The Social Acceptance of Automated Driving Systems: Safety Aspects

A contribution to responsible innovation by using a referendum format, discrete choice model experiment to measure the social acceptance of ADS by Dutch citizens with corresponding heterogeneity

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in **Complex Systems Engineering and Management**

Faculty of Technology, Policy and Management

by

Bart Overakker

Student number: 4015630

To be defended in public on October 12th 2017

Graduation Committee

Chairperson	: Prof. dr. ir. C. G. Chorus, Section TLO
First Supervisor	: Dr. E. J. E. Molin, Section TLO
Second Supervisor	: Dr. F. Santoni de Sio, Section Ethics & Philosophy
External Supervisor	: Ir. J. van der Waard, KiM

Contents

Acronyms

ADS	Automated Driving System
AV	Automated Vehicle
VOT	Value of Time (or Value of Travel Time Savings)
VMT	Vehicle-Miles Travelled
WTP	Willingness to Pay

Executive Summary

Automated driving systems (ADS) can improve traffic safety, improve accessibility and reduce environmental impact (Shladover, 2016). On the contrary, on May 7^{th} 2016, a fatal accident with a Tesla on autopilot in U.S. Florida was a harsh reminder that the technology is still in its testing phase (Greenemeier, 2016). In complex technical systems like ADS technical failure may occur, which forms a serious threat to human well-being. Moreover, studies have shown that citizens are very concerned about deliberate misuse of ADS (Kyriakidis, Happee, & De Winter, 2015), e.g. people purposely abusing ADS to cause damage or even hurt someone. Therefore, according to many experts, the implementation of ADS does not only entail technical issues but also normative issues.

To bridge the gap between technical- and normative issues, responsible innovation can be applied (Santoni de Sio, 2016). Friedman et. al. (2006) propose value-sensitive design to achieve responsible innovation via technical-, empirical- and conceptual research. As following from responsible innovation and value-sensitive design and in particular its focus on empirical research as one of its necessary elements, in this research an effort is made to use empirical research methods to provide more insights in normative issues of automated driving systems (ADS). The focus lies on social acceptance, particularly with respect to traffic safety. Also accessibility, environmental impact and heterogeneity among citizens will be analysed. Social acceptance is defined as "a person's assent to the reality of a situation, recognizing a process or condition (often a negative or uncomfortable situation) without attempting to change it, protest, or exit" (Fish, 2014, page 1). The following questions will be answered in this research:

What is the social acceptance of automated driving systems from the perspective of safety, accessibility and environmental impact and what is the corresponding heterogeneity?

1. *What percentage of citizens thinks automated driving systems are socially accepted?*
2. *How is the social acceptance influenced by safety, accessibility and environmental impact?*
3. *Are traffic fatalities caused by automated vehicles valued differently than current traffic fatalities?*
4. *Is there heterogeneity in the social acceptance among citizens?*

A survey is chosen as research method since it is a relatively inexpensive, flexible method to achieve extensive information about characteristics of a population. After an extensive theoretical analysis, seven attributes were identified that possibly influence the social acceptance: level of automation, road exemption, travel time, emissions, human error fatalities, technical failure fatalities and deliberate misuse fatalities. After the experiment was fine-tuned by a pilot study, it was held among a representative sample of 510 Dutch adults during the spring of 2017. The respondents had to state if they were in favour or against ADS for each of the twelve hypothetical futures that were presented to them. In these hypothetical futures, the attributes were systematically varied and described as a change to the current situation.

Using a MNL RUM model, the results show that 63% of all citizens prefer ADS over the current system. It is therefore concluded that citizens have a high social acceptance and thus are rather positive towards ADS. Also, citizens prefer a system where human drivers are still in control and can intervene in case necessary.

Next, it is concluded that the social acceptance is mostly influenced by fatalities caused by automated vehicles (AVs), while travel time is the least important attribute. However, the differences in influence of the attributes were not substantial. Safety, accessibility and environmental impact are all important for the social acceptance. Nevertheless, technical failure fatalities weigh as much as 4 human error fatalities. For deliberate misuse fatalities this is a factor 5.5. Although these relations coincide with literature, the magnitude is larger than expected. It implies that ADS have to be very safe in order to reach social acceptance. Since AVs are still 'learning' how to drive, this might cause problems for current and future experiments.

A latent class choice model is estimated to answer the final research question. Results show that large heterogeneity exists among citizens in the social acceptance. Citizens can be segmented into three classes (% of citizens): automated driving enthusiasts (32%), central mass (52%) and risk-averse class (16%). Contradictory to average citizens, automated driving enthusiasts prefer high automation levels. Even so, they still weigh fatalities caused by AVs as much as 3 human error fatalities on average. The central mass shows similar results to the results of the MNL model estimated on the full sample. The risk-averse class has a strong dislike for fatalities caused by AVs. This class weighs technical failure fatalities (deliberate misuse fatalities) as much as 5.5 (10) human error fatalities.

In conclusion, primarily two discrepancies are identified that are critical for the implementation of ADS: 1) High social acceptance versus strong dislike for fatalities caused by AVs; 2) Citizens who are enthusiastic about ADS versus citizens who are risk-averse. They lead to the following recommendations:

The social acceptance for ADS is high, so it is recommended for policy makers to have a positive and active approach towards ADS. By conducting experiments for professional users, safety risks can be minimalized while a learning curve is ensured. Also technology producers and policy makers should intensify research into cooperate driving. According to experts, ADS and cooperate driving are inseparable (Shladover, 2016), but globally the research into cooperate driving is lacking (Roland Berger, 2017). Since cooperate driving can lead to an increased risk of deliberate misuse, it is deemed critical for the implementation of ADS. Next, policy makers and especially the RDW should review the licensing of AVs. Currently, hardly any restrictions are in place for the licensing of AVs, which can cause dangerous situations on public roads. Finally, information campaigns can help to make citizens aware of the risks and benefits of ADS.

1. Introduction

Automated vehicles (or self-driving vehicles, driverless cars and robotic cars) get a lot of attention by policy makers, technology producers and society in general. The reason why this new technology is so interesting has much to do with the performance of the current Dutch transport system. A couple of examples: 1) In 2016, there were 629 traffic fatalities in the Netherlands (SWOV, 2017b). For two consecutive years the number of fatalities has increased. Between the age of 15 and 24, it is the second biggest death cause after suicide (CBS.nl, 2017); 2) In 2016, congestion has increased by 10 per cent compared to 2015 (Verkeersinformatiedienst, 2016). Annual financial damage for the Dutch economy is estimated to be 1,1 billion euro (TNO, 2016); 3) Private cars are responsible for 12 per cent of total EU emissions of CO_2, the main greenhouse gas (European Commission, 2015). Air pollution due to traffic causes an economical damage of approximately 4 billion euro (Milieudefensie, n.d.).

Automated vehicles (AVs) can improve traffic safety, improve accessibility and reduce environmental impact (Shladover, 2016). One might say that they are the perfect solution to all above-mentioned problems. That is partly why technology is rapidly improving; Elon Musk: "I really consider automated driving a solved problem; I think we are probably less than two years away" (The Guardian, 2016; page 1). GM president Dan Ammann and Nissan CEO Carlos Ghosn say they will be shipping AVs by 2020 (The Guardian, 2016). Audi is bound to release their new A8 with automated driving capabilities in December 2017 (nu.nl, 2017). Today's cars are already partly automated: power steering, (adaptive) cruise control, lane keeping assist et cetera (Anderson et al., 2014). Also more advanced automation exists in the Netherlands: trucks driving with short headways called Truck Platooning (Bakermans, 2016); the automated shuttle bus in Rotterdam (Oomen, 2005); and the WEpods in Ede-Wageningen (WEpods.nl, n.d.). All these projects show great possibilities for the future.

On the contrary, on May 7th 2016, a fatal accident with a Tesla on autopilot in U.S. Florida was a harsh reminder that the technology is still in its testing phase (Greenemeier, 2016). Again on April 11th, 2017, an accident with a Tesla on autopilot occurred on the A1 highway in the Netherlands (Telegraaf, 2017). These accidents are not the first and will probably not be the last accidents caused by AVs. In complex technical systems like AVs technical failure may occur, which can form a serious threat to human well-being.


Figure 1: Autonomous Trap 001 (Bridle, 2017)

Moreover, studies have shown that people are very concerned about deliberate misuse of AVs (Kyriakidis et al., 2015), e.g. people purposely abusing AVs to cause damage or even hurt someone. An interesting illustration of the risk for deliberate misuse is a thought-provoking experiment by James Bridle in which an AV is trapped (see Figure 1). This scientist/artist has replicated road markings with salt which allow AVs to get into the trap, while they cannot get out. Since AVs are dependent on the information from their environment, people with bad intentions can purposely alter this environment to cause harm. Also, AVs are like driving computers which make them vulnerable for hacking (Loukas, 2015). Again, this could lead to unpredictable threats for human well-being.

Therefore, according to many experts, the implementation of AVs does not only entail technical issues but also normative issues. Technical issues are how to optimize design requirements to reach optimal social impacts; normative issues are the evaluation of these social impacts to norms and values. One example of normative issues is an experiment from MIT called the moral machine (Bonnefon, Shariff, & Rahwan, 2016), which is a modern version of the trolley problem (Foot, 1978). In this experiment, normative moral issues are shown to respondents where an AV must choose between two evils, such as killing two car passengers or five pedestrians. Respondents must choose which outcome is more acceptable.

