase in muscular tension of the trapezius muscle was associated with greater stress exposure (Lee et al., 2017). EMG system was also used for online detection of the level of drowsiness (Artanto, Sulistyanto, Pranowo, & Pramesta, 2017).

2.3.6 Advantages and Limitations

Table 4 brings an overview of the major advantages and disadvantages of the abovementioned psychophysiological measurement techniques. For better understanding of the advantages and limitations, we recommend to consult the suggested readings, especially a recent paper by Lohani et al. (2019).

Table 4 Advantages and limitations of chosen psychophysiological measures

Measure	Advantages	Limitations
EEG/ERP	High temporal resolution	Low spatial resolution
	Relatively inexpensive (cf. fMRI)	Might be uncomfortable
	Relatively unobtrusive (cf. fMRI)	Time consuming
	Portable	Potentially high signal-to-noise ratio (signal has to be cleaned from noise, such as eye-movements or line noise)
Pupillometry	Inexpensive	Dependent and influenced by ambient light
	Relatively unobtrusive	
ECG	Inexpensive	Body movements influences the data
	Reliable	HR and HRV are influenced by contextual factors (respiration, posture, engagement, motivation)
	Easy to record	
	Good signal-to-noise ratio	
	Applicable in field studies	

EDA	Inexpensive	Low temporal resolution
	Easy to record	Sensitive to physiological reactivity, respiration, mental effort, etc.

Suggested Reading

Psychophysiology in driving context:

- Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. *Frontiers in Human Neuroscience*, *13*(March), 1–27. https://doi.org/10.3389/fnhum.2019.00057

EEG/ERPs:

- Malik, A. S., & Amin, H. U. (2017). *Designing EEG Experiments for Studying the Brain: Design Code and Example Datasets* (1st ed.). Academic Press.
- Luck, S. J., & Kappenman, E. S. (2011). *The Oxford Handbook of Event-Related Potential Components (Oxford Library of Psychology)* (Illustrated ed.). Oxford University Press.

Pupillometry:

- Sirois, S., & Brisson, J. (2014). Pupillometry. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(6), 679–692. https://doi.org/10.1002/wcs.1323

ECG:

- Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*(September), 1–17. https://doi.org/10.3389/fpubh.2017.00258

EDA:

- Boucsein, W. (2011). *Electrodermal Activity* (2nd ed. 2012 ed.). Springer.

EMG:

- De Luca, C. (2006). Electromyography. *Encyclopedia of medical devices and instrumentation*.

2.4 Further Tools

In the following section, a list of different empirical strategies is presented, along with their potential and where they fall short. Depending on the empirical approach and the research question in mind, additional techniques can be applied to further enhance the validity of the

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

study, but more than that, such techniques extend the constraints of the research paradigm. This becomes highly valuable as some studies are solely dependent on such technologies.

2.4.1 Wizard of Oz Paradigm

Since most of technologies that AVs use in communicating with other road users are not commercially available, it is hard for participants or even researchers to have access to such technology. To solve this problem, Wizard of Oz (WoOZ) experiments are developed (Kelley, 1984). With the ability to simulate an under-developed system, WoOZ has been widely adopted to address various subjects (Benzmüller et al., 2003; Kelley, 1984; Andemach, Deville, & Mortier, 1993). In the field of AVs technologies, in order to collect user data from a hypothetical model or developing AV system, WoOZ experiments are frequently used (Fuest, Michalowski, Träris, Bellem, & Bengler, 2018; Jarosch, Paradies, Feiner, & Bengler, 2019; Müller, Weinbeer, & Bengler, 2019).

2.4.1.1 Vehicle Setups

Typical WoOZ experiments consist of a participant, a driving wizard and an interaction wizard, where the driving wizard is hidden from the participant and the interaction wizard acts as an investigator (Bengler, Omozik, & Müller, 2020). As interactions between humans and different levels of automation are designed and investigated, the WoOZ setups would be adjusted accordingly. Taking interactions inside vehicles for instance, studies considering SAE Level 2-4 automations would require the participant sitting in the front row of the vehicle, since manual control and take-over-requests are usually required (Wang, Sibi, Mok, & Ju, 2017). While for SAE Level 5 automations where driver is out-of-the-loop, the setup with passengers in the backseat is possible (Sherry, Beckwith, Esme, & Tanriover, 2018; Karjanto et al., 2018).

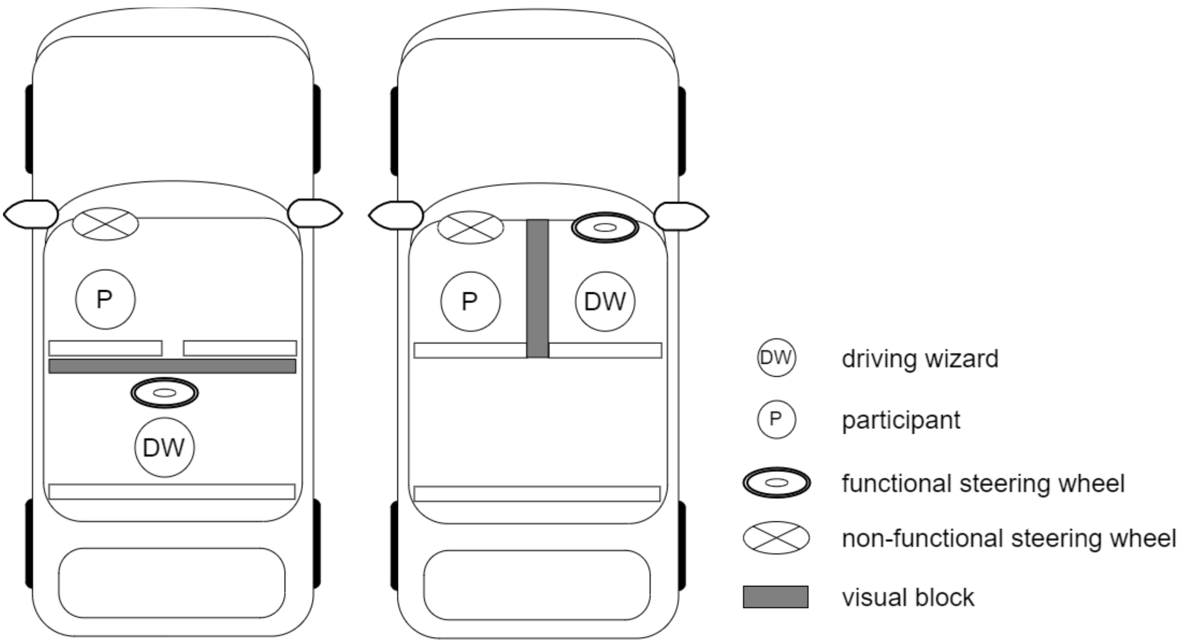

Figure 11 Wizard of Oz vehicle setup, available at Technical University Munich (Fuest et al., 2018)

2.4.1.2 Interaction Scenarios

As most of the automated vehicles available to the market fall within SAE Level 2-3, studies are focusing on in-vehicle interactions, where the role of the participant is a human driver inside an automated vehicle. During these experiments, the driving wizard would be controlling the vehicle during the autonomous mode, and requests for intervention when the system limit is reached. On the other hand, interactions outside the vehicle could still be investigated using WoOZ paradigm. To investigate the interaction between pedestrian and automated vehicle in real world, WoOZ paradigm is widely adapted as the risk for participants is minimized (Palmeiro et al., 2018; Habibovic, Andersson, Nilsson, Lundgren, & Nilsson, 2016). Road crossing is one of the major scenarios, where reactions of participants seeing an "automated vehicle" when crossing are evaluated. During the experiments, the driving wizard acts like a passenger in the passenger seat, while the driver seat is either empty or occupied with a driver doing non-driving-related-tasks (e.g., reading newspaper, using smartphone, watching video etc.). Decisions of participants on whether to cross the road would then be analysed, and the overall performance of the automation during the experiment would be evaluated (e.g., trust, acceptance, HMI design etc.).

2.4.1.3 Limitations

Among all AV experiment methods, WoOZ is by far the optimized and most realistic solution for real interactions data collection between human and AV. However, beside the complexity in experimental setups, critics and questions about the validity of such methods as well as whether solutions developed using WoOZ paradigms could be realized still exist. Furthermore, the problem of delayed response exists as the reaction time of wizard operator is not comparable to the real automated system (Wang et al., 2017), making the lagging response noticeable to participants in studies. The driving condition of wizard operators also impacts their performance, resulting in potential inconsistent driving experiences across the experiment (Pai et al., 2020). Nevertheless, these drawbacks could be mitigated with experimental designs and further researches regarding this topic

Suggested reading

First WoOZ experiment in natural language processing:

- Kelley, J. F. (1984, January). An iterative design methodology for user-friendly natural language office information applications. ACM Trans. Inf. Syst.,2(1),26–41. Retrieved from https://doi.org/10.1145/357417.357420

Analysis of components of WoOZ paradigm for automated vehicle and its relations to psychological testing:

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

- Müller, A. I., Weinbeer, V., & Bengler, K. (2019). Using the wizard of oz paradigm to prototype automated vehicles: Methodological challenges. New York, NY, USA: Association for Computing Machine

Evaluation and systematization of published experimental approaches, proposition of a specification language for the driving wizard's behaviour:

- Bengler, K., Omozik, K., & Müller, A. I. (2020). The Renaissance of Wizard of Oz (WoOz): Using the WoOz methodology to prototype automated vehicles. Proceedings of the Human Factors and Ergonomics Society Europe, 63-72

2.4.2 Augmented and Virtual Reality

Virtual reality simulation has been used in the automotive domain for experiments relating to both the driver and pedestrian side. Examples of driving simulators are discussed elsewhere in the document.

Pedestrian simulators have been developed to test out eHMIs. It is still however unclear whether these methods are as valid as naturalistic testing. Examples of pedestrian simulators include using screen-based setups (Ackermann, Beggiato, Schubert, & Krems, 2019; Chang, Toda, Igarashi, Miyata, & Kobayashi, 2018; Schwebel, Gaines, & Severson, 2008), mixed-reality setups (Maruhn, Dietrich, Prasch, & Schneider, 2020), Cave Automatic Virtual Environment (CAVE) simulation (e.g., Agarwal, 2019; Kaleefathullah et al., in press; Mallaro, Rahimian, O'Neal, Plumert, & Kearney, 2017) and head-mounted displays (Bazilinskyy, Kooijman, Dodou, & De Winter, 2020; Böckle, Brenden, Klingegård, Habibovic, & Bout, 2017; Deb, Carruth, Fuad, Stanley, & Frey, 2020; Deb, Carruth, Sween, Strawderman, & Garrison, 2017; De Clercq, Dietrich, Núñez Velasco, De Winter, & Happee, 2019; Hudson, Deb, Carruth, McGinley, & Frey, 2018; Otherson, Conti-Kufner, Dietrich, Maruhn, & Bengler, 2018; Schneider, & Bengler, 2020). See Feldstein, Lehsing, Dietrich, & Bengler, 2018 and Schneider, & Bengler, 2020, for more examples.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Figure 12 Coupled VR simulator, available at the Delft University of Technology (Bazilinskyy et al., 2020)

Augmented reality (AR) has the potential of providing a more realistic simulation environment by superimposing visual elements that pose the highest risk to participants (e.g., vehicles) over a real background (Tabone, De Winter, Ackermann, Bargman, Baumann, Deb, Emmenegger, Habibovic, Hagenzieker, Hancock, Happee, Krems, Lee, Martens, Merat, Norman, Sheridan, & Stanton, 2020). Recent work (Maruhn, Dietrich, Prasch, & Schneider, 2020) has in fact proposed a method which does exactly that. In this case, the participants experience virtual vehicles that are augmented on the actual streetscape.

2.4.2.1 Use Cases of AR

The use of AR technology in traffic research is diversified. There have been applications of AR inside the vehicle to allow for enhanced navigation (Rusch, Schall Jr, Lee, Dawson, & Rizzo, 2014) for example. See Riegler, Riener, & Holzmann, 2020 for a number of use cases of AR inside the vehicle. Other applications have included the use of AR for pedestrian navigation (Hesenius, Börsting, Meyer, & Gruhn, 2018; Montuwy, Cahour, & Dommes, 2018), city guides (Lakehal, Lepreux, Efstratiou, Christophe, & Nicolaou, 2020) and for crossing advise in urban traffic areas (Perez, Hasan, Shen, & Yang, 2018).

AR in the automotive user-interface domain is still a very active topic of research with a number of open challenges (Riegler, Riener, & Holzmann, 2020) that could likely be solved using such technology.

2.4.2.2 Opportunity for AR in SHAPE-IT

AR provides a potential solution to various problems that current eHMIs face. These include situations where multiple VRUs are encountered or multiple vehicles are communicating to one or more VRUs. Such a situation gives rise to ambiguity which may lead to a VRU inferring that a particular AV is signalling them to cross when in fact it is signalling to another road user it has detected across the road. Therefore, the first VRU might make an erroneous crossing decision, leading to an accident if the AV drives on. Another issue is that there is no standard

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

design for these eHMIs and so VRUs need to learn and infer each design encountered. Furthermore, in cases where text is used as a communication modality, the issue of readability and language barriers arises.

With AR, there is the potential to remove the eHMI from the exterior of the vehicle and place it on the VRU. This can be achieved through the use of Augmented Reality (AR) glasses which are able to place a virtual layer of information over the real environment. These devices are touted to be the replacement of the smartphone as a daily ubiquitous assistive technology and so their potential use in the future of urban road infrastructure should be explored. In this scenario, the wearer of the AR glasses would receive the message from an oncoming vehicle as a heads-up message on the device. Standardizations of the eHMI would be easier for manufacturers since software can be designed much quicker than any other hardware attachment to the vehicle. Moreover, the eHMI design can be updated much more easily as standards change. Should personalization and customization be the preferred route, each user of the AR glasses would be able to customize the eHMI message depending to their preference. Therefore, if for example, a French speaker prefers text messages, they would be able to receive a communication from an AV in French while crossing in Tokyo. This solution breaks down the language barrier and allows for better understandability according to preferred modality (some users may prefer visuals over text, etc.). Moreover, since each VRU would be receiving the communication from the AV individually, the multi-actor problem could potentially be solved. In this case, only the VRUs involved in the interaction with the AV would be signalled and so ambiguity would be lessened.

2.4.2.3 Limitations

Results has so far demonstrated the applicability of the presented setup, however, the authors cautioned that reported deviation from behaviour in the “real” condition should be interpreted with caution. There are also multiple technical challenges which still need to be overcome. These include issues of occlusion, reflections and shadows. Moreover, technological improvement in VR glass display technology in terms of resolution and cameras is needed. Although there is still more work to be done in this area in order, the authors concluded that results still demonstrated that AR is a promising tool to investigate research questions concerning pedestrian behaviour in a safe, controlled and realistic environment.

Suggested reading

- Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. C. F. (2020). Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. *Driving Simulation Conference Europe.* Antibes, France.
- De Clercq, K., Dietrich, A., Núñez Velasco, J. P., de Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human factors*, *61*(8), 1353-1370.
- Deb, S., Carruth, D. W., Sween, R., Strawderman, L., & Garrison, T. M. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied ergonomics*, *65*, 449-460.