The combination of technical- and normative issues cause for complex decision-making (Cuppen, 2012). This research will try to make sense of this complexity by using the framework of responsible innovation: methods that focus on including norms and values in the innovation process and the design of products. As following from responsible innovation and in particular its focus on empirical research as one of its necessary elements (Friedman, Kahn Jr., & Borning, 2006), an empirical research will be conducted to provide more insights in the social acceptance of automated driving systems. Social acceptance is defined as "a person's assent to the reality of a situation, recognizing a process or condition (often a negative or uncomfortable situation) without attempting to change it, protest, or exit" (Fish, 2014, page 1). Prominent in this research are discrepancies between technical- and normative issues. To provide insights in normative issues on a societal level, the social acceptance of citizens is measured, not only of potential AV-users. The focus will be on the following aspects:

Firstly the focus will be on safety, particularly on the cause of fatalities: human error fatalities, technical failure fatalities and deliberate misuse fatalities; How do fatalities caused by AVs influence the social acceptance? Are they valued differently than human error fatalities? Next to safety, a lot of research in the Netherlands is focused on overall efficiency of the transport system, e.g. improved accessibility and environmental impact. Although precise predictions are difficult to make, ministers of all 28 EU member states acknowledge in the Declaration of Amsterdam that AVs offer great potential to improve traffic flows, overall efficiency and environmental performance of the transport system (European Union, 2016). For some, traffic jams are a constant source of irritation and looming climate change has caused people to protest for new norms in environmental impact. These social impacts might therefore shape norms and values of citizens and influence the social acceptance.

Secondly, not only social impacts are important for the social acceptance, but also in which way AVs are implemented in the current system. To explore possible points of action for policy, the KiM Netherlands Institute for Transport Policy Analysis has devised a transition path to a future traffic- and transport system involving AVs, henceforth called an automated driving system (ADS). As the transition path of the KiM shows, different system designs are possible. Among others, these designs differ in automation levels of AVs and type of roads where AVs are deployed. Different designs lead to different impacts in safety, accessibility and environmental impacts, but also to other social impacts like privacy and joy of driving. It is therefore expected that the social acceptance is influenced by the design of ADS.

Finally, this research will try to measure heterogeneity in the social acceptance of ADS. According to the literature review on behavioural experiments for AVs by Becker and Axhausen, all 16 reviewed publications showed great heterogeneity in their sample in factors related to AVs (Becker & Axhausen, 2016). Although these publications mostly focused on consumer preferences, this research will determine if similar heterogeneity exists in the social acceptance of citizens.

Given the empirical nature of this research the focus will be on the transition phase, approximately the coming 10 to 15 years. Since it is assumed that fully automated driving is not realistic during this period, this research is restricted to automation level 3 and 4 as defined by SAE international (SAE International, 2014). For both levels the AV can drive itself; the main difference is that human drivers should be able to intervene in dangerous situations for level 3, while the human driver may conduct other tasks than driving for level 4. Also, during the transition phase low adoption rates of below 40 per cent are assumed. In this research, automated driving systems are thus defined as road traffic- and transport systems with relatively low percentage rates of level 3 and level 4 automated vehicles. The system herein is the socio-technical design in which AVs are embedded, i.e. the integration of the technical design with the institutional design.

1.1. Problem Statement

AVs can possibly resolve important social problems and technologies are improving rapidly. It demands policy decisions if-, how- and when AVs are on Dutch roads. The declaration of Amsterdam is proof of the urge for decision-making (European Union, 2016).

However, the decision-making is not straightforward. Not only do decision-makers (policy makers, technology producers) encounter numerous technical uncertainties, their decisions also have a big social impact. Decision-making for the implementation of ADS is described as a complex problem or wicked problem: problems that involve high social stakes and scientific uncertainties (Cuppen, 2012). "Complex issues concern a tangled web of related problems (multi-problem), lie across or at the intersection of many disciplines (multi-disciplinary) and the underlying processes interact on various scale levels and on different temporal scales (multi-scale)" (van Asselt Marjolein & Rijkens-klomp, 2002; page 168). Implementing ADS has all these characteristics: multi-problem (e.g. safety, accessibility and environmental impact), multi-disciplinary (e.g. technical and normative) and multi-scale (e.g. geographical, temporal and multi-actor). With these kinds of problems a technocratic view is too narrow. The decision-making process needs different kinds of knowledge, expertise and values (Cuppen, 2012).

To encounter some of these problems, Santoni de Sio recommends responsible innovation as a framework "...to prevent a situation where there is a disconnect between abstract moral discussions and the real world of engineering and policy" (Santoni de Sio, 2016; page 6). Responsible innovation is a general term for methods that focus on including norms and values in the innovation process and the design of products. It requires that all stakeholders including civil society are responsive to each other and take shared responsibility for processes and outcomes of research and innovation (Presidency of the Council of the European Union, 2014). Friedman et. al. (2006) propose to achieve responsible innovation via value-sensitive design, which goal is to integrate norms and values into technical systems and via technical-, empirical- and conceptual research.

To elaborate on the latter, responsible innovation and value-sensitive design require sufficient empirical research to include stakeholders' opinions, among others, citizens (Cuppen, 2012). Given the eminent social impacts, decision-makers must have knowledge of the social acceptance to ensure that norms and values of citizens are integrated into the design. Moreover, inclusion of citizens' values and opinions can enrich the decision-making of ADS.

1.2. Knowledge Gap

In literature, normative issues are described from two perspectives: 1) Conceptual research: moral issues raised and questioned by ethics scholars and 2) Empirical research: public's opinion measured and analysed by empirical social scientists. Ethics scholars focus on questions such as "how to prevent the transition to automated driving from negatively affect values like human accountability as well as the individual rights to life, physical integrity and privacy?" (Santoni de Sio, 2016) and "if an AV must choose in a split-second between killing two car passengers or five pedestrians, what should it choose?" (Bonnefon et al., 2016). These questions are addressed via conceptual analyses, normative reasoning, and interpretation of the existing evidence.

Empirical research is mostly based on consumer preferences and adoption levels. Some examples are: Satisfaction of in-vehicle technology (Abraham et al., 2017), AV adoption (Lavieri et al., 2017), willingness-to-pay (WTP) (Becker & Axhausen, 2016), driving behaviour (Jamson, Merat, Carsten, & Lai, 2013) and relocation patterns with ADS (Lavasani, Asgari, Jin, & Pinjari, 2017). The human subjects in these experiments are categorized as consumers rather than citizens. In other words, they are asked questions as if they are users of the technology or want to become users in the future.

This research contributes to empirical literature by being the first to measure the social acceptance primarily related to significant changes in traffic safety levels. Attention goes to three safety related variables: human error fatalities, technical failure fatalities and deliberate misuse fatalities. Particularly the focus on citizens, rather than consumers, is deemed both unique and interesting.

1.3. Objective

Primarily, the objective is to use empirical research methods to measure the social acceptance of automated driving systems from the perspective of safety. Also, influence of accessibility and environmental impact on the social acceptance is measured, including heterogeneity among citizens. The social acceptance can be used to withdraw insights in technical- and normative issues in the implementation of ADS.

The ultimate objective of this research is to use social acceptance to contribute to responsible innovation for ADS. Having insights in technical- and normative issues, this research will make recommendations for decision-makers to explore new possibilities and/or improve their work on ADS. Also, an actor analysis is conducted to explore their interrelations and direct recommendations to specific actors.

1.4. Research Question

The antecedent analysis has led to the following research questions:

What is the social acceptance of automated driving systems from the perspective of safety, accessibility and environmental impact and what is the corresponding heterogeneity?

1. *What percentage of citizens thinks automated driving systems are socially accepted?*
2. *How is the social acceptance influenced by safety, accessibility and environmental impact?*
3. *Are traffic fatalities caused by automated vehicles valued differently than current traffic fatalities?*
4. *Is there heterogeneity in the social acceptance among citizens?*

2. Theory

This chapter will explain theories and assumptions that set the foundation for this research. As introduced, the framework for this research is responsible innovation which is explained and motivated in the first paragraph. §2.2 will define the conceptual model of social acceptance, followed by a brief case study on the rise of the automobile to make responsible innovation and the social acceptance more tangible. In §2.4 an analysis is conducted of possible future automated driving systems to explore state-of-the-art research in technical- and normative issues. §2.5 will conduct an actor analysis to formulate directed recommendations for researchers, policy makers and engineers. Finally, §2.6 will summarize this theory section.

2.1. Responsible Innovation

Multiple definitions of responsible innovation are found in literature, like the examples below:

> *"The concept of innovation pertains both to the introduction of new products, processes and services and to organisational and societal renewal. This programme description defines innovation primarily as the use of application of the results of science and technology. Responsible innovation concerns research, development and design, and takes societal values, interests, needs, rights and welfare into consideration" (NWO, 2015).*
>
> *"Responsible research and innovation is an approach that anticipates and assesses potential implications and societal expectations with regard to research and innovation, with the aim to foster the design of inclusive and sustainable research and innovation" (EuroScientist, 2017).*
>
> *"Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society)" (von Schomberg, 2011).*
>
> *"Responsible Innovation is an activity or process which may give rise to previously unknown designs pertaining either to the physical world (e.g., designs of buildings and infrastructure), the conceptual world (e.g., conceptual frameworks, mathematics, logic, theory, software), the institutional world (social and legal institutions, procedures, and organization) or combinations of these, which - when implemented - expand the set of relevant feasible options regarding solving a set of moral problems" (van den Hoven, n.d.).*

Most definitions have common aspects: Firstly, responsible innovation is defined as an on-going process. It is not something that is applied at a certain moment in time, but rather continuously along the development path of the technology. For example, research and development should also be included in responsible innovation. Secondly, the definitions require norms and values to be integrated in innovations: societal values, interests, needs, rights, ethical acceptability, desirability et cetera.

Friedman et. al. (2006) propose to achieve responsible innovation via value-sensitive design. In this research, this methodology is used to explain social acceptance and place it within responsible innovation. In this methodology three types of research are identified: technical-, empirical- and conceptual research (see Figure 2).

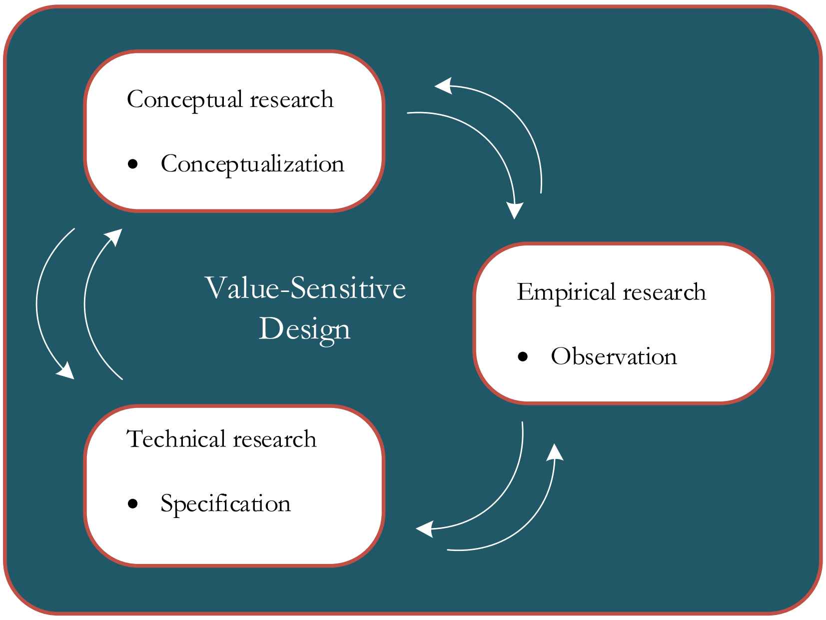


Figure 2: Value-sensitive design

Technical research is based on the specification of a system. It calculates and specifies design requirements that are needed to achieve certain goals. To meet the criteria of value-sensitive design, technical research should focus on the embodiment of norms and values into the technical design (Friedman et al., 2006). Norms and values can be retrieved from empirical and conceptual research.

Empirical research is a discipline that makes use of observation. Multiple methods can be used like interviews, surveys or case studies. Commonly used topics in transport are value-of-time (VOT) and willingness-to-pay (WTP). The topic of this research is social acceptance: "a person's assent to the reality of a situation, recognizing a process or condition (often a negative or uncomfortable situation) without attempting to change it, protest, or exit" (Fish, 2014). It relates to behaviour and is difficult to observe and even more difficult to predict, e.g. when is a new technology accepted? Social acceptance is also subjective: Individuals commonly have different information on a subject and different norms and values, so individuals will behave differently in similar situations. This research measures a status-quo of the social acceptance. Preferably, the social acceptance is measured frequently during the development of ADS to comply with the continuous nature of responsible innovation.

The gap between empirical- and conceptual research is the difference between observation and theory (Friedman et al., 2006). Conceptual research is typical for ethics scholars, who address moral issues via conceptual analyses, normative reasoning, and interpretation of the existing evidence. So why bother with conceptual research if a technology is already accepted? There are a couple of reasons why the social acceptance may fall short: "1) Acceptance may be based on wrong information; 2) Acceptance may be based on wrong reasons or values; 3) People may have no choice and; 4) Important parties may have no voice" (Poel, 2016). Therefore it is argued to also include ethical acceptability in value-sensitive design: "A reflection on a new technology that takes into account the moral issues that emerge from its introduction" (Taebi, 2016; page 2). Ethical acceptability can overcome the shortcomings of the social acceptance.

Hence, the goal of responsible innovation and value-sensitive design is to integrate norms and values into technical systems and via technical-, empirical- and conceptual research. The next paragraph will identify a conceptual model of the social acceptance.

2.2. Conceptual Model of the Social Acceptance

The aim of this paragraph is to find a conceptual model of the social acceptance. From the antecedent analyses, three criteria are extracted that need to be included in the conceptual model: 1) The conceptual model includes all citizens, not only users; 2) The social acceptance is subjective. Therefore, the conceptual model needs to allow individuals to have different information on the subject and different norms and beliefs; 3) The model should allow weighing the different factors that influence the social acceptance.

A number of conceptual models that include social acceptance in one way or another are found in literature. The *technology acceptance model* is an information systems theory by Davis on how users come to accept and use a technology (Davis, 1989). This model is based on users including factors like 'perceived usefulness' and 'perceived ease of use' and not on citizens. The *theory of planned behaviour* and *theory of reasoned action* both include factors related to the social acceptance as an independent variable for behaviour rather than a dependent variable (Madden, Ellen, & Ajzen, 1992). These and other conceptual models did not serve the purpose that is intended for this research.

The search led to the conceptual model for spatial behaviour of Timmermans (1980) shown in Figure 3. Although this model does not include the social acceptance, it does include the three criteria mentioned above. The model assumes that individuals make their decision in three steps: Firstly, individuals gather all the information that is known to them. It is not realistic to assume that individuals know all the complex factors that involve ADS. The physical environment of ADS is therefore reduced to the cognitive environment of ADS of a certain individual.

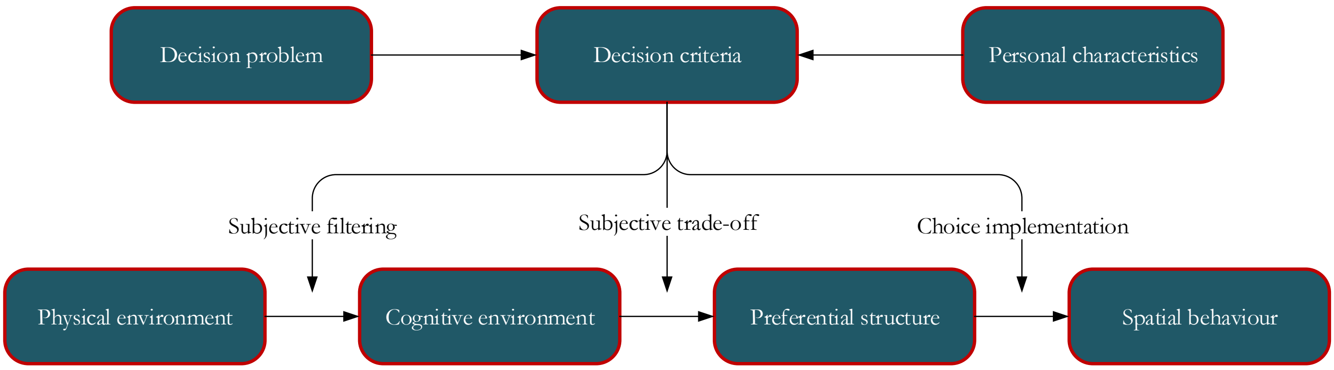


Figure 3: Generic conceptual model spatial behaviour (Timmermans, 1980)

Secondly, an individual makes a subjective trade-off of all the information known to him/her and weighs all the impacts to define his/her opinion of all alternatives (Timmermans, 1980). The result is a preferential structure: a picking order of all possible alternatives. In the third and final step of the thought process, an individual makes its actual choice. He/she applies a decision rule to this picking order of alternatives. The three steps in the though process are influences by decision criteria, which are based on the decision problem at hand at personal characteristics of the individual (Timmermans, 1980).

The result in the conceptual model of Timmermans (1980) is the spatial behaviour of individuals. However, since all three criteria are present in this model, it is assumed that this model can also be used for the social acceptance. Hence, it is assumed that citizens will gather the information they know about ADS, make a subjective trade-off of all impacts and finally make the choice if ADS is socially accepted.

The next paragraph will try to make responsible innovation and the social acceptance more tangible by conducting a case study about the rise of the automobile.

2.3. The Rise of the Automobile

In the Netherlands, the rise of the automobile started in 1896 and forty years later there were approximately 100.000 cars. Assumedly, the first car related fatal accident in the Netherlands happened on February 22nd 1899: An automobile frightened a crossing horse and carriage, after which the rider died in result of a fall (Bos, Groningen, Mom, & Vinne, 1996). Between 1905 and 1907 a number of fatal car accidents occurred that strongly drew the attention of the public (Vinne, 2007). Official statistics of traffic fatalities only started in 1926 when there were 518 fatalities. A traffic fatality is internationally defined as "someone that in effect of an accident on the public road, in which at least one vehicle is concerned, dies within thirty days as a result of that accident" (SWOV, 2016). As car usage grew exponentially after World War I, the number of traffic fatalities grew as well resulting in more than 3000 traffic fatalities annually in the seventies (see Figure 4).