- Feldstein, I. T., Lehsing, C., Dietrich, A., & Bengler, K. (2018). Pedestrian simulators for traffic research: state of the art and future of a motion lab. *International Journal of Human Factors Modelling and Simulation*, *6*, 250–265. doi:10.1504/ijhfms.2018.096128
- Hesenius, M., Börsting, I., Meyer, O., & Gruhn, V. (2018, September). Don't panic! guiding pedestrians in autonomous traffic with augmented reality. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 261-268).
- Perez, D., Hasan, M., Shen, Y., & Yang, H. (2019). Ar-ped: A framework of augmented reality enabled pedestrian-in-the-loop simulation. *Simulation Modelling Practice and Theory*, *94*, 237-249.
- Tabone, W., De Winter, J. C. F., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock. P. A., Happee, R., Krems, J., Lee, J. D., Martens, M., Merat, N., Norman, D. A., Sheridan, T. B.,& Stanton, N. A. (2020). *Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. Transportation Research Interdisciplinary Perspectives, 9, 100293.*

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

3 Models of Human/Agent Interaction

A model can be defined as a 'representation of construction and the working of some system of interest' (Maria, 1997). While the word 'modelling' refers to the process of developing a model, a model is in many ways similar to but a simpler form of a system that it represents. Many real-world problems can be defined as a dynamic interaction between competitive agents where agents' physical (i.e., position, orientation, and velocity) and internal states (i.e., navigational goals, behavioural traits, and 'mental model' of the surrounding environment) can be accommodated into a mathematical framework (Brown, D. et al., 2020). Road users and especially AVs' tasks such as state, intention, risk and trait estimation and also motion prediction and behaviour imitation could be addressed within this framework (Brown, D. et al., 2020). This framework can construe observations as a prediction of future events and act as foundations for ideas and help us to come up with questions regarding the observed behaviours (Calder et al., 2018). By defining and employing an appropriate model and considering important parameters in the related traffic scenarios, we could see what would happen in interaction scenarios considering agents' trajectories, decision time (response and reaction time) and social preference and therefore would be able to opt for better algorithm designs in this field as, for example, the research shows that heterogeneity of the drivers, pedestrians, vehicles, and road environment has not been considered in the model development of the most past related studies clearly (Amado et al., 2020).

This section is divided into four parts: At first, models which assume that humans are 'optimal' or 'rational' and/or those that based on some predefined rules are discussed. In the second part the models that try to describe not only the data, but also a hypothetical underlying process (or mechanisms) generating that data are discussed. That is, models based on psychological mechanisms underlying human behaviour. The third part covers more recent modelling approaches that are based on artificial intelligence. The fourth and final part models that are mostly used in traffic simulation are covered. Although this classification may be reasonable, some models clearly fit on more than one of the modelling categories. For such models we have chosen the place where we think it fit the best, but acknowledge that they (almost) equally well fit into another of the categories.

3.1 Rule-Based Models

Rule based models are often considered as mathematical models that are data driven. These modelling methods are widely used for different purposes in transportation engineering and traffic safety to predict terms like decision-making (e.g., route-choice modelling), acceleration deceleration behaviour, gaze pattern, and other safety measures. The overall goal of these models is to convert actual observations into expectations of future events.. In practical terms, these models are able to predict an anticipated outcome from a given set of parameters (Calder et al., 2018). Over time, rule-based models have developed thanks to a greater availability of data and increased computation power. Different considerations affect our decision for choosing the appropriate type of modelling for our research. First is the intended output of the model, then the type, variables, and quantity of data available. Due to the complexity of human behaviour, different models have been developed to address this complexity by incorporating different parameters and terms that may describe different aspects of human behaviour. All these models have limitations and advantages that are introduced in the last section of this chapter. In the following subsections, the most important modelling strategies in the section of rule based behavioural modelling are introduced in brief. Note that some of these model types (and examples of them) are used as model components in [models of human decision-making & behaviour] and other are mostly used in traffic simulations. These models could also have been placed in those section (section 3.2, and section 3.3), but are here provided as examples of rule-based models.

3.1.1 Types of models

Rule based models are often considered as mathematical models that are data driven. These modelling methods are widely used for different purposes in transportation engineering and traffic safety to predict terms like decision-making (route-choice modelling), acceleration deceleration behaviour, gaze pattern, and other safety measures. The objective of these models is to translate real observations into anticipation of future events. In practical terms, these models are able to calculate an anticipated result from a given set of variables (Calder et al., 2018). Over time, rule-based models have developed thanks to a greater availability of data and increased computation power. Different considerations affect our decision for choosing the appropriate type of modelling for our research. First is the intended output of the model, then the type, variables, and quantity of data available. Due to the complexity of human behaviour, different models have been developed to address this complexity by incorporating different parameters and terms that may describe different aspects of human behaviour. All these models have limitations and advantages that are introduced in the last section of this chapter. In the following subsections, the most important modelling strategies in the section of rule based behavioural modelling are introduced in brief.

3.1.1.1 Social Force Model

This type of modelling was originally developed by Dirk Helbing (1995), and is mainly used for pedestrian motion modelling (Helbing & Molnár, 1995). Since its early development, many researchers have tried to develop this modelling approach further and have added more terms to increase its efficiency. Social force models can describe the dynamic movement of agents (e.g., bicycles or pedestrians) in space or time (Huang et al., 2017). Among the dynamic pedestrian flow models, the social force models have been widely used in the areas like transportation station management, building evacuation, and safety management (Chen et al., 2018). Social force models use differential equations to explain the continuous movement of the agents in space. These equations and their related parameters can be described from the kinematics's perspective (positions and speed).

The concept of this modelling strategy is that agents (e.g., pedestrians) are limited by social forces and internal motivations. These forces include the driving forces for reaching the target, repulsive forces from other parties (e.g., other pedestrians), obstacles, boundaries, and attraction forces of companions and stores (Chen et al., 2018). Personal motivations include speed and acceleration, and on the other hand they are restricted by their situation and the environment that they are moving through. A pedestrian simulation software has been developed with this modelling approach for assessing pedestrian motion behaviour at different scenarios (Chen et al., 2018). The improved versions of social force models have a powerful capability for describing pedestrian movement and imitating self-organizing events. While there are some challenges with this modeling approach for validation application of the social force model in other motion-based scenarios like cycling (Chen et al., 2018).

3.1.1.2 Cellular Automata Model

Cellular automata models were developed by Nagel and Schreckenberg (1992) to simulate traffic flow (Nagel & Schreckenberg, 1992). Cellular automata models have the ability to use micro level motions of the traffic agents and relate them to macro level performance of the traffic network. This is in contrast with other modeling approaches which are either very detailed in their scope (microscopic models) or general in their application (macroscopic) for analyzing the traffic network. Cellular automata models have the capability of showing each vehicle interaction and connecting these interactions with traffic flow performance measures like speed and travel time (Saifallah Benjaafar, 1997).

In this modelling strategy, the road network is divided into small cells which are defined to incorporate cars and other road users. In each time step, each cell has two states, they are either occupied or not, depending on the presence of a vehicle in each cell. The current state of the cars is determined over time to reach their destination. In each time step, the current state of the cells will be updated with their attributed variables (e.g., lateral distance and forward gap). The updating rules are applied at the same time to all the involved vehicles at each iteration (Saifallah Benjaafar, 1997). There are two types of cellular automata models: deterministic and stochastic; the stochastic approach considers the inherent randomness in the vehicles' behavior in real traffic. In the deterministic approach of cellular automata, all the vehicles behave in the same way, and they may have the same speed limit which would be a quite simplified version of real traffic. Despite this deficiency, they are still a valuable tool for analyzing a fully automated traffic network where the speeds and accelerations of all the agents are predefined and controlled. In this modelling approach we can assign characteristics to each traffic agent. Doing so, they can imitate traffic flow more realistically (Lárraga et al., 2005). Since the work by Nagel and Schreckenberg (1992), several researchers have developed more advanced versions of this modelling approach, to simulate traffic flow more realistically (Belitsky et al., 2001).

3.1.1.3 Artificial Neural Network Models

Artificial neural network models are supervised learning systems that use mathematical models in their inner layers, and to some extent try to simulate the way that the human brain normally processes infromation. Information explaining a specific situation is taken as the input of this methodology, and a decision or answer is the outcome (Aghabayk et al., 2015). Among many subject areas in the field of transport, the highest proportion of using neural network models is in the context of driver behavior and autonomous vehicles (Dougherty, 1995). These models learn human behaviours from training data and are, in a perfect world, capable of extrapolating those human behaviours in a new situation. Artificial neural networks have been used in a number of driver-behaviour modelling studies, including modelling lane changing (Rahman et al., 2013) and car following (Khodayari et al., 2012).

Despite the wide application of this approach in different disciplines, there are two main disadvantages to this methodology, one is that Neural network models are totally data dependent and they need supervised training. The other drawback is that they require field collected traffic data for their calculation, although results in previous studies indicate that this approach can accurately predict driver's behavior like lane changing behavior (Huang, 2014). Another issue with this methodology is that it is a "black box" method for modeling, where humans generally are not able to discern the inner working procedure of neural network, to understand what different parts do. In other words, neural network models can be viewed by the users only by the input and output of the model without giving any explanations about the internal working procedure (Aghabayk et al., 2015). (A broad description of AI based modelling is provided in chapter 4 of this section.)

3.1.1.4 Fuzzy Logic Models

In general, a fuzzy logic system is a nonlinear mapping of an input data (features) vector into a scalar output (Mendel, 1995). The fuzzy logic theory is developed based on the idea that human thinking is not happening based on the numbers but rather based on the labels fuzzy sets (Kalinic & Krisp, 2019). The main idea behind fuzzy logic is that variables can be represented by fuzzy sets and consequently we avoid rigid binary values. As a result, Fuzzy logic approach gives us this opportunity to deal with vague and imprecise concepts (Kalinic & Krisp, 2019). One of the main characteristics of fuzzy logic models is that they take into account the actual and natural perception of variables and consider the uncertainty of behavioral events (Mendel, 1995). They translate nonlinear systems into IF-THEN rules. There are two main terms in fuzzy logic modelling approach, the first is the membership functions and the second is the fuzzy inference process (for more information related to these terms see this paper (Kalinic & Krisp, 2019). There are different types of membership functions including triangular, gaussian, sigmoidal and polynomial. One essential step in the fuzzy logic process is to establish a mechanism for defining how to relate the input data to output results. This will be answered by setting if-then rules.

In the fuzzy logic rules both objective and subjective aspects of an issue can be utilized to solve a real-world problem. It is worth noting that there is a traffic flow simulation software (FLOWSIM) that is developed mainly by using fuzzy logic rules (Kalinic & Krisp, 2019; Mendel, 1995). On the other hand, one must consider that there are some difficulties in determining fuzzy rules. One of the challenging parts of applying fuzzy logic models is defining input parameters since it needs profound knowledge and experience in the field of study. If the drivers' perceptions are not defined properly in the model, the model's output will be unrealistic and the prediciton will be inaccurate. Two major challenges in employing fuzzy logic rules are the validation of membership functions and determining fuzzy rules (Aghabayk et al., 2015).

3.1.1.5 Discrete Choice Models

This modelling approach explains and predicts the choice between two or more discrete options. This methodology, theoretically or empirically predicts choices made by people among a finite set of options. Discrete choice models have been widely used in marketing, transportation, economy, and other areas to study both revealed and stated preference data (Keane, 1997). Discrete choice models are widely used by Transport planners to estimate traffic demand. As a famous example of using discrete choice modelling, Ahmad et al (1999) categorized lane change decision making of drivers as mandatory, discretionary and forced merging (Casello & Usyukov, 2014; Moridpour et al., 2010). He proposed a model which predicts the probability of performing mandatory, discretionary, and forced lane changing at a given time. Another example of using discrete choice modelling in transportation engineering is the route choice behaviour of people when they are choosing between a limited number of alternatives for their trip including cycling, walking, and bus (Prato, 2009). There are different types of discrete choice models including binary logit, binary probit, multinomial logit, multinomial probit, nested logit, mixed logit (Nagel & Schreckenberg, 1992).

The most common type of discrete choice modelling has been multinomial logit (MNL) which is widely used for travel behaviour analysis. The attractiveness of MNL is coming from the fact that the probabilities are easy to estimate in this approach. The problem with this modelling approach is that they are making strong assumptions about an individual's behaviour. The most common issue about these assumptions is with the independence of irrelevant alternatives property, which says that if we introduce a new alternative to a set of choice, then the choice probabilities fall proportionately for the all the existing alternatives (Keane, 1997). As an example of using logit model in the context of human VRU interaction, Silvano et al (2016) investigated the probability of yielding or not when a bicycle is interacting with a vehicle at a roundabout (Silvano et al., 2016).

3.1.1.6 Game Theoretic Models

Game theoretic models were first introduced and formulated by John F. Nash (1950). There are four basic elements in game theory models: 1) player or participants of the game: the ones who decide about their strategy, 2) a set of strategies: the strategies that are available for the players to play 3) payoff functions: after players decide on which strategy to do, there is a payoff or result for their action, which shows gain or loss; and 4) order of playing: when players want to decide their strategies, there is a need to decide the orders. Sometimes the order of playing games happens at the same time and sometimes they make decision one after another.

Six common types of game theoretic models have been used frequently in transportation engineering including ordinary non-cooperative game, generalized Nash equilibrium game, Cournot game, Stackelberg game, bounded rationality game and Repeated game. Examples of using this modelling approach are, modelling lane changing using game theory (Talebpour et al., 2015) and interaction modelling between pedestrians and vehicles (Camara & Fox, 2020). Game theory's application is at situations when agents or parties should make rational decisions about their actions in relation to other parties (Zhang et al., 2010).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

3.1.2 Advantages and Limitations

Table 5 presents the most notable advantages and limitations of the above-mentioned modelling approaches. For better understanding of the limitations, please see the suggested reading section with recommended references.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table 5 Advantages and limitations of rule-based models

Model type	Advantages	Limitations
Cellular automata model	Simplicity in modelling, small number of variables	Assuming all agents behave in a same way, difficulty in parameter calibration
	High capabilities in traffic simulation in multi lane highways. Other behavioural models like lane changing can be easily accounted for in the modelling framework	
Game theory	Simplicity in modelling, small number of variables	Doesn't consider driver variability in the model, binary answers, difficulty in parameter calibration
	useful method in cases that the firms are independent	In case of mixed strategies, the method of solving games is very comlicated
	Providing a systematic quantitative approach for deciding the best strategy in competitive situations	Determining pay-off functions and values of players of the game is difficult
		Assuming all players are rational
Artificial neural network models	Attempts to capture driver's variability by training data	Completely data driven and require supervised training
		Require large amount of data
Discrete choice models	Decide based on maximum gained utility	Require calculating the probability functions to determine the utility of each choice
	Probabilistic results instead of binary answers	Making strong assumptions in their theory
	Easy to implement and interpret	
Fuzzy logic models	A powerful solution for solving complex problems in all fields of life, as it resembles human reasoning and decision making.	Time-consuming to define and develop fuzzy rules and membership functions. It's hard to interpret the outputs of the modelling
	High precision	For more complex scenarios it requires to define more fuzzy grades which result to increase exponentially the rule
	Based on linguistic model	Restricted number of usages of input variables
	Using simple mathematics for nonlinear, integrated and complex systems	

Suggested reading:

- Aghabayk, K., Sarvi, M., & Young, W. (2015). A State-of-the-Art Review of Car-Following Models with Particular Considerations of Heavy Vehicles. *Transport Reviews*, *35*(1), 82–105.
- Calder, M., Craig, C., Culley, D., de Cani, R., Donnelly, C. A., Douglas, R., Edmonds, B., Gascoigne, J., Gilbert, N., Hargrove, C., Hinds, D., Lane, D. C., Mitchell, D., Pavey, G., Robertson, D., Rosewell, B., Sherwin, S., Walport, M., & Wilson, A. (2018). Computational modelling for decision-making: Where, why, what, who and how. *Royal Society Open Science*, *5*(6).
- Chen, X., Treiber, M., Kanagaraj, V., & Li, H. (2018). Social force models for pedestrian traffic–state of the art. *Transport Reviews*.
- Zhang, H., Su, Y., Peng, L., & Yao, D. (2010). A review of game theory applications in transportation analysis. *Proceedings of ICCIA 2010.*

3.2 Models based on psychological mechanisms underlying human behaviour

Here, two behavioural modelling approaches are discussed briefly: Evidence Accumulation Models (EAMs) and Behavioural Game Theory (BGT).