In the period 1900 to 1945, the Dutch government saw traffic safety as a local problem that did not need central coordination (Bax, 2012). The Dutch organisation ANWB was first to raise public awareness to traffic safety, initially from the perspective of cyclists. The main idea at that time was that car drivers were to blame and irresponsible. Interestingly it was not the government, but societal pressure that put traffic safety on the political agenda (Bax, 2012).

In the U.S.A. the number of traffic fatalities also grew with the rise of the automobile. "In the first four years after World War I, more Americans died in auto accidents than had been killed during battle in Europe" (Oatman-Stanford, 2014; page 2). Also in the U.S.A. people would protest and blame car drivers for the lack of traffic safety and the government was slow to react on it. Since there were no significant protests against- or limitations to car manufacturers, they produced numerous cars. The most famous being Ford's Model T which sold in millions (Oatman-Stanford, 2014).

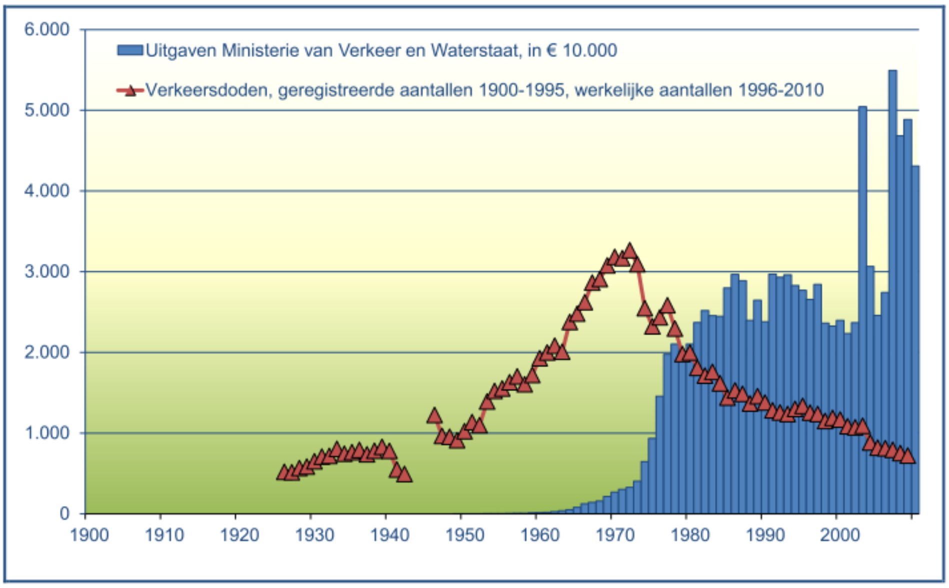


Figure 4: Traffic Fatalities and Expenditures in Traffic Safety of the Ministry of Transport, Public Works and Water Management (Bax, 2012)

As shown in Figure 4, the expenditures on traffic safety of the Dutch central government grew significantly in the seventies. The result was a steady decline in number of fatalities. In 2016, there were 629 traffic fatalities in the Netherlands (SWOV, 2017b). Over 70 per cent of all fatal accidents involved a car, so approximately 440 fatalities. Furthermore, approximately 50 per cent of these 440 fatalities are not car drivers, but pedestrians, cyclists or other travellers. In the Netherlands, between the age of 15 and 24, traffic accidents are the second biggest death cause, after suicide (CBS.nl, 2017).

Other criteria for traffic safety are number of (severe) injuries and fatalities per billion km. For the latter, the car is actually a very safe road modality (see Table 1). In the table, a dash means that there are no records available and examples of other modalities are busses, trains and vehicles for the disabled.

Table 1: Traffic Fatalities and Hospitalizations in 2015 in the Netherlands (SWOV, 2013, 2017b)

Modality	Fatalities in 2016 [#]	Car involved with fatal accidents [%]	Fatalities with car involved [#]	Hospitalized [#]	Fatalities per billion km [#]
Car	231	100	231	2835	2
Bicycle	189	75	142	1960	14
Pedestrian	51	57	29	148	16
Motorcycle/Moped	89	-	-	2791	74
Other modalities	69	-	-	5593	-
Total	629		402	13327	

In hindsight, did responsible innovation apply to the rise of the automobile? Well, the rise of the automobile was big step forward from a technological perspective. Research was based on innovation of the automobile which was very successful. People were able to move faster to further destinations.

However, partly because of the passive approach by governments, it resulted in unsafe traffic situations worldwide. From the societal pressure peaking in the seventies it can be argued that the amount of traffic fatalities was not accepted. Traffic safety did improve significantly in the Netherlands since the seventies and stagnated at just over 600 fatalities for the last few years. Currently, the Netherlands has one of the highest road safety levels worldwide (SWOV, 2017a). The current number of fatalities seems to have become the new norm. Although effort is being made to make the Dutch roads safer, there are no major protests by society. It is therefore argued that the current transport system has a neutral social acceptance.

From an international perspective, the Dutch safety record has made good progress which took years of research and effort by many, but one might still question if more than 600 traffic fatalities a year is ethically acceptable. It has been convincingly argued that it is simply morally unacceptable to die while using the transport systems and that the system designers have the moral responsibility to prevent the realization of (fatal) accidents (Nihlén Fahlquist, 2006). It is possible that people accept traffic fatalities because they got used to them, which does not necessarily make it acceptable. It may also imply that there is a moral obligation for policy makers to promote the implementation of ADS if they are that much safer than the current transport system (Santoni de Sio, 2016).

This brief history of the rise of the automobile illustrates the lack of responsible innovation, especially in the early beginning of the automobile. The automobile was well embedded in every day's life while norms and values were not embedded in the traffic system. The number of fatalities reached high above the norms of being socially accepted and being ethically acceptable. It also illustrates that the rise of the automobile was a multi-actor problem: car manufacturers, (central) government, citizens and many more, which made the implementation more complex.

For the implementation of ADS extensive research on current traffic safety is already available and significant governmental investments are being made. It is therefore unlikely that AVs will significantly deteriorate current safety records. Even so, ADS might introduce new unpredictable risks for human life, e.g. technical failure and deliberate misuse. Therefore, an effort must be made to prevent similar unaccepted safety levels for ADS.

2.4. Possible Futures of Automated Driving Systems

This paragraph will analyse automated driving systems to explore possible future system designs with their technical- and normative issues. The aim is to identify design factors and realistic levels for the social impacts that can be used in the empirical experiment. The conceptual framework of transport systems by van Van Wee, Annema, & Banister (2013) in Figure 5 is normally used for policy analysis. However, it provides a nice structured overview of transport systems and includes safety, accessibility and environment impacts. Therefore, this framework is used to structure this chapter.

The social impacts will be analysed first: 'safety' §2.4.1, 'accessibility' §2.4.2 and 'the environment' §2.4.3. The following three paragraphs will discuss design features that may influence the social acceptance: 'volume, composition of traffic and transport, division over time and space' §2.4.4, 'technology' §2.4.5 and 'way of using vehicles' §2.4.6. The remaining three factors in the model are categorized as intermediate factors and are discussed in the paragraph about accessibility. Each factor will have a different section for technical- and normative issues, although some overlap between the two is unavoidable. Even so, it helps to clearly identify possible tension between technical- and normative issues.