EAMs suggest that evidence for a particular response is integrated by single or multiple accumulators over time and by some rate known as ‘drift rate’: The rate at which sensory information reaches a boundary (a decision boundary) is determined by the quality of evidence extracted from the stimulus or memory (Ratcliff, Roger et al., 2016). The accumulation process is noisy which means in each time step, the evidence may direct attention to one or the other of the two boundaries (or two poles of one boundary), but more frequently to the correct than the incorrect one. The boundary defines the amount of evidence that should be accumulated before a response is made (Ratcliff, Roger et al., 2016). A response is chosen when the evidence for one alternative reaches some level of evidence that triggers a decision. These models usually have two major applications – the first is to accumulate informative evidence for/against every competing hypothesis and second, to accumulate affective assessments for/against each of several courses of action (Busemeyer et al., 2019). Overall, EAMs have been used so far in simple driving tasks like response time and psychomotor vigilance test (Ratcliff, Roger et al., 2014), brake response (Svärd et al., 2020), steering control scenarios (Markkula et al., 2018a), collision threat detection task and time to collision estimation

(Daneshi et al., 2020), driver gap acceptance in turns (Zgonnikov et al., 2020), pedestrian crossing decision (Sargoni and Manley, 2020), detection response task (Howard et al., 2020), AV-human interactions in take over and crossing scenarios (Markkula et al., 2018b) and modelling cognitive load in driver distraction context (Castro et al., 2019).

BGT employs experimental evidence to make computational models of human cognitive limitations, social utility and learning rules aware of 'how people actually behave in strategic situations' (Camerer, 2003). One of the most important components of this model is the theory of how people make a choice in one-shot games or in the first round of a repeated game and this is where related studies indicate that Nash equilibrium in conventional game (GT) theory (i.e., a set of decision strategies indicating individuals cannot make their gain better by unilaterally changing the strategy in non-cooperative games) is often a poor description of human players' behaviour especially in unrepeated normal-form games (Wright and Leyton-Brown, 2017). In traffic interactions, for instance, crossing scenarios involving a pedestrian and an AV can be considered as non-cooperative simultaneous repeated games which can be solved using mixed-strategy algorithms. These models include but not limited to level-k reasoning (Stahl et al., 1995), cognitive hierarchy theory (Camerer et al., 2004), logit quantal response equilibrium (McKelvey et al., 1995), noisy introspection (Goeree et al., 2004) and the dual accumulator model (Golman et al., 2019). Concerning traffic interactions, a few studies have been employed the mentioned models such as Level-k reasoning (Albaba and Yildiz, 2020) and cognitive hierarchy reasoning (Li et al., 2019) and showed that they can simulate the traffic conditions well.

3.2.1 Advantages and Limitations

Most of the normative models that have been discussed in the previous section assume that road users act mostly like moving objects without considering each other's intentions before taking every decision which makes it hard for one to account for interdependencies and have a meaningful level of model order (i.e., the problem of infinite regress). Behavioural models can account for this shortcoming by considering each agent's intention and preference in each decision task. Moreover, unlike conventional GT, BGT suggests that people are not totally self-interested (i.e., preferences are highly context-dependent) and they do not make decisions based on absolute outcomes but according to a heuristic estimate of the potential value of losses and gains (Kahneman and Tversky, 1979).

While EAMs provide an ample level of detail of decision-making process, they can do that for a very constrained set of tasks and typically are considered as a single-decision models which suggests they may not be a good framework for all type of interaction scenarios. In addition, both BGT and EAM are in their infancy with regard to traffic interactions. especially considering vehicle-pedestrian scenarios; more work is needed to understand the potential of these models in this context.

Suggested Reading

- Bhatia, S. (2013). Associations and the accumulation of preference. *Psychological review*, *120*(3), 522.

- Bhatia, S. (2016). The dynamics of bidirectional thought. *Thinking & Reasoning*, *22*(4), 397-442.
- Camerer, C. F., & Fehr, E. (2006). When does" economic man" dominate social behavior?. *Science*, *311*(5757), 47-52.
- Colman, A. M., Pulford, B. D., & Lawrence, C. L. (2014). Explaining strategic coordination: Cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning. *Decision*, *1*(1), 35.
- Evans, N. J., & Wagenmakers, E. J. (2019). Evidence accumulation models: Current limitations and future directions.
- Purcell, B. A., & Palmeri, T. J. (2017). Relating accumulator model parameters and neural dynamics. *Journal of mathematical psychology*, *76*, 156-171.

3.3 Artificial Intelligence Based Models

Artificial Intelligence (AI) can be defined as intelligence representation by machines or systems. Machine Learning (ML) and Deep Learning (DL) are two most common tools used for AI. Unlike rule-based or statistic-based methods, AI-based methods for modelling rely more on large datasets, and where the model can learn directly from the data.

AI-based methods are applied in various field, including the classification, regression, 2-dimentional image segmentation, and object detection. It can be widely used in the field of intelligent vehicles. In the SHAPE-IT project, we are concerned about the interaction between humans and automated vehicles. AI-based modelling from this perspective mainly includes the following research questions:

RQ1: How AI based modelling can be used to investigate vehicle-VRU interaction? For this research question, we mainly focus on Artificial Intelligence (AI) and Human Factor. This research question corresponds to the VRU behaviour, and can be divided into several sub-questions, including: a) what is the state-of-the-art algorithm of pedestrian behaviour prediction? b) what is the influence of vehicles on pedestrian's low-level information, or pedestrian's trajectory? c) what is the influence (interaction) of vehicles on pedestrian's high-level information, or pedestrian's crossing intention? d) how do the pedestrians interact with automated vehicles, and how could we use this information to support human factor design?

RQ2: How can AI based modelling be used to investigate vehicle-vehicle interaction?

RQ3: How can AI based modelling be used to investigate vehicle-driver interaction?

RQ4: How are AI based modelling affected by Human factors, and how can it be used in automation design?

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

3.3.1 Available Tools

For AI-based modelling, there are some available tools and methods to assist us. Deep Learning (DL) is one of the most powerful tools. Because these tools can learn features and build models directly from dataset we feed in, they can be generally used to solve various tasks.

Convolutional Neural Networks are very powerful for extracting features from images directly. These networks can perform tracking (Ma et al., 2015; Wang et al., 2015), trajectories predictions (Casas et al., 2018) and intention classifications.

In the field of the pedestrian trajectory prediction, recent research on DL has showed the potential of learning features directly from data. The methods based on Recurrent Neural Networks and their improved version - Long short-term memory, are preferred by many researchers because of their strong ability to handle the trajectory sequence information (Alahi et al., 2016; Xue et al., 2018).

Generate Adversarial Networks are increasingly widely used for pedestrian trajectories and behaviour prediction recently (Gupta et al., 2018; Li et al., 2019). This kind of methods can not only provide a solution to regress the ground truth in training data but can also generate other possible solution which not in the training data. These networks can overcome the difficulties in approximating intractable probabilistic computation but may not be easy to get to converge during training.

However, when there are not enough data that can be used for DL training, the usage of DL can cause overfitting. In this situation, machine learning methods can be used, especially in intention classifications, such as Support Vector Machine and Random Forest. For regression tasks, the Linear Regression (LR) is also a useful tool to use.

3.3.2 Advantages and Limitations

Using AI-based modelling has some advantages compared to the use of rule-based methods. Generally speaking, the rule-based or statistic-based model usually requires the knowledge of the experts to design the features and models, and they are hard to generalized to another dataset. But with AI-based model, the deep learning structure can learn the features and model parameters directly from dataset, and can be easily generalized to other scenarios.

Besides, the hand-crafted models are difficult to be designed completed, and not easy to capture the complex feature of the data distribution, especially when the features of the data are non-linear. By contrast, deep learning networks can extract the complex feature of the datasets, and can deal with non-linear property better.

Take pedestrians behaviour prediction for example, the great challenge of pedestrians' intention prediction is that pedestrians can change their direction and velocity suddenly (Volz et al., 2015) and they tend to interact with the other road users and surroundings. Therefore, it is difficult to reliably predict the intentions of pedestrians by hand-crafted features. The AI based modelling methods can avoid this problem – the model can learn the features directly from data, and build the model implicitly with neural networks.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

There are also limitations of the AI based modelling of pedestrian behaviour. Because the neural networks usually model human behaviour as black box, it is hard to explain the models. Also, when the models do not work well, it is also not easy to modify and tune the AI model. Some other limitations are: a) AI based models usually need a large amount of labelled data, b)the annotation could be time consuming and expensive, and c) the model training time can be substantial for complicated AI models.

3.3.3 Limitations from the Perspective of Human Factors

AI components behave differently from conventional components. In particular, there is a certain level of uncertainty that depends on the accuracy of the AI model. Thus, for a system such as an AV, one cannot specify the exact behaviour (since in fact the behaviour should depend on partially unanticipated automatic decisions that the system makes in a given context). Thus, one must find ways to specify the corridor in which an AI-based system can make decisions. This depends on a good definition of the context in which the system operates and the quality or suitability of data in this context (See Section 4 on Requirements Engineering).

At the moment, several gaps in AI-based modelling research exist in the state of the art:

There is no established method to define a corridor of acceptable behaviour. Goal-models can work, but tend to fail scaling to the complexity of AVs. Scenario-based approaches are by example, and thus do not offer sufficient completeness for building safety cases. System-centric feature requirements do not scale or to reach a sufficient level of completeness (i.e., one would need millions of system requirements to describe simple goals).

There is no established method to define the context of operation. ODDs (operational design domains) are fashionable in the automotive field, but not yet integrated in development methods. In particular, they leave many open questions with respect to safety argumentation (Gyllenhammar et al., 2020).

There is no established framework to describe the quality of input or output data in a given context.

In addition, data-driven approaches and AI modelling do not sufficiently cover all aspects of human factors and should be complemented with problem-based requirements derived from human factors. Because people are not limited in their behaviour, they can be way more than in a set of data. Humans have intelligence which allow humans to act differently. In AI based automated vehicles, decision-based algorithms work in similar ways, they have capability to learn by themselves but they follow pre-defined behaviour and struggle to deal with completely new situations. Inability to adjust to new situations and inability to fully comprehend and learn human behaviour will surface as system failures or bugs in the real-world, if not mitigated through a systematic approach to understanding human factors and system requirements.

Suggested Reading

- Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social LSTM: Human trajectory prediction in crowded spaces. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2016.110
- Casas, S., Luo, W., & Urtasun, R. (2018). IntentNet: Learning to Predict Intention from Raw Sensor Data. CoRL.
- Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., & Alahi, A. (2018). Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2018.00240

3.4 Simulation of traffic interaction

The word ‘simulation’ describes an approximate imitation of a certain situation or process. With advancement in technology computer simulations become more and more popular in different scientific domains. Traffic is one domain that has extensively used computer simulations. This section provides a broad overview of different simulation and modelling approaches for both motorised and non-motorised traffic.

3.4.1 Motorised traffic simulation

Motorised traffic simulation models are broadly classified as macroscopic , microscopic and mesoscopic simulations. Macroscopic models describe traffic as a continuum flow and only considers aggregated variables such as average flow, average density and average speed (van Wageningen-Kessels et al., 2015). Microscopic models, on the other hand, describe how individual vehicles behave within the network, Mesoscopic models fall between macroscopic and microscopic models and simulate vehicle movements in aggerate terms such as probability distributions (van Wageningen-Kessels et al., 2015). Macroscopic and mesoscopic simulations are out of the scope of this document, as their application in SHAPE-IT seems not to be likely, so we will only discuss microscopic models of traffic simulation.

A microscopic simulation model is composed of physical components and associated behaviour models. The physical components include a road network, road users and traffic control systems, while driving behaviour and route choice models are part of behavioural models. Simulations of individual vehicle units in a microscopic model are based on car following (longitudinal movement), lane changing (lateral movement) and gap acceptance models.

Car following models describe the process through which individual drivers follows each other (Brakstone and McDonald, 1999). Car following models are further categorized mainly into three different categories; (a) safe-distance models (e.g., Pipes, 1966; Gipps, 1981), (b) stimulus-response models (Treiber et al., 2000), (c) psychophysical or action point models

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

(Wiedemann, 1974). Table 6 shows some of the most famous models in each category and their main characteristics.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Table 6 Car following models

Category	Characteristics	Famous models
Safe-distance models	Vehicles adjust speed and maintain a safe distance from their leader	Pipes, Gipps
Stimulus-Response models	Acceleration of drivers is reaction to (a) own current speed, (b) distance to leader and (c) relative speed to the leader	Gazis-Herman-Rothery model, Intelligent driver model, optimal velocity model
Psychophysical or action point models	Drivers react when the change is large enough to be perceived	Wiedemann car following models

The second most common behaviour observed in traffic after the car following is lane changing. Lane changing can be categorized as mandatory and discretionary. A mandatory lane change occurs when the driver wants to change lane in anticipation of a turn (left or right) or to avoid a lane closure in the downstream. While a discretionary lane occurs when the driver is not satisfied in driving within same lane (e.g., due to slow moving vehicle ahead) and desires a faster speed, large following distance etc. (Vechione et al., 2017). In a microscopic simulation, lane changing models describe the decision process of driver during a lane change in a given time. Rahman and Chowdhury (2013), broadly classifies lane changing models into; (a) changing models for adaptive cruise control, and (b) computer simulation. The lane changing models for simulation are further categorized as rule-based models, discrete-choice-based models, incentive-based models and artificial intelligence models (see Table 7).

Table 7 Lane-changing models for simulation

Category	Model	Reference(s)
Rule-based model	Gipps model	Gipps (1986)
	CORSIM model	Halati et al., (1997)
	ARTEMiS model	Hidas (2005)
	Cellular-Automata model	Rickert et al., (1996); Nagel et al., (1998)
	Game theory model	Kita (1999)
Discrete-choice-based models	Ahmed’s model	Ahmed (1999)
	Toledo et al’s model	Toledo et al. ()
Incentive-based models	MOBIL	Kesting et al. (2007)
	LMRS	Schakel et al. (2012)
Artificial intelligence-based models	Fuzzy-logic-based models	Das & Bowles (1999)
	Artificial neural network model	Dumbuya et al. (2009)

Another important component of a microscopic simulation is the gap acceptance model. Gap acceptance is used to determine the number vehicles that can pass through a conflicting point in a given time. Gap acceptance is an important parameter in simulation of vehicle movements at unsignalized intersections, roundabouts and lane changing on freeways (Ozbay et al., 2014). Gap acceptance models captures driver’s decision making and determines the size of a gap that a driver might accept or reject while aiming to merge or cross the intersection.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

3.4.2 Non-motorised traffic simulation

The non-motorised modes of transport consist of walking, cycling and other varieties of human-powered transportation (Biggar, 2019). However, the scope of this section is limited to study of the pedestrian movement (walking) using modelling approaches. Although the literature on pedestrian simulation is limited compared to motorised traffic, it is gaining a significant attention nowadays. For example, Kouskoulis & Antoniou (2017), presented a system review of pedestrian simulation models with a focus on emergency situations.