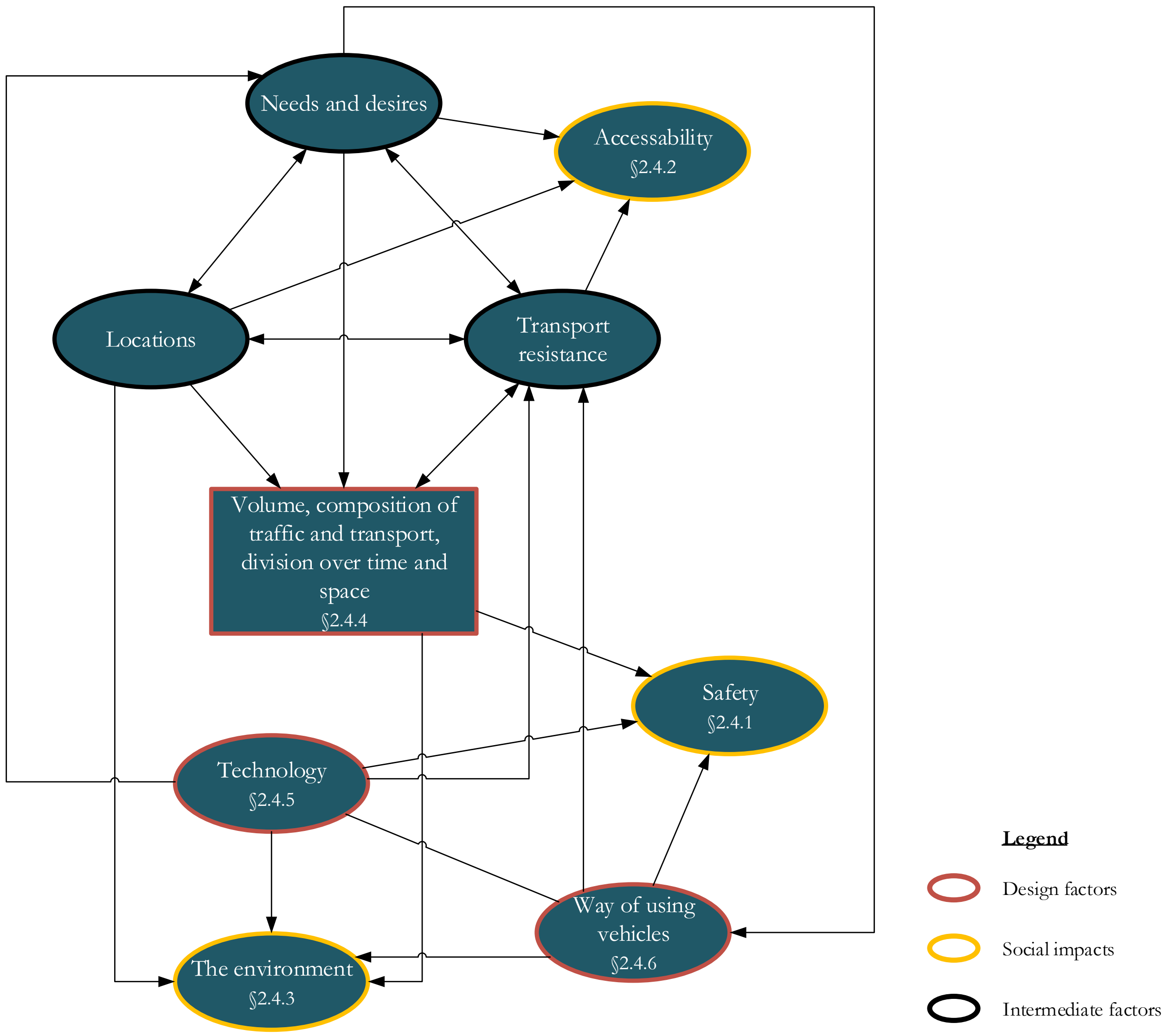


Figure 5: Conceptual framework for factors having an impact on transport volumes and the impact of the transport system on accessibility, the environment and safety (Van Wee et al., 2013)

2.4.1. Traffic Safety

Technical

On May 7th, 2016, a fatal accident with a Tesla on autopilot in U.S. Florida was a harsh reminder that the technology is still in its testing phase (Greenemeier, 2016). Again on April 11th, 2017, an accident with a Tesla on autopilot occurred on the A1 highway in the Netherlands (Telegraaf, 2017). Both accidents raised awareness by the public of the risks, very similar to the first accidents with the automobile.

On the other hand, on December 27th 2016, a video emerged that shows a Tesla that registered an accident ahead before it actually happened and applied the brakes accordingly (NOS, 2016). It is debatable whether the human driver had the capability to avoid the crashing cars. Nevertheless, it spoke to people's imagination on what this technology can bring, since the video got over half a million views on the website dumpert.nl alone.

The vehicle that currently has the best safety record is Waymo, formerly Google's AV project. This vehicle showed 124 discrepancies in software or hardware in 1.023.330 kilometres in 2016, a 19 per cent decrease since the year before (Davies, 2017). Although this is not a good measure for the safety performance of the technology, it does show that progress is being made. It is likely that AVs will be safer than human driven vehicles in the near future, especially when AVs would be operating under predefined conditions in a protected environment.

The future is unpredictable and so is the implementation of ADS. It is therefore unknown which effects AVs have on safety. It is not the goal of this research to determine the safety of AVs. It is assumed that there will become a point where ADS are safer than our current transport system. It is the goal to determine how much safer they need to be to reach social acceptance.

Normative

A first reason why the social acceptance of ADS might differ from the current system in terms of safety is because human error fatalities are likely to be replaced by fatalities caused by AVs. Technical failure is defined as software or hardware failure of an AV. If sensors have a misdetection or algorithms misinterpret the data, terrible accidents could happen. In regret theory, it is stated that losses loom larger than gains of equal magnitude (Chorus, 2017; Loomes & Sugden, 1983). A decrease in human error fatalities is a gain compared to the current system; an increase in fatalities caused by AVs is a loss compared to the current system. It is therefore possible that citizens think that fatalities caused by AVs are worse than human error fatalities.

Secondly, the vulnerability of the system for deliberate misuse could influence the social acceptance. Deliberate misuse is described as the risk of people using the system with bad intentions. Being able to hack into the car and use it for terrorist attacks is a grim prospect. AVs are like driving computers and are likely to become connected to other cars, infrastructure or a cloud. This will make them more vulnerable for hacking (Loukas, 2015).

Next, deliberate misuse fatalities could weigh heavier than human error- or technical failure fatalities because of the difference between safety and security. Safety is being free from unintentional danger (technical failure, human error) and security is being free from intentional attack like criminal activity (deliberate misuse). According to many experts, security weighs heavier than safety (Pearsall & Hanks, 2001). It is therefore possible that deliberate misuse fatalities are valued worse than technical failure fatalities and human error fatalities.

Fourthly, technical failure and deliberate misuse propose possible problems in responsibility. Who is to blame in a crash caused by automated driving, the manufacturer or the human driver? In Europe, a human driver should be fully operating the vehicle at any time (Vienna Convention, 1986). However, experts say this responsibility might shift to AV manufacturers in case automated driving is used (Anderson et al., 2014). Interestingly, both the driver and Tesla are held responsible for the abovementioned fatal crash in the U.S. Florida according to the investigators (NTSB, 2017). The responsibility gap might influence the valuation of technical failure- and deliberate misuse fatalities.

It is also hypothesized that there is heterogeneity in the valuation of fatalities among citizens. There might be citizens who prefer ADS when its absolute safety record is better than the current transport system, regardless the cause of fatalities. In Figure 6, this is labelled as high acceptance. On the other hand, there might be citizens that strongly penalize technical failure fatalities and deliberate misuse fatalities. They would only want ADS if there are hardly any fatalities caused by AVs, which is labelled as low acceptance.

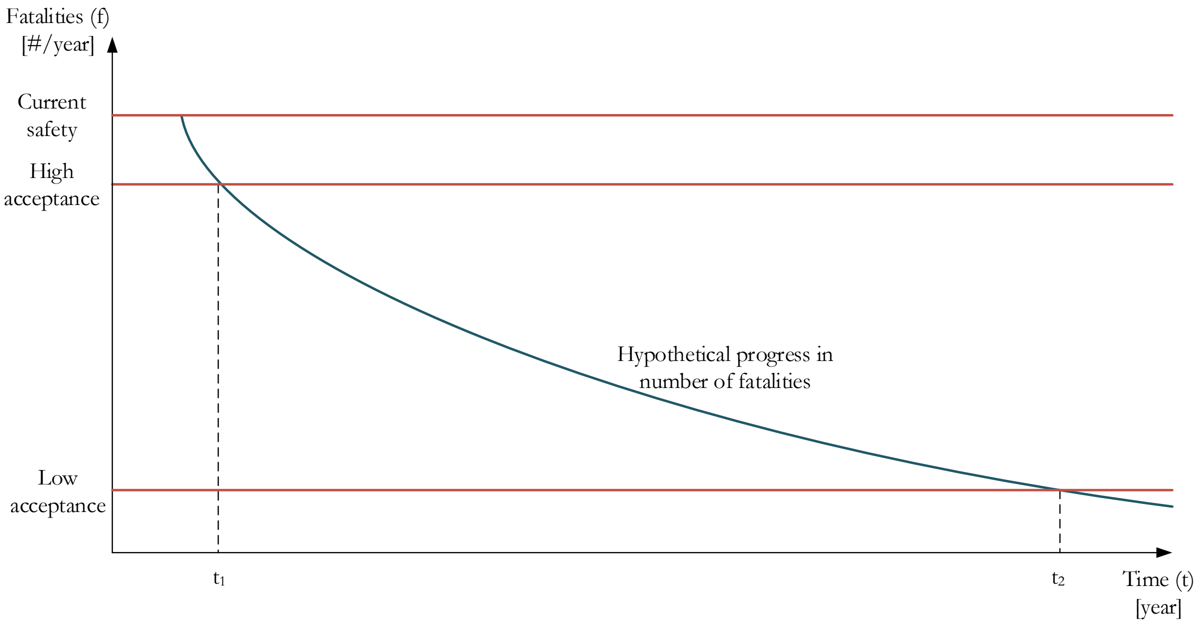


Figure 6: Difference in implementation time corresponding to high and low acceptance of traffic safety

If we would take the progress in discrepancies of the automated vehicle Waymo (an annual decrease of 19%) as a measurement of progress in number of fatalities of ADS and assume a difference of 500 fatalities between high- and low acceptance, the difference between t_1 and t_2 would still be approximately 15 years. In other words, the difference between high acceptance and low acceptance could mean a 15 year gap in the accepted moment in time to implement ADS. The social acceptance could therefore provide valuable insights for policy makers when to start implementation.

2.4.2. Accessibility

Technical

Accessibility is influenced by locations, transport resistance and needs & desires. These three factors are influenced by each other, but also by many other factors like way of using vehicles, technology and safety. This complexity makes the effect of AVs on accessibility difficult to assess. Numerous publications try to answer the question how it will change in the transition to ADS with a wide range of estimates.

Some examples of the complexity: 1) According to the literature review of van den Berg & Verhoef (2016) predictions for capacity change with ADS vary from no effect to four times the current capacity. Higher capacity allows for more traffic volume, thus possibly increasing the accessibility (Anderson et al., 2014). 2) ADS can cause a decrease in transport resistance. When travellers can use their time in the vehicle for other purposes than driving, this can cause an increase in value-of-time (VOT) (Correia & van Arem, 2016), which subsequently may increase total vehicle-miles-travelled (VMT) and decrease accessibility. 3) Different spatial planning may change the division of traffic over time and space, changing the accessibility (Zakharenko, 2015). Ideally, all these factors (and many more) are known before predictions can be made how ADS will change accessibility.

Normative

For many decades, congestion has been a problem in the Netherlands with thousands of people stuck in a traffic jam each day. On May 1st, 2017, a publication stated that the Dutch roads will only get busier in the coming decades (Ministerie van Infrastructuur en Milieu, 2017). The congestions can seriously hamper the accessibility of travellers. For some, traffic jams are a constant source of irritation. It is therefore hypothesized that increased accessibility will cause an increased social acceptance.

2.4.3. Environmental Impacts

Technical

Driving behaviour has a large effect on the environmental impact of traffic. "Depending on road type and technology, fuel consumption increased by up to 40% for aggressive driving compared to normal driving. Again, this was more pronounced for emissions, with increases up to a factor 8" (De Vlieger, De Keukeleere, & Kretzschmar, 2000). It is expected that AVs will drive more efficiently and cause less emissions. Subsequently, if AVs reduce congestion they will further improve driving efficiency and therefore the environment. Nevertheless, as mentioned in the analysis of accessibility, the impact on total VMT is unknown which can thus also increase environmental impact.

Normative

Given the Paris Agreement from November 2016, climate change and global greenhouse gasses are on the political agenda (United Nations, 2017). If ADS can decrease the environmental impact of the transport system, they could be more prominently placed on the political agenda. There is an increasing awareness of sustainability among citizens (GfK, 2016). Also, between 2015 en 2017, there has been an increase of 32% in registered electric vehicle in the Netherlands (RVO, 2017). Consequently, it is possible that a reduction in environmental impact can increase social acceptance.

2.4.4. Volume and Composition of Traffic and Transport, Division over Time and Space

Technical

Central in the conceptual framework by van Van Wee, Annema, & Banister (2013) is the volume and composition of traffic and transport, division over time and space. This factor is responsible for some of the fundamental risk factors such as speed differences and mass differences in traffic (Van Wee et al., 2013). Similar to accessibility, this factor is influenced by locations, transport resistance and needs & desires. The technical analysis is therefore not repeated.

However, a very important issue of ADS that is mentioned in multiple publications is the concept of mixed traffic. Mixed traffic entails traffic that consists of automated- and non-automated traffic. During the transition period, it can be beneficial to only allow AVs on specific roads to avoid AVs having to counteract with other traffic. For example, the WEpods in Ede-Wageningen partially drive on segregated lanes (WEpods.nl, n.d.). This allows for a learning curve without proposing too much risk. The next step could be to only allow AVs on highways. 'Letting go on highways' is often mentioned in literature (KiM, 2017). The uniformity of traffic on highways – approximately the same speed and direction – creates good conditions for the technology. However, the relatively high speed limits on highways create a greater risk of damage and fatalities in case of a failure. This risk is especially large for regional roads with no separation between roadways.

Another possibility is to allow automated driving on streets with relatively low speed limits. However, roads with low speed limits are often subject to complex traffic situations, with different modalities, speeds and directions. The more complex conflict situations an AV can encounter, the more sophisticated the technology needs to be. Especially in the Netherlands with busy bicycle, public transport and moped traffic in urban environments, this could cause a problem (De Vlieger et al., 2000).

Normative

Different rationalities between human drivers and software can cause people to be hesitant to trust AVs, while AVs have problems in anticipating human behaviour (van Loon & Martens, 2015). Also, communication issues may arise between AVs and non-automated traffic (Hagenzieker, 2015). Therefore, dangerous situations may arise with mixed traffic. Mixed traffic has a lot to do with the roads that are exempted for AVs, e.g. on which roads is an AV allowed to drive on auto-pilot. Since currently there are hardly AVs on the road, the preference of citizens for mixed traffic is largely unknown. Although some experiments take into account passion for driving or certain traffic conditions, they have not measured the preferences for the type of roads on which automated driving should be implemented (Becker & Axhausen, 2016).

2.4.5. Technology

Technical

The technology of ADS is vastly progressing. Without going in too much detail like sensors and software, this paragraph will analyse the major 'roadblocks' that need to be overcome by technology. Shladover (2016) identifies the following two which will be discussed: 1) The technology needs to improve to allow higher levels of automation; 2) Shladover states that automated driving cannot exist without cooperate driving.

The SAE classification consists of six levels, where level 0 is no automation at all (see Figure 7). Levels 1 and 2 consist of advanced driving assistance systems (ADAS), like Automated Cruise Control, Lane Keeping Assist, Park Distance Control and Traffic Jam Assistant, where level 2 is obviously more advanced than level 1 (De Winter, Happee, Martens, & Stanton, 2014). These levels are not considered actual ADS. Starting from level 3, SAE International classifies the automation as ADS. In literature the acronyms SAD (semi- or conditional automated driving), HAD (Highly automated driving) and FAD (fully automated driving) or also often used for levels 3, 4 and 5 respectively (Bakermans, 2016; De Winter et al., 2014; Jamson et al., 2013). The definition of SAE International will be used, focussed on automation level 3 and level 4.

The complexity of the needs of technology is explained by Shladover (2016). Commonly, AVs use three systems to 'sense' their environment: Lidar, Sonor and camera-imaging. If AVs sense a balloon along its path, it might register this object as unsafe since its systems cannot specify the weight. The AV might stop or make a dangerous manoeuvre to avoid the balloon. However, hitting the balloon is no problem, so avoiding it is a false alarm or false positive. On the other hand, when AVs encounter a small brick, it can seriously damage the vehicle or change its trajectory. So in this case, AVs must avoid collision. If not, it is a missed detection or false negative. Ideally, an AV has near-zero false positives and near-zero false negatives, which is very difficult to reach simultaneously (Shladover, 2016).

In lower automation levels, a human driver is more capable of avoiding false positives or false negatives. The driver is more committed to driving thus better capable to intervene (De Winter et al., 2014). Therefore, it is more difficult to guarantee the safety for higher levels of automation, since humans will have less control to avoid false positives or false negatives.

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
***Human driver* monitors the driving environment**						
0	No Automation	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver Assistance	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial Automation	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
***Automated driving system* ("system") monitors the driving environment**						
3	Conditional Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes
4	High Automation	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	Full Automation	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Figure 7: SAE levels (SAE International, 2014).

However, counter intuitively, the relation between level of automation and safety might not be linear. Multiple studies have shown that level 3 automation can be very dangerous (De Winter et al., 2014; Merat, Jamson, Lai, Daly, & Carsten, 2014). With this level of automation the human driver acts as a back-up system, meaning he/she should intervene in extreme situations. This is the same level of automation commonly used in aviation. The plane can fly itself for most of the time and pilots need to monitor the process of flying. This has proven to be very difficult for humans (Wees & Brookhuis, 2005). Unsurprisingly, pilots are heavily trained to keep their attention at monitoring the system by extensive checklists. Therefore, Shladover (2016) mentions the possibility to skip automation level 3 and go straight to level 4.

Also, AVs may use cooperate driving: communication with certain infrastructure, other vehicles or a cloud to increase awareness of vehicles on the road, coordinate headways, and thus increase efficiency and safety (Behere, Törngren, & Chen, 2013). Shladover (2016) mentions five reasons for the importance of coordination: "1) It compensates for sensor limitations; 2) It provides additional information about other vehicles not measurable by remote sensors; 3) It enables advance alerts about hazards and intentions and negotiation of manoeuvres; 4) It can verify that other vehicles have been seen; and 5) It enables system-level coordination and management".

However, cooperate driving adds to the technical complexity of the system so it might also increase the risk of technical failure. More importantly, external communication affects the vulnerability of deliberate misuse, because of the increase in communication links (Loukas, 2015). Simply said, when AVs are connected to their environment they become more vulnerable for (deliberate or accidental) false information from that environment.

How and when certain levels of automation will be allowed is currently a very important topic. The Tesla autopilot (level 2) already caused some accidents as discussed in §2.4.1. The new Audi A8, bound to be released autumn 2017, is said to have automation level 3 (nu.nl, 2017). It might not be long before car manufacturers claim to have safe vehicles with automation level 4 or even 5. Technically, there is definitely a push for automated driving technologies.

Normative

People are hesitant to high level of automation. Most people prefer ADAS or automation level 3 over automation level 4 or 5 (Becker & Axhausen, 2016). Also legally, it may prove difficult to shift to full automation. According to the Vienna Convention (1986), the human driver needs to be in control of the vehicle at all times. It might lead to a system where a certain interaction between driver and machine is required to make use of the best of both worlds. Dangerous situations may arise when the driver gets bored or driving becomes too complex.