The pedestrian simulation models are categorized as macroscopic and microscopic. The macroscopic models treat the pedestrian crowds as fluid or continuum which respond to local influences (Xia et al., 2009; Guo et al., 2010). On the other hand, the microscopic models consider pedestrians as discrete individuals (Guo et al., 2010; Teknomo et al., 2016).

Guo et al. (2010) further categorizes the microscopic simulation models into a) continuous, b) discrete and c) semi-continuous. The continuous models include, social force and optimal control theory models. While discrete models include lattice gas and cellular automata models. As the name suggests, continuous models describe the continuous movement of pedestrians in time and space. While, space and time are discretized to resemble movement of pedestrians. However, in case of semi-continuous models the time is considered as discrete while the space as continuous. For a more detailed understanding please refer to Guo et al. (2010).

Pedestrian simulations have been used to study wide a variety of problems. Over the past few decades, pedestrian simulations were largely used to evaluate the effectiveness of proposed policies for improvement of the pedestrian facilities (Lovas, 1994; Teknomo et al., 2002). In addition to the simulation of normal pedestrian behaviour, simulations have also been applied in evacuation and panic research (e.g., Helbing et al., 2002).

3.4.3 Advantages and Limitations

Traffic simulations have been extensively used to test and evaluate a proposed strategies of traffic improvement before the implementation. One of the biggest advantages of using traffic simulation is the saving of both cost and time compared to field methods. Traffic simulations also provides opportunities to evaluate the effectiveness of number different strategies within a shortest possible time. Traffic simulation has been widely used to answer many questions regarding traffic signal optimization (e.g., Stevanovic et al., 2016), modelling lane changing and merging (e.g., Hidas, 2002), evaluation of advanced traffic management system. Traffic simulations have also been widely used to study the impact of connected and autonomous vehicles on traffic flow (e.g., Talebpour and Mahmassani, 2016; Lu et al., 2020). More recently, researchers integrated traffic simulations with driving simulators to produce more realistic traffic scenarios in the driving simulator (That & Casas, 2011; Jeihani et al., 2017).

Although traffic simulation has several advantages, it has also various limitations. Microscopic traffic simulation is considered as a cost effective and time saving approach to conduct detailed analysis of complex traffic situations. However, the validity of its results depends mainly on two things; (a) if the model is able to replicate the behaviour of road users observed in the real situations, and (b) how well it is calibrated to the real world. Calibration is the process of changing default parameters to reduce the error between simulated results and the actual results to a required threshold value (typically based on real-world data). The behaviour of pedestrians in comparison to the drivers is more complex and is easily affected by the surroundings (Guo et al., 2010). Hence, it is very difficult to validate and calibrate pedestrian simulation models.

3.4.4 Available Tools

Table 8 presents the most common software options available in the field of traffic modelling and simulation.

Table 8 Traffic modelling and simulation software

Name	Scope	Pedestrian Simulation	License type
PTV VISSIM	Microscopic	Yes	Commercial
AIMSUN	Hybrid	Yes	Commercial
CORSIM	Microscopic	Yes	Commercial
SUMO	Microscopic	Yes	Open-source
PARAMICS	Microscopic	Yes	Commercial
PTV VISUM	Macroscopic	-	Commercial
TRANSIMS	Microscopic	-	Open-source

Suggested Reading

Traffic simulation in general:

- Barceló, J. (2010). *Fundamentals of traffic simulation* (Vol. 145, p. 439). New York: Springer; Lieberman, E. B. (2014). Brief history of traffic simulation. *Traffic and Transportation Simulation*, *17*
- van Wageningen-Kessels, F., Hoogendoorn, S. P., Vuik, K., & van Lint, H. (2015). Traffic Flow Modeling: a Genealogy. *Transportation Research Circular. N E-C195. April*, 1.

Car Following Models:

- Brackstone, M., & McDonald, M. (1999). Car-following: a historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(4), 181-196.

Lane changing models:

- Moridpour, S., Sarvi, M., & Rose, G. (2010). Lane changing models: a critical review. *Transportation letters*, *2*(3), 157-173

- Rahman, M., Chowdhury, M., Xie, Y., & He, Y. (2013). Review of microscopic lane-changing models and future research opportunities. *IEEE transactions on intelligent transportation systems*, *14*(4), 1942-1956.

Gap acceptance models:

- Akçelik, R. (2007, December). A review of gap-acceptance capacity models. In The 29th Conference of Australian Institutes of Transport Research (CAITR 2007), University of South Australia, Adelaide, Australia (pp. 5-7).

Pedestrian Simulation:

- Lu, L., Ren, G., Wang, W., Chan, C. Y., & Wang, J. (2016). A cellular automaton simulation model for pedestrian and vehicle interaction behaviors at unsignalized mid-block crosswalks. *Accident Analysis & Prevention*, *95*, 425-437.
- Liu, M., Zeng, W., Chen, P., & Wu, X. (2017). A microscopic simulation model for pedestrian-pedestrian and pedestrian-vehicle interactions at crosswalks. *PLoS one*, *12*(7), e0180992.
- Suh, W., Henclewood, D., Greenwood, A., Guin, A., Guensler, R., Hunter, M. P., & Fujimoto, R. (2013). Modeling pedestrian crossing activities in an urban environment using microscopic traffic simulation. *Simulation*, *89*(2), 213-224.

Traffic simulation tools:

- Kotusevski, G., & Hawick, K. A. (2009). A review of traffic simulation software
- Jones, S. L., Sullivan, A. J., Cheekoti, N., Anderson, M. D., & Malave, D. (2004). Traffic simulation software comparison study. *UTCA report*, *2217*.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

4 Requirements Engineering

Automated systems are playing an important role in our daily life, with the advancement of technology and available functionality we are getting more dependent on these systems. To take the full advantage of these automated systems, human factor experts provide certain requirements which are sometimes not considered in full by developers. For example, human factors experts provide user experience vision and usability evaluation which helps to increase the system safety, trust and acceptance.

The companies try to integrate human factors, but it is not clear that how to include and communicate this knowledge (Dul et al., 2012) to the developers, particularly in large scale agile development. In large scale development, human factors knowledge is often neglected because of communication challenges (Dresner, 2015), difficulties of including user experience (Larusdottir et al., 2017), iterative development and due to fast delivery and short release time.

These factors lead us to find a solution to integrate human factors requirements within the state-of-the-art requirements engineering process, aiming to remove the communication gap between developers and human factors experts.

Requirements Engineering (RE) is a systematic approach to reduce the likelihood of a development of the wrong solution (i.e., one that does not solve the problem). RE does not only answer questions about what system to build and how to engineer it, but also whether we need to engineer this system why and to what extent. Thus, RE must not only obtain a technical perspective but inherently a social perspective as well. While not directly a modelling approach, the sum of all requirements can be seen as a conceptual model of the conditions and capabilities that a system must possess and for this reason, we list it here in this report.

4.1 The Importance of Requirements

Within system and software engineering, RE is generally considered to be a key factor of success. Despite of its significance, this discipline is not well understood across multidisciplinary projects. Particularly when we talk about non-functional or quality requirements because different professionals may have discrete perspectives. Depending on the contextual factors (Schneider et al., 2018), this accounts for both academia and industry (Vogelsang et al., 2020). The main purpose of requirement engineering is to make sure that the system under development meets customers' expectations with minimum cost and time.

Studies show that poor requirement management can be the single largest cause of many software failures. From all around the world there are many real-world examples of software failures, which have catastrophic results including loss of both financial and human lives due to improper requirement engineering. Further studies show that cost of error correction increases with the development phases of SDLC (software development life cycle).

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

As technology is continuously changing and as the complexity of systems is increasing, approaches to requirements engineering also need to evolve. Traditionally companies were using plan-driven approaches for requirement engineering, where requirements are analysed and specified in a specific phase. Recently, many companies transition to value-driven (or: agile) approaches, where systems are built iteratively. This allows these companies to validate that a system indeed solves its purpose in small increments and to provide value early and, through regular updates, consistently, to the customers. While such approaches promise companies to be more responsive to changing requirements and to achieve fast delivery of new functions, it requires to rethink how requirements are engineered in parallel to system development. This is especially true for complex systems such as automated vehicles.

4.2 Types of Requirements

According to IEEE, a requirement is a condition or capability needed by a user to solve a problem or to achieve an objective (IEEE, 1990). Glinz (2007), suggests to distinguish different types of requirements, most importantly functional and non-functional requirements (see Figure 13; note that the term non-functional requirements are discouraged but widely used to describe [quality] attributes and constraints).

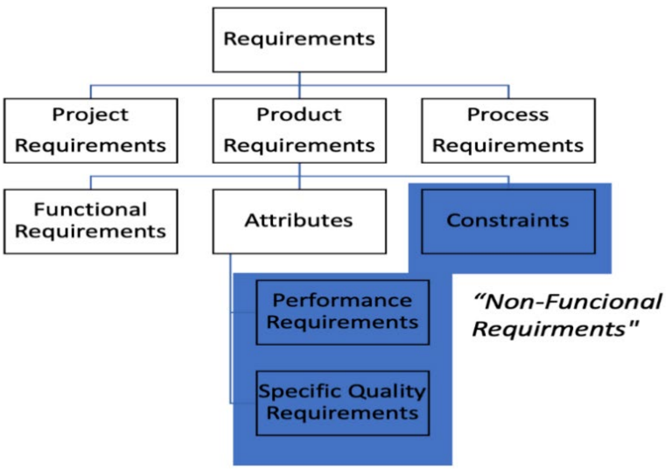

Figure 13 Requirements taxonomy according to Glinz (2007)

Functional requirements

Functional requirements include all requirements that a system must possess to enable the user to achieve their tasks. We can also say that functional requirements are all the functionalities of the system that users use to accomplish their work. For example, vehicle should provide an alert to driver on detection of an object through rear camera sensors.

Non-functional requirements

Following the taxonomy proposed by Glinz (2007), non-functional requirements are attributes and constraints of the system. System attributes includes performance requirements and other specific quality requirements as shown in Figure 13. For example, a non-functional requirement could specify how quickly the system should provide alert to driver, when a threat is detected by camera sensors.

4.3 Requirements Engineering Activities

Requirement engineering process includes following activities (Leffingwell & Widrig, 2003)

Requirement Elicitation

During Requirements Elicitation, requirement engineers discover and gather stakeholders (customer, user and others) requirements (needs and wants) using different elicitation techniques (such as interviewees, prototyping, storyboarding etc). While this is typically the first activity to perform in requirements engineering, it is usually necessary to return to this activity periodically.

Requirement Analysis

In this process, the elicited requirements are analysed. The activity involves reviewing the requirements to remove ambiguities, conflicts and inconsistences. It is examined if there is any missing or extra requirement. The related requirements are grouped together and processed to different categories for better clarity and understanding.

Requirement Specification

This is a process where requirements are captured in one or more ways including natural language, formal modelling approaches or mathematical expressions. The stakeholders' requirements are specified in a document or graphical model. Requirements that are specified should be complete, clear, correct and consistent.

Requirement Validation

The process of confirming that the specified requirements are according to customers and user's needs. We ask the customers and users to check and confirm if the requirements are correct and complete.

Requirement Management

Requirement management is a process to manage all the activities of RE. It includes documenting, analysing, prioritizing, tracing, tracking, controlling and reviewing requirement changes.

4.4 Customer (Problem-based) and Supplier (System-based) Requirements

Problem-based requirements are based on the customer/user requirements. It includes all the activities that user must be able to accomplish with the system (Wiegers, 2003) and describe requirements from the perspective of the users and the problems that the system should address. These requirements are used as input to the system requirements. Gathering customer requirements is a challenging task, partially, because users have difficulties to articulate what they actually want. Moreover, information they provide may be incomplete or conflicting. In addition, engineers tend to think too early about the possible solutions. Instead, customer requirements should as far as possible be formulated independent from potential solutions; they should describe the problem-space without unnecessarily constraining the solution space.

In contrast, system requirements are built on the basis of user requirements but from the perspective of the system under construction. They tend to describe in more detail which properties the system should have to fulfil the customer requirements. System requirements are used by developers to develop the system

Both, customer and system requirements can be of different types, for example, functional and non-functional requirements (Sommerville, 2019). Often, a customer requirements document is written by the customer and potential suppliers can propose system requirements specifications as part of their offer.

4.5 Limitations

Transitioning to agile, value-driven and continuous development methods, AV development companies are currently re-designing their approaches to requirements engineering. In addition, the raise of AI-technology in modern vehicles adds new complexity to defining requirements due to their inductive training nature (see the section 3.3.3 for some examples relating human factors requirements and AI). Moreover, while adopting agile methods, software teams aim to shorten time to market, and are at risk to focus too much on the technical perspective and to neglect human factor aspects.

Human factor knowledge often reveals requirements from the perspective of (human) users for a system under development. Requirements Engineering is in principle concerned with this, but few concrete suggestions exist with respect to how to practically integrate knowledge about human factors in RE for agile work. This is concerning, especially because of the multi-disciplinary nature of AV development that requires to align the opinion of experts from many domains (such as mechanical, hardware, software, as well as human factors). Given these challenges, it is unclear how HF knowledge can be systematically captured as requirements.

Future work needs to investigate how HF knowledge can be systematically captured as requirements in agile AV development.

Suggested Reading

General:

- Belani, H., Vukovic, M., & Car, Ž. (2019, September). Requirements engineering challenges in building AI-based complex systems. In *2019 IEEE 27th International Requirements Engineering Conference Workshops (REW)* (pp. 252-255). IEEE.
- Sommerville, I. (2019). Software Engineering, Global Edition.
- Wiegers, K., & Beatty, J. (2013). *Software requirements*. Pearson Education.

How to rely on effective use cases for defining requirement:

- Cockburn, A. (2001). *Writing effective use cases*. Pearson Education India.

Pragmatic approach to customer/supplier specifications in tender projects:

- Lauesen, S. (2002). *Software requirements: styles and techniques*. Pearson Education.

Comprehensive overview of techniques and modeling approaches:

- Pohl, K. (2010). *Requirements engineering: fundamentals, principles, and techniques*. Springer Publishing Company, Incorporated.

Approach for RE in large-scale agile development:

- Leffingwell, D. (2010). *Agile software requirements: lean requirements practices for teams, programs, and the enterprise*. Addison-Wesley Professional.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

5 Method champions

Due to the multidisciplinary essence of the SHAPE-IT project, many possible approaches and methods are available to the ESRs. It might be overwhelming at times. Therefore, a network of Method champions (MCs) and supporting method champions (SMCs) was created.

A MC is an ESR who uses a specific method extensively as part of their project. Therefore, each ESR was assigned to be a MC for such method. His or her role is to be a discussion partner for other ESRs who consider employing the respective method, or who is interested in knowing more about the method. MC is also responsible for training others in the respective method, either online or at network-wide training events. All ESRs are a leading MC for one method. Moreover, SMCs were assigned to the leading MCs. The role of a SMCs is to support the activities of the leading MC when necessary, especially when preparing a formal training for other ESRs. Table 9 presents all ESRs with their MC and SMC roles.

Table 9 Overview of the ESRs, method champions, and authorship

	Name	Method Champion	Supporting Method Champion for
ESR1	Nikol Figalova	Neuroergonomics	Driving simulator (planning and conducting)
			Methodology for setting requirements
			Self-report data
ESR2	Naomi Mbelekani	Trust and Acceptance Assessment related to AVs	On road studies
			Neuroergonomics
			AI based modelling
			Safety Assessment
ESR3	Chi Zhang	AI based modelling	Road user modelling part 2 (behavioural models)
			Safety Assessment
ESR4	Yue Yang	Pedestrian/cycling/other simulator	Road user modelling part 1 (rule-based models)
ESR5	Chen Peng	Driving simulator (planning and conducting)	Driving simulator (data assessment)
			Self-report data
ESR6	Mohamed Nasser	Driving simulator (data assessment)	Naturalistic driving studies
			Test-Track Studies and Wizard of Oz studies
			Driving simulator (planning and conducting)
			Neuroergonomics
ESR7	Liu Yuan-Cheng	Test-Track Studies and Wizard of Oz studies	Driving simulator (planning and conducting)
			Driving simulator (data assessment)
			Trust and Acceptance Assessment related to Avs
ESR8	Amna Pir Muhammed	Methodology for setting requirements	Safety Assessment
			Self-report data
ESR9	Wilbert Tabone	AR, VR studies	-
ESR10	Siri Hegna Berge	Self-report data	Pedestrian/cycling/other simulator
			AR, VR studies
			Trust and Acceptance Assessment related to Avs
ESR11	Sarang Jokhio	Naturalistic driving studies	Road user modelling part 1 (rule-based models)
			Road user modelling part 2 (behavioural models)
ESR12	Xiaolin He	On road studies	Driving simulator (planning and conducting)
			Driving simulator (data assessment)
ESR13	Amir Hossein Kalantari	Models based on psychological mechanisms	Pedestrian/cycling/other simulator
			Rule-based models
ESR14	Ali Mohammadi	Rule-based models	Models based on psychological mechanisms
ESR15	Xiaomi Yang	Safety Assessment	Rule-based models
			AI based modelling
			Models based on psychological mechanisms

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Conclusion

SHAPE-IT is a unique project which brings together researchers from various institutions across Europe, with diverse educational backgrounds, and different fields of expertise. Thanks to this diversity, we are able to see the problems of human factors aspects in the design and integration of AVs in cities of the future from a holistic perspective and to come with innovative ideas. The interdisciplinary approach is a must in order to facilitate development of safe and user-centred automated vehicles for urban environments. However, it brings certain challenges concerning the research methodology. We try to overcome these challenges by organizing and systemizing the empirical and modelling approaches employed throughout the project

This is a multi-author document, prepared by the 15 ESRs involved in SHAPE-IT. The aim was to provide a brief overview of the methods that will be employed in their research. Each section was prepared by the ESRs that plan to work with such method or approach, with co-authors as appropriate (e.g., similar and complementary knowledge and interest). The text is intended to be informative and easy to follow, rather than exhausting and comprehensive. It serves as a source of information and references for anyone interested in the SHAPE-IT activities, and in human factors research related to AV design in general.

References

Ackermann, C., Beggiato, M., Schubert, S., & Krems, J. F. (2019). An experimental study to investigate design and assessment criteria: What is important for communication between pedestrians and automated vehicles? *Applied ergonomics*, *75*, 272-282.

Advances in Intelligent Systems and Computing, vol 964 (pp. 145–156) Cham: Springer. doi:10.1007/978-3-030-20503-4_13

Agarwal, R. (2019). Validation of a Pedestrian Simulator for interaction between Pedestrians and Autonomous Vehicles.

Aghabayk, K., Sarvi, M., & Young, W. (2015). A State-of-the-Art Review of Car-Following Models with Particular Considerations of Heavy Vehicles. *Transport Reviews*, *35*(1), 82–105. https://doi.org/10.1080/01441647.2014.997323

Ahmed, K. I. (1999). Modeling drivers' acceleration and lane changing behavior (Doctoral dissertation, Massachusetts Institute of Technology).Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social LSTM: Human trajectory prediction in crowded spaces. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2016.110

Albaba, B. M., & Yildiz, Y. (2020). Driver Modeling through Deep Reinforcement Learning and Behavioral Game Theory. *arXiv preprint arXiv:2003.11071*.

Albert, M., Lange, A., Schmidt, A., Wimmer, M., & Bengler, K. (2015). Automateddriving–Assessment of interaction concepts under real driving conditions. *Procedia Manufacturing*, *3*, 2832-2839.

Alyavina, E., Nikitas, A., & Tchouamou Njoya, E. (2020). Mobility as a service and sustainable travel behaviour: A thematic analysis study. Transportation Research Part F: Traffic Psychology and Behaviour, 73, 362–381. https://doi.org/10.1016/j.trf.2020.07.004

Amado, H., Ferreira, S., Tavares, J. P., Ribeiro, P., & Freitas, E. (2020). Pedestrian–Vehicle Interaction at Unsignalized Crosswalks: A Systematic Review. *Sustainability*, *12*(7), 2805.

Andemach, T., Deville, G., & Mortier, L. (1993). The design of a real world wizard ofoz experiment for a speech driven telephone directory information system. In Third european conference on speech communication and tech

Antonson, H., Mårdh, S., Wiklund, M., & Blomqvist, G. (2009). Effect of surrounding landscape on driving behaviour: A driving simulator study. Journal of Environmental Psychology, 29(4), 493–502. https://doi.org/https://doi.org/10.1016/j.jenvp.2009.03.005

Attig, C., Wessel, D., & Franke, T. (2017). Assessing personality differences in human-technology interaction: An overview of key self-report scales to predict successful interaction. Communications in Computer and Information Science, 713, 19–29. https://doi.org/10.1007/978-3-319-58750-9_3

Artanto, D., Sulistyanto, M. P., Pranowo, I. D., & Pramesta, E. E. (2017). Drowsiness detection system based on eye-closure using a low-cost EMG and ESP8266. 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 235–238. https://doi.org/10.1109/icitisee.2017.8285502

Aycin, M. F., & Benekohal, R. F. (1999). Comparison of car-following models for simulation. *Transportation Research Record*, *1678*, 116–127. https://doi.org/10.3141/1678-15

Bärgman, J. (2016). Methods for analysis of naturalistic driving data in driver behavior research. Chalmers University of Technology.

Barnard, Y., & Carsten, O. (2010, April). Field Operational Tests: challenges and methods. In Proceedings of European Conference on Human Centred Design for Intelligent Transport Systems, Eds edn. HUMANIST publications, Lyon (pp. 323-332).

Barnard, Y., Innamaa, S., Koskinen, S., Gellerman, H., Svanberg, E., & Chen, H. (2016). Methodology for Field Operational Tests of Automated Vehicles. Transportation Research Procedia, 14, 2188–2196. https://doi.org/10.1016/j.trpro.2016.05.234

Bartneck, C., Kulić, D., Croft, E., & Zoghbi, S. (2009). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics*, *1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3

Bazzan, A. L. (2007). Traffic as a complex system: Four challenges for computer science and engineering. Proceedings of the XXXIV SEMISH. Citeseer, 2128.

Bazilinskyy, P., Dodou, D., & De Winter, J. (2019). Survey on eHMI concepts: The effect of text, color, and perspective. Transportation Research Part F: Traffic Psychology and Behaviour, 67, 175–194. https://doi.org/10.1016/j.trf.2019.10.013

Bazilinskyy, P., Kooijman, L., Dodou, D., & De Winter, J. C. F. (2020). Coupled simulator for research on the interaction between pedestrians and (automated) vehicles. *Driving Simulation Conference Europe.* Antibes, France.

Bazilinskyy, P., Stapel, J., de Koning, C., Lingmont, H., de Lint, T. S., van der Sijs, T. C., van den Ouden, F. C., Anema, F., & de Winter, J. C. F. (2018). Graded auditory feedback based on headway: An on-road pilot study. Varieties of Interaction: From User Experience to Neuroergonomics, 4959, 1–11.

Beggiato, M., Hartwich, F., & Krems, J. (2018). Using Smartbands, Pupillometry and Body Motion to Detect Discomfort in Automated Driving. Frontiers in Human Neuroscience, 12(September), 1–12. https://doi.org/10.3389/fnhum.2018.00338

Beggiato, M, Hartwich, F. and Krems, J. (2019). Physiological correlates of discomfort in automated driving. *Transportation Research Part F* 66. 445–458. doi.org/10.1016/j.trf.2019.09.018

Bengler, K., Omozik, K., & Müller, A. I. (2020). The Renaissance of Wizard of Oz (WoOz): Using the WoOz methodology to prototype automated vehicles. Proceedings of the Human Factors and Ergonomics Society Europe, 63-72.

Benzmüller, C., Fiedler, A., Gabsdil, M., Horacek, H., Kruijff-Korbayov, I., Pinkal, M., Wolska, M. (2003). A wizard-of-oz experiment for tutorial dialogues in mathematics. In Proceedings of the aied workshop on advanced technologies for mathematics education

Bellem, H., Klüver, M., Schrauf, M., Schöner, H. P., Hecht, H., & Krems, J. F. (2017). Can We Study Autonomous Driving Comfort in Moving-Base Driving Simulators? A Validation Study. Human Factors, 59(3), 442–456. https://doi.org/10.1177/0018720816682647

Bellem, H., Thiel, B., Schrauf, M., & Krems, J. F. (2018). Comfort in automated driving: An analysis of preferences for different automated driving styles and their dependence on personality traits. Transportation Research Part F: Traffic Psychology and Behaviour, 55, 90–100. https://doi.org/10.1016/j.trf.2018.02.036

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. Journal of Neuroscience Methods, 190(1), 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Benzmüller, C., Fiedler, A., Gabsdil, M., Horacek, H., Kruijff-Korbayov, I., Pinkal, M., . . .Wolska, M. (2003). A wizard-of-oz experiment for tutorial dialogues in mathemat-ics. In Proceedings of the aied workshop on advanced technologies for mathematics education

Biggar, M. (2020) Non-motorized Transport: Walking and Cycling. In: Leal Filho W., Azul A., Brandli L., Özuyar P., Wall T. (eds) Sustainable Cities and Communities. Encyclopedia of the UN Sustainable Development Goals. Springer, Cham. https://doi.org/10.1007/978-3-319-71061-7_1-1

Blana, E. (1996). Driving Simulator Validation Studies: A Literature Review.

Boda, C., Dozza, M., Bohman, K., Thalya, P., Larsson, A., & Lubbe, N. (2018). Modelling how drivers respond to a bicyclist crossing their path at an intersection: How do test track and driving simulator compare? *Accident Analysis & Prevention, 111*, 238-250. doi:10.1016/j.aap.2017.11.032

Boucsein, W. (2011). Electrodermal Activity (2nd ed. 2012 ed.). Springer.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks: London, New Delhi: Sage.

Brackstone, M., & McDonald, M. (1999). Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour, 2(4), 181-196. https://doi.org/10.1016/S1369-8478(00)00005-X

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2020). Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. Counselling and Psychotherapy Research, capr.12360. https://doi.org/10.1002/capr.12360

Brookhuis, K. A., & de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. Accident Analysis and Prevention, 42(3), 898–903. https://doi.org/10.1016/j.aap.2009.06.001

Brown, K., Driggs-Campbell, K., & Kochenderfer, M. J. (2020). Modeling and prediction of human driver behavior: A survey. *arXiv preprint arXiv:2006.08832*.

Brouwer, A.-M., Hogervorst, M. A., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. Journal of Neural Engineering, 9(4), 045008. https://doi.org/10.1088/1741-2560/9/4/045008

Busemeyer, J. R., Gluth, S., Rieskamp, J., & Turner, B. M. (2019). Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends in cognitive sciences*, *23*(3), 251-263.

Cafiso, S., & Pappalardo, G. (2020). Safety effectiveness and performance of lane support systems for driving assistance and automation – Experimental test and logistic regression for rare events. Accident Analysis and Prevention, 148(August 2020), 105791. https://doi.org/10.1016/j.aap.2020.105791

Calder, M., Craig, C., Culley, D., de Cani, R., Donnelly, C. A., Douglas, R., ... & Hinds, D. (2018). Computational modelling for decision-making: where, why, what, who and how. *Royal Society open science*, *5*(6), 172096.

Camerer, C. F. (2003). Behavioural studies of strategic thinking in games. *Trends in cognitive sciences*, *7*(5), 225-231.

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861-898.

Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures*, *12*(1), 25-44.

Carsten, O. (2014). Field Operational Tests: challenges and methods

Casas, S., Luo, W., & Urtasun, R. (2018). IntentNet: Learning to Predict Intention from Raw Sensor Data. CoRL.

Casello, J. M., & Usyukov, V. (2014). Modeling Cyclists' Route Choice Based on GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2430*(1), 155–161. https://doi.org/10.3141/2430-16

Castermans, J. (2017). Overview of the Data Collection. UDRIVE Deliverable 30.1. EU FP7 Project UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D30.1

Castleberry, A., & Nolen, A. (2018). Thematic analysis of qualitative research data: Is it as easy as it sounds?. *Currents in pharmacy teaching & learning*, *10*(6), 807–815. https://doi.org/10.1016/j.cptl.2018.03.019

Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of experimental psychology: human perception and performance*, *45*(6), 826.

Chang, K. (2015). Motion Analysis. *E-Design,* 391-462. doi:10.1016/b978-0-12-382038-9.00008-9

Chang, C. M., Toda, K., Igarashi, T., Miyata, M., & Kobayashi, Y. (2018, September). A video-based study comparing communication modalities between an autonomous car and a pedestrian. In *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 104-109).

Cleij, D., Venrooij, J., Pretto, P., Pool, D.M., Mulder, M. and Bülthoff, H.H. (2018, February). Continuous Subjective Rating of Perceived Motion Incongruence During Driving Simulation. *IEEE Transactions on Human-Machine Systems,* Vol 48, No. 1

Cockburn, A. (2001). *Writing effective use cases*. Pearson Education India.

Colombaroni, C., & Fusco, G. (2014). Artificial neural network models for car following: Experimental analysis and calibration issues. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, *18*(1), 5–16. https://doi.org/10.1080/15472450.2013.801717

Daneshi, A., Azarnoush, H., & Towhidkhah, F. (2020). A one-boundary drift-diffusion model for time to collision estimation in a simple driving task. *Journal of Cognitive Psychology*, *32*(1), 67-81.

Danner, S.; Pfromm, M.; Bengler, K. (2020). Does Information on Automated Driving Functions and the Way of Presenting It before Activation Influence Users' Behavior and Perception of the System? *Information*, *11*, 54.

Das, S., & Bowles, B. A. (1999, June). Simulations of highway chaos using fuzzy logic. In 18th International Conference of the North American Fuzzy Information Processing Society-NAFIPS (Cat. No. 99TH8397) (pp. 130-133). IEEE.

Dey, D., Habibovic, A., Löcken, A., Wintersberger, P., Pfleging, B., Riener, A., . . . Terken, J. (2020). Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives,* *7*, 100174. doi:10.1016/j.trip.2020.100174

De Clercq, K., Dietrich, A., Núñez Velasco, J. P., de Winter, J., & Happee, R. (2019). External human-machine interfaces on automated vehicles: Effects on pedestrian crossing decisions. *Human factors*, *61*(8), 1353-1370

De Luca, C. (2006). Electromyography. Encyclopedia of medical devices and instrumentation.