To overcome these issues from an ethical perspective, Santoni de Sio (2016) claims to use the concept of meaningful human control. It requires a meaningful human control over the behaviour of the system. This can be done by either an appropriate design of a partial automation system, or by an appropriate design of a supervised automation system. The aim is to have sufficient human input in the task of driving to ensure safety and responsibility.

2.4.6. Way of Using Vehicles

Technical

Driving behaviour is seen as a major cause of accidents worldwide. Most car accidents in the Netherlands occur because of driving under influence, distractions, eating or drinking and aggressing driving (auto-en-vervoer.infonu.nl, 2015). In the U.S.A., approximately 95 per cent of car accidents occur because of human error (Shladover, 2016). It would therefore seem logical that when humans do not drive the vehicle, roads would become safer. Secondly, driving behaviour can seriously influence congestion and the emissions of the vehicle.

Normative

This paragraph will explore socio-demographics or other personal characteristics that can influence the way people use vehicles. It is possible that these characteristics can explain potential differences between citizens in the social acceptance of ADS. The characteristics are extracted from literature based on two criteria: 1) They applied to the social acceptance, safety, accessibility and/or the environment and 2) They were significant in the conducted research. The following characteristics met those criteria:

Gender, Age & Level of Education

Numerous publications identify different effects between socio-demographics and automated driving. A lot of these studies are about WTP for or WTU. For example, people with a high level of education are more familiar with AVs (Lavasani et al., 2017). They also have a higher WTP. Kyriakidis et al. (2015) found neither clear age- nor gender effects. However, they did find that females were more worried about fully automated driving.

Modal Choice
As already mentioned in §2.4.4, mixed traffic situations can be troublesome for automated driving. Also behavioural problems between different modes of traffic can occur (Hagenzieker, 2015). Do cyclists feel safe when they cross an automated vehicle, even if they have right of way?

Residential Area
Lavasani et al. (2017) found that residents from urban areas showed more interest in using AVs compared to residents in rural areas.

Familiarity with ADAS
ADAS are advanced driving assistance systems, like adaptive cruise control and lane departure warning. Multiple studies found a positive correlation of higher automation levels (mostly ADAS) in current vehicle on intention-to-use and WTP for AVs (Becker & Axhausen, 2016).

Crash Experience
Bansal et al. (2016) found a significant and positive correlation between the number of crashes experienced by an individual and their WTP for AVs. However, Lavasani et al. (2017) found that this correlation was dependent on the severity of the crash. They found a negative correlation between 'minor injuries' and WTP, but a positive correlation between 'major incapacitating' with WTP. Nevertheless, the former research had 347 observations, while the latter only had 144 observations.

Technology & Sustainability Enthusiasm
Finally, it is possible that citizens who have great interest in technology will be more positive towards ADS. They enjoy technological gadgets and might therefore be more interested in AVs. The effect of sustainability enthusiasm is unknown. One might reason that they are in favour of ADS, because they expect that emissions would go down. On the contrary, they might experience every innovation in the car industry as negative since it might stimulate more car-usage.

2.5. Actor Analysis in the Netherlands

So far, this research has covered responsible innovation, the conceptual model, the rise of the automobile and an analysis of possible future automated driving systems. An actor analysis is used to formulate the recommendations for researchers, policy makers and engineers later in this research. It fulfils the objective of this research to contribute to responsible innovation. Figure 8 shows a simplified image of all the actors relevant for the implementation of ADS and their interrelations. They are categorized in policy and research, technology producers and citizens.

Citizens
The focus of this research lies on citizens who can be split into two groups: (potential) users and non-users. The amount of potential users and the frequency of their usage determine demand. As mentioned, multiple publications try to analyse the demand by measuring WTP and WTU of potential users. It is not a goal of this research to identify users and non-users. The division is only made to emphasize that this research will not only focus on users, but also on non-users. Hence the social acceptance of citizens is measured, regardless if someone intends to use AVs.

This research will not focus on the demand. It rather focuses on preferences of social impacts and ultimately social acceptance of ADS. Nevertheless, an estimation of the size of the user group can be made with the results of this research. In the literature review of Becker & Axhausen (2016), the size of the group that intended to use AVs is between 40 and 60 per cent, dependent on methodology and frequency of use.

To meet the criteria of responsible innovation, the social acceptance has to be aligned with technical development, policy and research. However, the social acceptance may change over time and citizens might influence each other. The results of this research will measure the current status-quo and interpret what this means for the implementation of ADS. In order to make ADS a responsible innovation, the social acceptance needs to be measured frequently to determine the change in social acceptance.

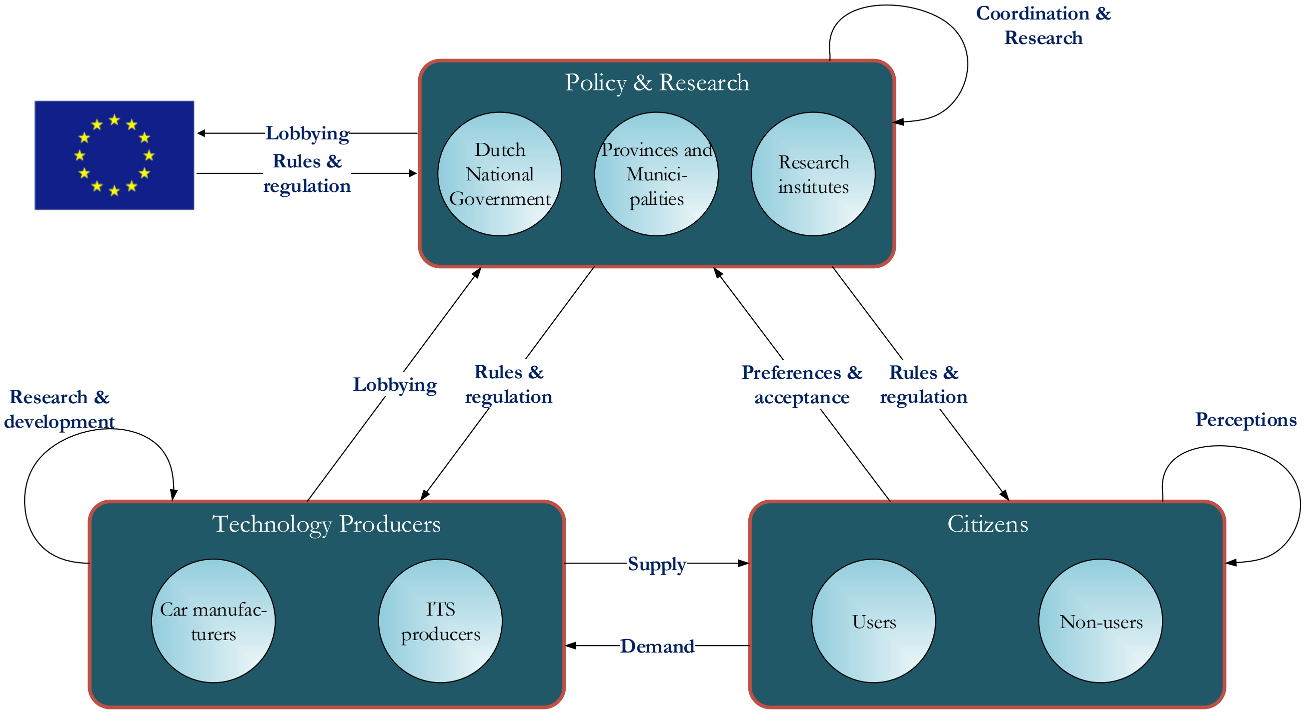


Figure 8: Actors and their interrelations

Policy and Research

The actors in policy and research are strictly taken not all public entities. However, there is a strong connection between policy makers and researchers because the current world of ADS mainly consists of testing and research. Therefore, all actors in policy and research are considered as policy makers. Moreover, they are considered to benefit most from insights of this research.

To start in the top left of the figure, the European Union forces rules and regulations onto the Dutch government. In return, the Dutch government lobbies to influence this legislation. A great example is the Declaration of Amsterdam, which lays down agreements on the steps necessary for the development of ADS in the EU (European Union, 2016).

Next, policy and research is arbitrarily divided into three groups: the National Government, Provinces & Municipalities and Research Institutes. The main actor in the national government is the Ministry of Infrastructure and the Environment, since it is the driving factor behind lobbying, coordination, research and rules and regulation. The second group are the provinces and municipalities, who show great interest in the risks and benefits of ADS and focus on specific projects in their region (STAD, 2017). Finally, the research institutes are actors like the RDW, TNO, KiM, several universities and SWOV. The most important actors in the category policy and research are participating in the Spatial and Transport impacts of Automated Driving (STAD) project (STAD Event, n.d.).

One that is mentioned specifically is the RDW, the Dutch Vehicle Authority. It has the task of licensing vehicles before they are allowed on the public road. Although human drivers have to be in control of the vehicle at all times, vehicles with high automation levels can still be licensed by the RDW. The risk remains that people will use self-driving features on public roads. An example is the fatal accident in U.S. Florida on May 7th, 2016.