Deb, S., Carruth, D. W., Fuad, M., Stanley, L. M., & Frey, D. (2020). Comparison of child and adult pedestrian perspectives of external features on autonomous vehicles using virtual reality experiment. In N. Stanton (Ed.), *Advances in Human Factors of Transportation. AHFE 2019.*

Deb, S., Carruth, D. W., Sween, R., Strawderman, L., & Garrison, T. M. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied ergonomics*, *65*, 449-460.

Diaz-Piedra, C., Sebastián, M. V., & Di Stasi, L. L. (2020). EEG theta power activity reflects workload among army combat drivers: An experimental study. Brain Sciences, 10(4). https://doi.org/10.3390/brainsci10040199

Dietrich, A., Bengler, K., Feldstein, I. T., & Lehsing, C. (2018). Pedestrian simulators for traffic research: state of the art and future of a motion lab. *International Journal of Human Factors Modelling and Simulation*, *6*(4), 250. https://doi.org/10.1504/ijhfms.2018.10017411

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proceedings of the National Academy of Sciences, 113(10), 2636-2641. https://doi.org/10.1073/pnas.1513271113

Dlugosch, C., Conti, A. S., & Bengler, K. (2013). Driver Distraction through Conversation Measured with Pupillometry. 198–204. https://doi.org/10.17077/drivingassessment.1488

Dresner, D. (2015). ISO standards for user centered design and the specification of usability.

Du, N., Yang, X. J., & Zhou, F. (2020). Psychophysiological responses to takeover requests in conditionally automated driving. Accident Analysis and Prevention, 148(October). https://doi.org/10.1016/j.aap.2020.105804

Dul, J., Bruder, R., Buckle, P., Carayon, P., Falzon, P., Marras, W. S., ... & van der Doelen, B. (2012). A strategy for human factors/ergonomics: developing the discipline and profession. *Ergonomics*, *55*(4), 377-395

Dumbuya, A., Booth, A., Reed, N., Kirkham, A., Philpott, T., Zhao, J., & Wood, R. (2009). Complexity of traffic interactions: Improving behavioural intelligence in driving simulation scenarios. In Complex systems and self-organization modelling (pp. 201-209). Springer, Berlin, Heidelberg.

Dużmańska, N., Strojny, P. ł., & Strojny, A. (2018). Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness. Frontiers in Psychology, 9, 1–10. https://doi.org/10.3389/fpsyg.2018.02132

Eriksson, G., Svenson, O., & Eriksson, L. (2013). The time-saving bias: Judgements, cognition and perception. Judgment and Decision Making, 8, 492–497.

Eisma, R., Dickinson, A., Goodman, J., Mival, O., Syme, A., Tiwari, L. (2003) Mutual inspiration in the development of new technology. *In Proceedings of Include 2003*

ERTRAC (2017). Automated Driving Roadmap. European Road Transport Research Advisory Council, Brussel.

Faas, S., Mathis, L., Baumann, M. (2020). External HMI for self-driving vehicles: Which information shall be displayed? *Transportation Research Part F*, *68*, 171-186

Feierle, A.; Danner, S.; Steininger, S.; Bengler, K. (2020). Information Needs and Visual Attention during Urban, Highly Automated Driving—An Investigation of Potential Influencing Factors. *Information*, *11*, 62.

Feldstein, I. T., Lehsing, C., Dietrich, A., & Bengler, K. (2018). Pedestrian simulators for traffic research: state of the art and future of a motion lab. *International Journal of Human Factors Modelling and Simulation*, *6*, 250–265. doi:10.1504/ijhfms.2018.096128

Feldstein, I. T., & Dyszak, G. N. (2020). Road crossing decisions in real and virtual environments: A comparative study on simulator validity. *Accident Analysis and Prevention*, *137*(February), 105356. https://doi.org/10.1016/j.aap.2019.105356

Feng, S., Feng, Y., Yan, X., Shen, S., Xu, S., & Liu, H. X. (2020). Safety assessment of highly automated driving systems in test tracks: A new framework. *Accident Analysis & Prevention*, *144*, 105664.

Fox, B. H. (1960). Engineering and psychological uses of a driving simulator. *Highway Research Board Bulletin*, *261*, 14–37. http://dx.doi.org/

French, D. J., West, R. J., Elander, J., & WILDING, J. M. (1993). Decision-making style, driving style, and self-reported involvement in road traffic accidents. Ergonomics, 36(6), 627-644.

Fridman, L., Mehler, B., Xia, L., Yang, Y., Facusse, L. Y., & Reimer, B. (2017). To Walk or Not to Walk: Crowdsourced Assessment of External Vehicle-to-Pedestrian Displays. ArXiv. http://arxiv.org/abs/1707.02698

Fridman, L., Brown, D. E., Glazer, M., Angell, W., Dodd, S., Jenik, B., ... & Reimer, B. (2019). MIT advanced vehicle technology study: Large-scale naturalistic driving study of driver behavior and interaction with automation. IEEE Access, 7, 102021-102038.

Fuest, T., Michalowski, L., Träris, L., Bellem, H., & Bengler, K. (2018). Using the driving behavior of an automated vehicle to communicate intentions-a wizard of oz study. *In 2018 21st international conference on intelligent transportation systems (itsc)*(pp.3596–3601).

Glinz, M. (2007). On non-functional requirements. In *15th IEEE International Requirements Engineering Conference (RE 2007)* (pp. 21-26). IEEE.

Goeree, J. K., & Holt, C. A. (2004). A model of noisy introspection. *Games and Economic Behavior*, *46*(2), 365-382.

Gold, C., Happee, R., & Bengler, K. (2018). Modeling take-over performance in level 3 conditionally automated vehicles. Accident Analysis and Prevention, 116, 3–13. https://doi.org/10.1016/j.aap.2017.11.009

Golman, R., Bhatia, S., & Kane, P. B. (2019). The dual accumulator model of strategic deliberation and decision making. *Psychological Review*.

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. Transportation Research Part B: Methodological, 15(2), 105-111. https://doi.org/10.1016/0191-2615(81)90037-0

Gössling, S., Cohen, S. A., & Hares, A. (2016). Inside the black box: EU policy officers' perspectives on transport and climate change mitigation. Journal of Transport Geography, 57, 83–93. https://doi.org/10.1016/j.jtrangeo.2016.10.002

Guo, Y., Wu, Y., Lu, J., & Zhou, J. (2019). Modeling the Unobserved Heterogeneity in E-bike Collision Severity Using Full Bayesian Random Parameters Multinomial Logit Regression. *Sustainability*, *11*(7), 2071. https://doi.org/10.3390/su11072071

Guo, R. Y., Wong, S. C., Huang, H. J., Zhang, P., & Lam, W. H. (2010). A microscopic pedestrian-simulation model and its application to intersecting flows. *Physica A: Statistical Mechanics and its Applications*, *389*(3), 515-526.

Gupta, A., Johnson, J., Fei-Fei, L., Savarese, S., & Alahi, A. (2018). Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2018.00240

Gyllenhammar, M., Johansson, R., Warg, F., Chen, D., Heyn, H.-M., Sanfridsson, M., … Ursing, S. (2020). Towards an Operational Design Domain That Supports the Safety Argumentation of an Automated Driving System. In Proceeding of the 10th European Congress on Embedded Real Time Software and Systems. Toulouse. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-267132

Habibovic, A., Andersson, J., Nilsson, M., Lundgren, V. M., & Nilsson, J. (2016, June). Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches. In 2016 IEEE Intelligent Vehicles Symposium (IV) (pp. 32-37). IEEE.

Habibovic, A., Lundgren, V. M., Andersson, J., Klingegård, M., Lagström, T., Sirkka, A., ... & Larsson, P. (2018). Communicating intent of automated vehicles to pedestrians. *Frontiers in psychology*, *9*, 1336.

Hafner, R. J., Walker, I., & Verplanken, B. (2017). Image, not environmentalism: A qualitative exploration of factors influencing vehicle purchasing decisions. Transportation Research Part A: Policy and Practice, 97, 89–105. https://doi.org/10.1016/j.tra.2017.01.012

Hagenzieker, M. P., van der Kint, S., Vissers, L., van Schagen, I. N. L. G., de Bruin, J., van Gent, P., & Commandeur, J. J. F. (2020). Interactions between cyclists and automated vehicles: Results of a photo experiment. Journal of Transportation Safety & Security, 12(1), 94–115. https://doi.org/10.1080/19439962.2019.1591556

Halati, A., Lieu, H., & Walker, S. (1997). CORSIM-corridor traffic simulation model. In Traffic Congestion and Traffic Safety in the 21st Century: Challenges, Innovations, and OpportunitiesUrban Transportation Division, ASCE; Highway Division, ASCE; Federal

Highway Administration, USDOT; and National Highway Traffic Safety Administration, USDOT..

Hankey, J. M., Perez, M. A., & McClafferty, J. A. (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets. Virginia Tech Transportation Institute.

Hartwich, F., Witzlack, C., Beggiato, M., & Krems, J. F. (2019). The first impression counts–A combined driving simulator and test track study on the development of trust and acceptance of highly automated driving. *Transportation research part F: traffic psychology and behaviour*, *65*, 522-535.

Hecht, T.; Kratzert, S.; Bengler, K. (2020). The Effects of a Predictive HMI and Different Transition Frequencies on Acceptance, Workload, Usability, and Gaze Behavior during Urban Automated Driving. *Information*, *11*, 73.

Helbing, D., & Molnár, P. (1995). Social force model for pedestrian dynamics. *Physical Review E*, *51*(5), 4282–4286. https://doi.org/10.1103/PhysRevE.51.4282

Helbing, D., Farkas, I. J., Molnar, P., & Vicsek, T. (2002). Simulation of pedestrian crowds in normal and evacuation situations. *Pedestrian and evacuation dynamics*, *21*(2), 21-58.

Hesenius, M., Börsting, I., Meyer, O., & Gruhn, V. (2018, September). Don't panic! guiding pedestrians in autonomous traffic with augmented reality. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct* (pp. 261-268).

Herpers, R., Heiden, W., Kutz, M., Scherfgen, D., Hartmann, U., Bongartz, J., & Schulzyk, O. (2008). FIVIS bicycle simulator - An immersiv game platform for physical activities. ACM Future Play 2008 International Academic Conference on the Future of Game Design and Technology, Future Play: Research, Play, Share, 244–247. https://doi.org/10.1145/1496984.1497035

Hidas, P. (2005). Modelling vehicle interactions in microscopic simulation of merging and weaving. Transportation Research Part C: Emerging Technologies, 13(1), 37-62.

Hidas, P. (2002). Modelling lane changing and merging in microscopic traffic simulation. Transportation Research Part C: Emerging Technologies, 10(5-6), 351-371.

Hoffman, J. D., Lee, J. D., Brown, T. L., & McGehee, D. V. (2002). Comparison of Driver Braking Responses in a High-Fidelity Simulator and on a Test Track. Transportation Research Record, 1803(1), 59–65. https://doi.org/10.3141/1803-09

Holleman, G. A., Hooge, I. T. C., Kemner, C., & Hessels, R. S. (2020). The 'Real-World Approach' and Its Problems: A Critique of the Term Ecological Validity. *Frontiers in Psychology*, *11*, 1–2. https://doi.org/10.3389/fpsyg.2020.00721

Howard, Z. L., Evans, N. J., Innes, R. J., Brown, S. D., & Eidels, A. (2020). How is multi-tasking different from increased difficulty?. *Psychonomic Bulletin & Review*

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Hudson, C. R., Deb, S., Carruth, D. W., McGinley, J., & Frey, D. (2018, July). Pedestrian perception of autonomous vehicles with external interacting features. In *International Conference on Applied Human Factors and Ergonomics* (pp. 33-39). Springer, Cham.

Huang, L. (2014). A neural network based modeling and simulation of bicycle conflict avoidance behaviors at non- Signalized intersections. *2014 10th International Conference on Natural Computation, ICNC 2014*, *51108192*, 160–165. https://doi.org/10.1109/ICNC.2014.6975828

Huang, L., Wu, J., You, F., Lv, Z., & Song, H. (2017). Cyclist social force model at unsignalized intersections with heterogeneous traffic. *IEEE Transactions on Industrial Informatics*, *13*(2), 782–792. https://doi.org/10.1109/TII.2016.2597744

IEEE (1990). IEEE standard glossary of software engineering terminology. IEEE Std 610.12-1990, pages 1–84.

International Transport Forum. (2019). New Directions for Data-Driven Transport Safety. International Transport Forum. https://www.itf-oecd.org/sites/default/files/docs/new-directions-data-driven-transport-safety_0.pdf

Jarosch, O., Paradies, S., Feiner, D., & Bengler, K. (2019). Effects of non-driving relatedtasks in prolonged conditional automated driving–a wizard of oz on-road approachin real traffic environment.Transportation research part F: traffic psychology andbehaviour,65, 292–305

Jeihani, M., NarooieNezhad, S., & Kelarestaghi, K. B. (2017). Integration of a driving simulator and a traffic simulator case study: Exploring drivers' behavior in response to variable message signs. *IATSS research*, *41*(4), 164-171.

Kamel, N., & Malik, A. S. (2015). EEG/ERP analysis: methods and applications. CRC Press.

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I* (pp. 99-127).

Kaleefathullah, A. A., Merat, N., Lee, Y. M., Eisma, Y. B., Madigan, R., Garcia, J., & De Winter, J. (in press). External Human-Machine Interfaces can be misleading. An examination of trust development and misuse in a CAVE-based pedestrian simulation environment. *Human Factors*.

Karjalainen, P., Pirjola, L., Heikkilä, J., Lähde, T., Tzamkiozis, T., Ntziachristos, L., Keskinen, J., & Rönkkö, T. (2014). Exhaust particles of modern gasoline vehicles: A laboratory and an on-road study. Atmospheric Environment, 97, 262–270. https://doi.org/10.1016/j.atmosenv.2014.08.025

Käthner, I., Wriessnegger, S. C., Müller-Putz, G. R., Kübler, A., & Halder, S. (2014). Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain–computer interface. Biological Psychology, 102, 118–129. https://doi.org/10.1016/j.biopsycho.2014.07.014

Kalinic, M., & Krisp, J. M. (2019). Fuzzy inference approach in traffic congestion detection. *Annals of GIS*, *25*(4), 329–336. https://doi.org/10.1080/19475683.2019.1675760

Kaptein, N. A., Theeuwes, J., & van der Horst, R. (1996). Driving Simulator Validity: Some Considerations. *Transportation Research Record: Journal of the Transportation Research Board*, *1550*(1), 30–36. https://doi.org/10.1177/0361198196155000105

Karjanto, J., Yusof, N. M., Wang, C., Terken, J., Delbressine, F., & Rauterberg, M. (2018). The effect of peripheral visual feedforward system in enhancing situation awareness and mitigating motion sickness in fully automated driving. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 678-692.

Kelley, J. F. (1984). An iterative design methodology for user-friendlynatural language office information applications.ACM Trans. Inf. Syst.,2(1),26–41. Retrieved fromhttps://doi.org/10.1145/357417.357420doi: 10.1145/357417.357420

Kesting, A., Treiber, M., & Helbing, D. (2007). General lane-changing model MOBIL for car-following models. Transportation Research Record, 1999(1), 86-94.

Khodayari, A., Ghaffari, A., Kazemi, R., & Braunstingl, R. (2012). A modified car-following model based on a neural network model of the human driver effects. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *42*(6), 1440–1449. https://doi.org/10.1109/TSMCA.2012.2192262

Kita, H. (1999). A merging–giveway interaction model of cars in a merging section: a game theoretic analysis. Transportation Research Part A: Policy and Practice, 33(3-4), 305-312.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Research Reviews, 29(2–3), 169–195. https://doi.org/10.1016/s0165-0173(98)00056-3

Kolekar, S., de Winter, J. and Abbink, D. (2020). Which parts of the road guide obstacle avoidance? Quantifying the driver's risk field. *Applied Ergonomics* 89. doi.org/10.1016/j.apergo.2020.103196

Kouskoulis, G., & Antoniou, C. (2017). Systematic review of pedestrian simulation models with a focus on emergency situations. *Transportation Research Record*, *2604*(1), 111-119.

Kundinger, T., Riener, A., Sofra, N., & Weigl, K. (2020, March). Driver drowsiness in automated and manual driving: insights from a test track study. In *Proceedings of the 25th International Conference on Intelligent User Interfaces* (pp. 369-379).

Kyriakidis, M., de Winter, J. C. F., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., Martens, M. H., Bengler, K., Andersson, J., Merat, N., Reed, N., Flament, M., Hagenzieker, M., & Happee, R. (2019). A human factors perspective on automated driving. Theoretical Issues in Ergonomics Science, 20(3), 223–249. https://doi.org/10.1080/1463922X.2017.1293187

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. *Procedia Manufacturing, 3*, 2403-2409. doi:10.1016/j.promfg.2015.07.499

Lakehal, A., Lepreux, S., Efstratiou, C., Christophe, K., & Nicolaou, P. (2020). Investigating Smartphones and AR Glasses for Pedestrian Navigation and their Effects in Spatial Knowledge Acquisition.

Langdridge, D., & Hagger-Johnson, G. (2009). Introduction to Research Methods and Data Analysis in Psychology. Pearson Education.

Lárraga, M. E., del Río, J. A., & Alvarez-Icaza, L. (2005). Cellular automata for one-lane traffic flow modeling. *Transportation Research Part C: Emerging Technologies*, *13*(1), 63–74. https://doi.org/10.1016/j.trc.2004.12.001

Lauesen, S. (2002). *Software requirements: styles and techniques*. Pearson Education.

Larusdottir, M., Gulliksen, J., & Cajander, Å. (2017). A license to kill–Improving UCSD in Agile development. *Journal of Systems and Software*, *123*, 214-222.

Lawson, A. R., Pakrashi, V., Ghosh, B., & Szeto, W. Y. (2013). Perception of safety of cyclists in Dublin City. *Accident Analysis & Prevention*, *50*, 499-511.

Lee, B. G., Chong, T. W., Lee, B. L., Park, H. J., Kim, Y. N., & Kim, B. (2017). Wearable Mobile-Based Emotional Response-Monitoring System for Drivers. IEEE Transactions on Human-Machine Systems, 47(5), 636–649. https://doi.org/10.1109/thms.2017.2658442

Lee, O., Dialynas, G., de Winter, J., Happee, R., Schwab, A. L., De Winter, J. ;, & Happee, R. ; (2017). Description of a model based bicycle simulator. International Cycling Safety Conference, 6th Annual, 4–8. http://pure.tudelft.nl/ws/files/28598463/58_lee_dialynas_winter_happee_schwab.pdf

Leffingwell, D., & Widrig, D. (2003). *Managing Software Requirements: A Use Case Approach, Second Edition Publisher: Addison Wesley Pub ISBN: 0-321-12247-X Pages*, *544*.

Leffingwell, D. (2010). *Agile software requirements: lean requirements practices for teams, programs, and the enterprise*. Addison-Wesley Professional.

Li, J., Ma, H., & Tomizuka, M. (2019). Conditional Generative Neural System for Probabilistic Trajectory Prediction. IEEE International Conference on Intelligent Robots and Systems. https://doi.org/10.1109/IROS40897.2019.8967822

Likert, R. (1932). A technique for the measurement of attitudes. Archives of psychology.

Löcken, Andreas, Carmen Golling, and Andreas Riener. 2019. “How Should Automated Vehicles Interact with Pedestrians? A Comparative Analysis of Interaction Concepts in Virtual Reality.” in Proceedings - 11th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2019.

Lohani, M., Payne, B. R., & Strayer, D. L. (2019). A review of psychophysiological measures to assess cognitive states in real-world driving. Frontiers in Human Neuroscience, 13(March), 1–27. https://doi.org/10.3389/fnhum.2019.00057

Lotz, A., Russwinkel, N., & Wohlfarth, E. (2020). Take-over expectation and criticality in Level 3 automated driving: a test track study on take-over behavior in semi-trucks. *Cognition, Technology & Work*, 1-12.

Løvås, G. G. (1994). Modeling and simulation of pedestrian traffic flow. *Transportation Research Part B: Methodological*, *28*(6), 429-443.

Li, S., Li, N., Girard, A., & Kolmanovsky, I. (2019, December). Decision making in dynamic and interactive environments based on cognitive hierarchy theory, Bayesian inference, and predictive control. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (pp. 2181-2187). IEEE.

Li, J., Ma, H., & Tomizuka, M. (2019). Conditional Generative Neural System for Probabilistic Trajectory Prediction. IEEE International Conference on Intelligent Robots and Systems. https://doi.org/10.1109/IROS40897.2019.8967822

Liu, N., Nikitas, A., & Parkinson, S. (2020). Exploring expert perceptions about the cyber security and privacy of Connected and Autonomous Vehicles: A thematic analysis approach. *Transportation Research Part F: Traffic Psychology and Behaviour*, 66–86. https://doi.org/10.1016/j.trf.2020.09.019

Liu, Y.-C. (2003). Effects of using head-up display in automobile context on attention demand and driving performance. Displays, 24(4), 157–165. https://doi.org/https://doi.org/10.1016/j.displa.2004.01.001

Louw, T., Hajiseyedjavadi, F. H., Jamson, H., Romano, R., Boer, E., & Merat, N. (2019). The Relationship between Sensation Seeking and Speed Choice in Road Environments with Different Levels of Risk. April, 29–35. https://doi.org/10.17077/drivingassessment.1671

Lu, Q., Tettamanti, T., Hörcher, D., & Varga, I. (2020). The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation. Transportation Letters, 12(8), 540-549.

Lundgren, Victor Malmsten, Azra Habibovic, Jonas Andersson, Tobias Lagström, Maria Nilsson, Anna Sirkka, Johan Fagerlönn, Rikard Fredriksson, Claes Edgren, Stas Krupenia, and Dennis Saluäär. 2017. "Will There Be New Communication Needs When Introducing Automated Vehicles to the Urban Context?" Advances in Intelligent Systems and Computing 484:485–97.

Ma, C., Huang, J. Bin, Yang, X., & Yang, M. H. (2015). Hierarchical convolutional features for visual tracking. Proceedings of the IEEE International Conference on Computer Vision. https://doi.org/10.1109/ICCV.2015.352

Malik, A. S., & Amin, H. U. (2017). Designing EEG Experiments for Studying the Brain: Design Code and Example Datasets (1st ed.). Academic Press.

Mallaro, S., Rahimian, P., O'Neal, E. E., Plumert, J. M., & Kearney, J. K. (2017, November). A comparison of head-mounted displays vs. large-screen displays for an interactive pedestrian simulator. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (pp. 1-4).

Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. Accident Analysis & Prevention, 95, 209–226. https://doi.org/10.1016/j.aap.2016.07.007

Markkula, G., Boer, E., Romano, R., & Merat, N. (2018a). Sustained sensorimotor control as intermittent decisions about prediction errors: Computational framework and application to ground vehicle steering. *Biological cybernetics*, *112*(3), 181-207.

Markkula, G., Romano, R., Madigan, R., Fox, C. W., Giles, O. T., & Merat, N. (2018b). Models of human decision-making as tools for estimating and optimizing impacts of vehicle automation. *Transportation research record*, *2672*(37), 153-163.

Maruhn, P., Dietrich, A., Prasch, L., & Schneider, S. (2020, March). Analyzing Pedestrian Behavior in Augmented Reality—Proof of Concept. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (pp. 313-321). IEEE

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and economic behavior*, *10*(1), 6-38.

Melcher, V., Rauh, S., Diederichs, F., Widlroither, H., & Bauer, W. (2015). Take-Over Requests for Automated Driving. Procedia Manufacturing, 3, 2867–2873. https://doi.org/https://doi.org/10.1016/j.promfg.2015.07.788

Mendel, J. M. (1995). Fuzzy Logic Systems for Engineering: A Tutorial. *Proceedings of the IEEE*, *83*(3), 345–377. https://doi.org/10.1109/5.364485

Montuwy, A., Cahour, B., & Dommes, A. (2018, April). Older Pedestrians Navigating With AR Glasses and Bone Conduction Headset. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-6).

Moridpour, S., Sarvi, M., & Rose, G. (2010). Lane changing models: a critical review. *Transportation letters*, *2*(3), 157-173

Morton, R. D., & White, M. J. (2013). Revised reinforcement sensitivity theory: The impact of FFFS and stress on driving. Personality and Individual Differences, 54(1), 57–63. https://doi.org/https://doi.org/10.1016/j.paid.2012.08.005

Mirvis, D. M., & Goldberger, A. L. (2001). Electrocardiography. Heart disease, 1, 82-128.

Missonnier, P., Deiber, M.-P., Gold, G., Millet, P., Gex-Fabry Pun, M., Fazio-Costa, L., Giannakopoulos, P., & Ibáñez, V. (2006). Frontal theta event-related synchronization: comparison of directed attention and working memory load effects. Journal of Neural Transmission, 113(10), 1477–1486. https://doi.org/10.1007/s00702-005-0443-9

Mullen, N., Charlton, J., Devlin, A., & Bedard, M. (2011). Simulator validity: Behaviours observed on the simulator and on the road. In D. L. Fisher, M. Rizzo, J. K. Caird, & J. D. Lee (Eds.), Handbook of Driving Simulation for Engineering, Medicine and Psychology (1st ed., pp. 1-18). CRC Press.

Mühl, K., Strauch, C., Grabmaier, C., Reithinger, S., Huckauf, A., & Baumann, M. (2019). Get Ready for Being Chauffeured: Passenger's Preferences and Trust While Being Driven by Human and Automation. *Human Factors.* https://doi.org/10.1177/0018720819872893

Müller, A. I., Weinbeer, V., & Bengler, K. (2019). Using the wizard of oz paradigm toprototype automated vehicles: Methodological challenges. New York, NY, USA:Association for Computing Machine

Nagel, K., Wolf, D. E., Wagner, P., & Simon, P. (1998). Two-lane traffic rules for cellular automata: A systematic approach. Physical Review E, 58(2), 1425.

Nazemi, M., van Eggermond, M. A., Erath, A., & Axhausen, K. W. (2018). Studying cyclists' behavior in a non-naturalistic experiment utilizing cycling simulator with immersive virtual reality. Arbeitsberichte Verkehrs-und Raumplanung, 1383.

Nazemi, M. ;, Van Eggermond, M. A. B. ;, Erath, A. ;, Schaffner, D. ;, Joos, M. ;, Axhausen, K. W., Nazemi, M., Van Eggermond, M., Erath, A., Schaffner, D., & Joos, M. (2019). Studying bicyclists' perceived level of safety using a cycling simulator combined with immersive virtual reality. https://doi.org/10.3929/ethz-b-000356563

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis. International Journal of Qualitative Methods, 16(1), 160940691773384. https://doi.org/10.1177/1609406917733847

Nuñez Velasco, J. P., Farah, H., van Arem, B., & Hagenzieker, M. P. (2019). Studying pedestrians' crossing behavior when interacting with automated vehicles using virtual reality. Transportation Research Part F: Traffic Psychology and Behaviour, 66, 1–14. https://doi.org/10.1016/j.trf.2019.08.015

O'Hern, S., Oxley, J., & Stevenson, M. (2017). Validation of a bicycle simulator for road safety research. Accident Analysis and Prevention, 100, 53–58. https://doi.org/10.1016/j.aap.2017.01.002

Oliveira, L., Burns, C., Luton, J., Iyer, S., & Birrell, S. (2020). The influence of system transparency on trust: Evaluating interfaces in a highly automated vehicle. *Transportation research part F: traffic psychology and behaviour*, *72*, 280-296.

Otherson, I., Conti-Kufner, A. S., Dietrich, A., Maruhn, P., & Bengler, K. (2018). Designing for automated vehicle and pedestrian communication: Perspectives on eHMIs from older and younger persons. *Proceedings of the Human Factors and Ergonomics Society Europe*, 135-148.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). The measurement of meaning (No. 47). University of Illinois press.

Oxley, J. A., Ihsen, E., Fildes, B. N., Charlton, J. L., & Day, R. H. (2005). Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. Accident Analysis and Prevention, 37(5), 962–971. https://doi.org/10.1016/j.aap.2005.04.017

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Ozbay, K., Yang, H., Morgul, E. F., Mudigonda, S., & Bartin, B. (2014, January). Big data and the calibration and validation of traffic simulation models. In Traffic and Transportation Simulation-Looking Back and Looking Ahead: Celebrating 50 Years of Traffic Flow Theory, a Workshop (pp. 92-122). Transportation Research Board Washington, DC.

Özkan, T., & Lajunen, T. (2005). Why are there sex differences in risky driving? The relationship between sex and gender-role on aggressive driving, traffic offences, and accident involvement among young Turkish drivers. Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 31(6), 547-558.

Paddeu, D., Parkhurst, G., & Shergold, I. (2020). Passenger comfort and trust on first-time use of a shared autonomous shuttle vehicle. Transportation Research Part C, 115(March), 102604. https://doi.org/10.1016/j.trc.2020.02.026

Pai, G., Widrow, S., Radadiya, J., Fitzpatrick, C. D., Knodler, M., & Pradhan, A. K. (2020). A Wizard-of-Oz experimental approach to study the human factors of automated vehicles: Platform and methods evaluation. *Traffic injury prevention,* 1-5.

Palinko, O., & Kun, A. L. (2012). Exploring the effects of visual cognitive load and illumination on pupil diameter in driving simulators. Eye Tracking Research and Applications Symposium (ETRA), 413–416. https://doi.org/10.1145/2168556.2168650

Palmeiro, A. R., van der Kint, S., Vissers, L., Farah, H., de Winter, J. C., & Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. Transportation research part F: traffic psychology and behaviour, 58, 1005-1020.

Parker, D., Stradling, S. G., & Manstead, A. S. (1996). Modifying beliefs and attitudes to exceeding the speed limit: an intervention study based on the theory of planned behavior 1. Journal of Applied Social Psychology, 26(1), 1-19.

Paxion, J., Galy, E., & Berthelon, C. (2014). Mental workload and driving. Frontiers in Psychology, 5, 1–11. https://doi.org/10.3389/fpsyg.2014.01344

Perez, D., Hasan, M., Shen, Y., & Yang, H. (2019). Ar-ped: A framework of augmented reality enabled pedestrian-in-the-loop simulation. *Simulation Modelling Practice and Theory*, *94*, 237-249.

Pfleging, B., & Schmidt, A. (2015). driving-related activities in the car: Defining driver activities for manual and automated driving. In *Workshop on experiencing autonomous vehicles: Crossing the boundaries between a drive and a ride at CHI* (Vol. 15).

Pipes, L. A. (1966). Car following models and the fundamental diagram of road traffic. Transportation Research/UK/.

Pohl, K. (2010). *Requirements engineering: fundamentals, principles, and techniques*. Springer Publishing Company, Incorporated.

Poisson, C., Barré, J., Bourmaud, G., & Forzy, J. F. (2020, April). Driver Behavior in Conditional Automation: Comparison of Driving Simulator and Wizard of Oz Conditions. *In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-7).

Portouli, E., Bekiaris, E., Papakostopoulos, V., & Maglaveras, N. (2007). On-road experiment for collecting driving behavioural data of sleepy drivers. Somnologie, 11(4), 259–267. https://doi.org/10.1007/s11818-007-0319-3

Protzak, J., & Gramann, K. (2018). Investigating Established EEG Parameter During Real-World Driving. Frontiers in Psychology, 9, 1–11. https://doi.org/10.3389/fpsyg.2018.02289

Radhakrishnan, V., Merat, N., Louw, T., Lenné, M. G., Romano, R., Paschalidis, E., Hajiseyedjavadi, F., Wei, C., & Boer, E. R. (2020). Measuring drivers' physiological response to different vehicle controllers in highly automated driving (HAD): Opportunities for establishing real-time values of driver discomfort. Information (Switzerland), 11(8), 1–14. https://doi.org/10.3390/INFO11080390

Rahman, M., Chowdhury, M., Xie, Y., & He, Y. (2013). Review of microscopic lane-changing models and future research opportunities. *IEEE Transactions on Intelligent Transportation Systems*, *14*(4), 1942–1956. https://doi.org/10.1109/TITS.2013.2272074

Reason, J. T., and Brand, J. J. (1975). Motion Sickness. Oxford: Academic Press.

Riccio, G. E., and Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural stability. Ecol. Psychol. 3, 195–240. doi: 10.1207/s15326969eco0303_2

Scholar, C. K. (n.d.). *Ling Huang*. 13–19.

Stapel, J., Mullakkal-babu, F. A., & Happee, R. (2017). Driver behavior and workload in an on-road automated vehicle. Proceedings of Road Safety & Simulation International Conference, October, 1–10.

Szalay, Z., Tettamanti, T., Esztergár-Kiss, D., Varga, I., & Bartolini, C. (2018). Development of a test track for driverless cars: vehicle design, track configuration, and liability considerations. *Periodica Polytechnica Transportation Engineering*, *46*(1), 29-35.

Ratcliff, R., & Strayer, D. (2014). Modeling simple driving tasks with a one-boundary diffusion model. *Psychonomic bulletin & review*, *21*(3), 577-589.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in cognitive sciences*, *20*(4), 260-281.

Rickert, M., Nagel, K., Schreckenberg, M., & Latour, A. (1996). Two lane traffic simulations using cellular automata. Physica A: Statistical Mechanics and its Applications, 231(4), 534-550.

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Riegler, A., Riener, A., & Holzmann, C. (2020, November). A Research Agenda for Mixed Reality in Automated Vehicles. In *19th International Conference on Mobile and Ubiquitous Multimedia* (pp. 119-131).

Rossner, P., & Bullinger, A. C. (2020). How Do You Want to be Driven? Investigation of Different Highly-Automated Driving Styles on a Highway Scenario. Advances in Intelligent Systems and Computing, 964, 36–43. https://doi.org/10.1007/978-3-030-20503-4_4

Rusch, M. L., Schall Jr, M. C., Lee, J. D., Dawson, J. D., & Rizzo, M. (2014). Augmented reality cues to assist older drivers with gap estimation for left-turns. *Accident Analysis & Prevention*, *71*, 210-221.

Saldaña, J. (2009). Coding Manual for Qualitative Researchers. SAGE Publications Inc. https://www.researchgate.net/profile/Mohammed_Ishaq4/post/Tips_and_guidelines_for_qualitative_coding_and_text_analysis_using_Nvivo/attachment/5ab9f2ffb53d2f0bba5a8440/AS%3A608695847231489%401522135807434/download/Saldana+-+2013+-+The+Coding+Manual+for+Qua

Sargoni, O. T., & Manley, E. (2020, November). A sequential sampling model of pedestrian road crossing choice. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoSpatial Simulation* (pp. 10-19).

Schepers, J. P., Kroeze, P. A., Sweers, W., & Wüst, J. C. (2011). Road factors and bicycle–motor vehicle crashes at unsignalized priority intersections. Accident Analysis & Prevention, 43(3), 853-861.

Schneider, S., & Bengler, K. (2020). Virtually the same? Analysing pedestrian behaviour by means of virtual reality. *Transportation Research Part F: Traffic Psychology and Behaviour*, *68*, 231–256. doi:10.1016/j.trf.2019.11.005

Schneider, S., Wollersheim, J., Krcmar, H., & Sunyaev, A. (2018). How do requirements evolve over time? A case study investigating the role of context and experiences in the evolution of enterprise software requirements. *Journal of Information Technology*, *33*(2), 151-170.

Schramka, F., Arisona, S., Joos, M., & Erath, A. (2017). Development of virtual reality cycling simulator. *Arbeitsberichte Verkehrs-und Raumplanung*, *1244*.

Schwebel, D. C., Gaines, J., & Severson, J. (2008). Validation of virtual reality as a tool to understand and prevent child pedestrian injury. *Accident Analysis & Prevention*, *40*(4), 1394-1400.

Schwalm, M., Keinath, A., & Zimmer, H. D. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. Human Factors for Assistance and Automation, May 2014, 1–13.

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. Frontiers in Public Health, 5(September), 1–17. https://doi.org/10.3389/fpubh.2017.00258

SHAPE-IT

This project has received funding from the European Community's
Horizon 2020 Framework Programme under grant agreement 860410

Shechtman, O., Classen, S., Awadzi, K., & Mann, W. (2009). Comparison of Driving Errors Between On-the-Road and Simulated Driving Assessment: A Validation Study. *Traffic Injury Prevention*, *10*(4), 379-385. doi: 10.1080/15389580902894989

Sherry, J., Beckwith, R., Esme, A.A., & Tanriover, C. (2018). Getting things done in an automated vehicle. In Social Robots in the Wild Workshop at the 13th Annual ACM/IEEE International Conference on Human-Robot Interaction.

Simpson, G., Johnston, L., & Richardson, M. (2003). An investigation of road crossing in a virtual environment. Accident Analysis and Prevention, 35(5), 787–796. https://doi.org/10.1016/S0001-4575(02)00081-7

Sirois, S., & Brisson, J. (2014). Pupillometry. Wiley Interdisciplinary Reviews: Cognitive Science, 5(6), 679–692. https://doi.org/10.1002/wcs.1323

Sommerville, I. (2019). Software Engineering, Global Edition.

Stahl, D. O., & Wilson, P. W. (1995). On players′ models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218-254.

Strauch, C., Mühl, K., Patro, K., Grabmaier, C., Reithinger, S., Baumann, M., & Huckauf, A. (2019). Real autonomous driving from a passenger's perspective: Two experimental investigations using gaze behaviour and trust ratings in field and simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 15–28. https://doi.org/10.1016/j.trf.2019.08.013

Svärd, M., Markkula, G., Bärgman, J., & Victor, T. (2020). Computational modeling of driver pre-crash brake response, with and without off-road glances: Parameterization using real-world crashes and near-crashes 10.31234/osf.io/6nkgv

Tabone, W., De Winter, J. C. F., Ackermann, C., Bärgman, J., Baumann, M., Deb, S., Emmenegger, C., Habibovic, A., Hagenzieker, M., Hancock. P. A., Happee, R., Krems, J., Lee, J. D., Martens, M., Merat, N., Norman, D. A., Sheridan, T. B.,& Stanton, N. A. (2021). Vulnerable road users and the coming wave of automated vehicles: Expert perspectives. Transportation Research Interdisciplinary Perspectives, 9, 100293.

Talebpour, A., & Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transportation Research Part C: Emerging Technologies, 71, 143-163. https://doi.org/10.1016/j.trc.2016.07.007

Tenhundfeld, N, L., et al. (2019). Trust and Distrust of Automated Parking in a Tesla Model X. *Human Factors* Vol. XX, No. X, Month XXXX, pp. 1–17. DOI: 10.1177/0018720819865412

Teknomo, K. (2006). Application of microscopic pedestrian simulation model. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(1), 15-27.

Teknomo, K., Takeyama, Y., & Inamura, H. (2016). Review on microscopic pedestrian simulation model. *arXiv preprint arXiv:1609.01808*.

That, T. N., & Casas, J. (2011). An integrated framework combining a traffic simulator and a driving simulator. *Procedia-Social and Behavioral Sciences*, *20*, 648-655.

Treiber, Martin, Ansgar Hennecke, and Dirk Helbing. "Congested traffic states in empirical observations and microscopic simulations." Physical review E 62, no. 2 (2000): 1805.

Toledo, T., Koutsopoulos, H. N., & Ben-Akiva, M. E. (2003). Modeling integrated lane-changing behavior. Transportation Research Record, 1857(1), 30-38.

Václav, L., Petr, Z., & Aleš, Z. (2017). The Use of Driving Simulators in Psychological Research. Sibirskiy Psikhologicheskiy Zhurnal, 64, 65–75. https://doi.org/10.17223/17267080/64/4

Van Der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research Part C: Emerging Technologies, 5(1), 1–10. https://doi.org/10.1016/s0968-090x(96)00025-3

van Nes, N., Bärgman, J., Christoph, M., & van Schagen, I. (2019). The potential of naturalistic driving for in-depth understanding of driver behavior: UDRIVE results and beyond. Safety Science, 119, 11-20. https://doi.org/10.1016/j.ssci.2018.12.029

van Schagen, I., & Sagberg, F. (2012). The Potential Benefits of Naturalistic Driving for Road Safety Research: Theoretical and Empirical Considerations and Challenges for the Future. Procedia - Social and Behavioral Sciences, 48, 692–701. https://doi.org/10.1016/j.sbspro.2012.06.1047

van Schagen, I., Welsh, R., Backer-Grondahl, A., Hoedemaeker, M., Lotan, T., Morris, A., ... & Winkelbauer, M. (2011). Towards a large scale European Naturalistic Driving study: final report of PROLOGUE: deliverable D4. 2.

van Wageningen-Kessels, F., Hoogendoorn, S. P., Vuik, K., & van Lint, H. (2015). Traffic Flow Modeling: a Genealogy. Transportation Research Circular. N E-C195. April, 1.

Vechione, M., Balal, E., & Cheu, R. L. (2018). Comparisons of mandatory and discretionary lane changing behavior on freeways. International journal of transportation science and technology, 7(2), 124-136. https://doi.org/10.1016/j.ijtst.2018.02.002

Vogelsang, A., Eckhardt, J., Mendez, D., & Berger, M. (2020). Views on quality requirements in academia and practice: commonalities, differences, and context-dependent grey areas. *Information and Software Technology*, *121*, 106253.

Volz, B., Mielenz, H., Agamennoni, G., & Siegwart, R. (2015). Feature Relevance Estimation for Learning Pedestrian Behavior at Crosswalks. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC. https://doi.org/10.1109/ITSC.2015.144

Walker, F., Martens, M., Dey, D., Pfleging, B., Eggen, B., & Terken, J. (2019). Feeling-of-safety slider: Measuring pedestrian willingness to cross roads in field interactions with vehicles. Conference on Human Factors in Computing Systems - Proceedings, 1–6. https://doi.org/10.1145/3290607.3312880

Wang, J., Li, K., & Lu, X. (2014). Effect of Human Factors on Driver Behavior. *Advances in* Intelligent Vehicles, 111-157. doi:10.1016/b978-0-12-397199-9.00005-7

Wang, L., Ouyang, W., Wang, X., & Lu, H. (2015). Visual tracking with fully convolutional networks. Proceedings of the IEEE International Conference on Computer Vision. https://doi.org/10.1109/ICCV.2015.357

Wang, P., Sibi, S., Mok, B., & Ju, W. (2017). Marionette: Enabling On-Road Wizardof-Oz Automated Driving Studies. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (pp. 234-243). New York: ACM

Weinbeer, V., Muhr, T., Bengler, K., Baur, C., Radlmayr, J., & Bill, J. (2017). Highly automated driving: How to get the driver drowsy and how does drowsiness influence various take-over aspects?

Wiegers, K., & Beatty, J. (2013). *Software requirements*. Pearson Education.

Wiegers, (2003) Software Requirements: Practical Techniques for Gatheringand Managing Requirements Throughout the Product Development Cycle

Wilde, G. J. (1976). Social interaction patterns in driver behavior: An introductory review. *Human factors, 18*(5), 477-492.

Willig, C. (2008). Introducing Qualitative Research in Psychology. Open University Press. https://ocw.upj.ac.id/files/Textbook-PSI-308-Introducing-Qualitative-Research-in-Psychology.pdf

Wiedemann, R. (1974). Simulation des Strassenverkehrsflusses. Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe, Band 8, Karlsruhe, Germany.

World Health Organization. (2015). *Global status report on road safety 2015*. World Health Organization.

Wright, J. R., & Leyton-Brown, K. (2017). Predicting human behavior in unrepeated, simultaneous-move games. *Games and Economic Behavior*, *106*, 16-37.

Xia, Y., Wong, S. C., & Shu, C. W. (2009). Dynamic continuum pedestrian flow model with memory effect. *Physical Review E*, *79*(6), 066113.

Xue, H., Huynh, D. Q., & Reynolds, M. (2018). SS-LSTM: A Hierarchical LSTM Model for Pedestrian Trajectory Prediction. Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018. https://doi.org/10.1109/WACV.2018.00135

Yongji Luo, Bin Jia, Jun Liu, W. H. K. L. X. L. and Z. G. (2013). Modeling the interactions between car and bicycle in heterogeneous traffic. *Journal of Advanced Transportation*.

Yusof, N. M., Karjanto, J., Terken, J., Delbressine, F., Hassan, M. Z., & Rauterberg, M. (2016). The exploration of autonomous vehicle driving styles: Preferred longitudinal, lateral, and vertical accelerations. AutomotiveUI 2016 - 8th International Conference on Automotive

User Interfaces and Interactive Vehicular Applications, Proceedings, 245–252. https://doi.org/10.1145/3003715.3005455

Zander, T. O., Andreessen, L. M., Berg, A., Bleuel, M., Pawlitzki, J., Zawallich, L., Krol, L. R., & Gramann, K. (2017). Evaluation of a Dry EEG System for Application of Passive Brain-Computer Interfaces in Autonomous Driving. Frontiers in Human Neuroscience, 11, 1–16. https://doi.org/10.3389/fnhum.2017.00078

Zgonnikov, A., Abbink, D., & Markkula, G. (2020). Should I stay or should I go? Evidence accumulation drives decision making in human drivers. 10.31234/osf.io/p8dxn

Zhang, H., Su, Y., Peng, L., & Yao, D. (2010). A review of game theory applications in transportation analysis. *Proceedings of ICCIA 2010 - 2010 International Conference on Computer and Information Application*, 152–157. https://doi.org/10.1109/ICCIA.2010.6141559

Zhao, D., Wang, W., Li, C., Li, Z., Fu, P., & Hu, X. (2013). Modeling of passing events in mixed bicycle traffic with cellular automata. *Transportation Research Record*, *2387*, 26–34. https://doi.org/10.3141/2387-04

Zheng, J., Suzuki, K., & Fujita, M. (2012). Evaluation of Car-following Models Using Trajectory Data from Real Traffic. *Procedia - Social and Behavioral Sciences*, *43*, 356–366. https://doi.org/10.1016/j.sbspro.2012.04.109

Zheng, Z. (2014). Recent developments and research needs in modeling lane changing. *Transportation Research Part B: Methodological*, *60*, 16–32. https://doi.org/10.1016/j.trb.2013.11.009