Policy makers could have different attitudes towards ADS. During the rise of the automobile, their attitude changed from a passive to an active attitude (Bax, 2012). Also, there is a large variety in policy makers who have a positive attitude and a negative attitude towards ADS (Anderson et al., 2014). If you combine these two scales it results in four stands on policy (see Table 2). These four stands should not be considered deterministic, so combinations are possible. For example, policy makers could be positively active towards ADS on highways, but negatively active towards ADS in urban areas.

Table 2: Perspectives on Policy & Research

	Passive	Active
Negative	Wait and see. Chances are that policy will come too late.	Inflict extensive limitations to technology producers and high penalties to citizens who still use automated driving technologies.
Positive	Conduct own research, but wait for technology producers to come up with the technology. Only intervene when certain norms are overwritten	Make laws and regulations more flexible to allow ADS, cooperate with technology producers to create ADS and/or start campaigns to increase awareness.

According to consultancy firm Roland Berger, the Netherlands is the leader in the development of disruptive technologies in the automotive industry including ADS (Roland Berger, 2017). It is based on an international research on, among others, the implementation of experiments and intelligent transport systems (ITS). To enable market launch for AVs, "the Netherlands shows the most advance approach for developing type approval procedures by embedding lessons learned from test fleets into legislative process" (Roland Berger, 2017; page 14). It indicates that the Dutch government generally takes a positive and active stand towards ADS. There are benefits in being the innovator, like a good international reputation and economic benefits. However, the experiments do propose new and unpredictable risks for human well-being. Partly depending on the social acceptance, it can be questionable if the Netherlands should be the test-bed for ADS. Why should we not copy best practices from other countries?

Technology Producers

Another category entails the producers of the technology, who are divided into two groups: the car manufacturers and the producers of ITS. Together, they produce hardware and software which is needed for automated driving. The best known automated car manufacturers are Tesla and Google's Waymo. However, mostly all large car and truck manufacturers are focusing on some level of automation (Bakermans, 2016; Davies, 2017).

A famous Dutch producer of ITS is TomTom, which purposely develops 3D maps that assist AVs to locate themselves on the road (Kasteleijn, 2017). Another important actor is Mobileye, which develops sensors for AVs. On March 13th 2017, this company was bought by Intel for €14,3 billion dollar (De Financiele Telegraaf, 2017). The fact that practically the whole car market is interested in ADS and high investments are made in the technology, testifies of a certain technology push.

2.6. Summary of Theory

In the problem statement, it was argued that implementing ADS is a wicked-problem: multi-problem (safety issues, accessibility issues and environmental issues), multi-disciplinary (technical and normative) and multi-scale (geographical, temporal and multi-actor). To learn from the past, describe the present and to identify possible issues in the future, this chapter is summarized on a time scale in Table 3. It highlights the multi-disciplinary and multi-scale problems.

In Table 3, the past represents the current, non-automated traffic- and transport system from 1900 until today. The present represents the current status-quo of ADS, while the future represents possible futures of ADS.

Table 3: Summary of Theory

	Past	Present	Future
Technical	The rise of the automobile: Evolution of the automobile and infrastructure	Working prototypes of AVs: WE pods, Truck Platooning, Park Shuttle, Tesla/Audi et cetera	Future ADS, possibly differs in: Level of automation, road exemption, travel time, emissions, human error fatalities, technical failure fatalities, deliberate misuse fatalities
Empirical	It is argued that the number of fatalities in the past (seventies) were not socially accepted. It took effort and protests from the public to improve traffic safety	**The social acceptance of ADS by Dutch citizens which will be measured in this research,** possibly explained by: Gender, age, level of education, modal choice, residential area, familiarity with ADAS, crash experience, technology enthusiasm, sustainability enthusiasm	The social acceptance is likely to change in the future, partly because we will get more familiar with ADS if they are implemented. It will result in new norms and values and therefore social acceptance
Conceptual	It is argued that the number of fatalities in the past (seventies) were not ethically acceptable.	It is questionable if the current safety record is acceptable. It is therefore possible that policy makers have the moral obligation to implement ADS if they are safer than the current system	The concept of meaningful human control can be used to overcome ethical dilemmas. The voice of future generations should be included in conceptual research
Geographical	The same problems with traffic safety occurred mostly around the world. Currently, traffic safety differs around the world.	Given the contrast in traffic safety records of current systems around the world, citizens from different countries might have a different social acceptance of ADS	The implementation of ADS is likely to differ between countries. They can learn from each other by identifying best practices.
Multi-actor	Policy makers were lacking. Technology producers were not restricted. Awareness was raised by society (ANWB)	Policy makers have budgets, knowledge and research capabilities. Technology producers are not restricted.	According to the results of the empirical experiments, recommendations will be made to policy makers and technology producers.

3. Methodology

This chapter will construct the empirical experiment to measure the social acceptance, which will combine elements from the antecedent analyses. The experiment has been created with great care to make it understandable for average citizens without making them biased with too much information.

This chapter is structured as follows: Firstly, the format of the experiment is defined. Next, the attributes from the technical analysis are specified and attribute levels are determined, followed by the models that will be used to estimate the parameters. §3.5 will define the experimental design, i.e. the systematic variation of the attribute levels. The last paragraph will explain how the respondents are selected and how the sample compares to Dutch society in terms of socio-demographics.

3.1. Stated Choice Experiment

A survey is chosen as research method since it is a relatively inexpensive, flexible method to achieve extensive information about characteristics of a population. The goal is to estimate reliable and unbiased parameters. However, it has been argued that the social acceptance relates to behaviour and is difficult to observe or even predict. To be able to measure the social acceptance of ADS, it is compared to the social acceptance of the current transport system. In doing so, it can be determined what percentage of citizens prefer ADS over the current system. Since the social acceptance of the current transport system is assumed to be neutral, the percentage of citizens that think ADS are socially accepted is determined.

Respondents have to choose if they prefer a hypothetical future of ADS over the current transport system. The chosen format is single referendum contingent valuation. In this hypothetical referendum, respondents are asked as citizens to vote 'for' or 'against' a future with ADS. If respondents vote 'for', they must assume that this future becomes a reality. If respondents vote 'against', they must assume that the current system stays in place. Single referendum contingent valuation has gained widespread use in applications of (semi-)public goods (Green, Jacowitz, Kahneman, & McFadden, 1998). Examples are trade-offs in travel time, road-tax, traffic safety (Mouter & Chorus, 2016) and voting support for congestion charging (Hensher & Li, 2013)

It is hypothesized that the choices of respondents with respect to systematically varied attribute levels of ADS will change according to the social acceptance of these levels. It is also hypothesized that the choices differ between respondents because of different norms and values among citizens and because they all have different knowledge of ADS. These differences might be partly explained by personal characteristics. Hence, the experiment is a stated choice experiment that entails hypothetical futures of ADS as choice alternatives.

3.2. Attributes & Levels

In the theoretical analysis it has been argued that level of automation, mixed traffic and cooperate driving can lead to technical- and normative issues. It is therefore hypothesized that automation levels and the exemption of roads influence the social acceptance. Cooperate driving is not included for the following reasons: If Shladover is right, automated driving will not exist without cooperate driving so any social debate on the matter will be meaningless. Secondly, it is expected that most citizens will not have the required knowledge to make the link between cooperate driving and improved social impacts. For this research, cooperate driving is therefore considered a technical decision rather than a political decision.

Next, it is hypothesized that three social impacts have an effect on the social acceptance: safety, accessibility and environmental impact. Safety is divided in human error fatalities, technical failure fatalities and deliberate misuse fatalities. Accessibility is influenced by locations, transport resistance and needs & desires. For the scope of this research, it is assumed that locations and needs & desires are constant over time. Although change in spatial planning is expected on the long run with ADS, this effect is assumed to be negligible for the transition phase. Moreover, it is not deemed critical for the social acceptance. The focus will therefore be on the difference in transport resistance, most commonly expressed in travel time. For environmental impact, emissions are assumed to be the key attribute. Each of these attributes will be specified with their attribute levels. The safety attributes levels are scaled 1:10 to scale the parameters.

Table 4: Attribute levels ADS and effects

Attribute	Attribute levels	Coding	Sign	Explanation
Level of automation	• Automation level 3 • Automation level 4	• 0 • 1	aut	There are five levels of automation by the SAE, of which level 3, 4 and 5 are considered actual automated driving. For the near future, level 5 is assumed impossible for private vehicles, because of technology shortcomings and liability issues. That leaves automation level 3 and 4.
Road exemption	• Highways (100-130 km/h) • Regional roads (50-80 km/h) • Inside built environment (30-50 km/h)	• 0 • 1 • 2	rex	Two criteria determine the attribute levels: The speed and crossings with other traffic. Three categories are chosen that provide a sufficient distribution of these criteria.
Human error fatalities	• 300 • 400 • 500 • 600	• 30 • 40 • 50 • 60	hum	Currently, there are approximately 600 human error fatalities annually, which is the upper bound under the assumption that AVs are not implemented if they are less safe than conventional vehicles. 440 fatalities are the result of accidents with a car involved. It is assumed that, especially in the transition phase, AVs cannot prevent all of these 440 fatalities. An optimistic scenario is assumed to have a total of 300 fatalities annually.
Technical failure fatalities	• 0 • 40 • 80	• 0 • 4 • 8	tec	It is unknown how many accidents currently happen due to technical failure. The Bron geRegistreerde Ongevallen Nederland (BRON) do not record these failures. It is assumed that technical failure currently only contributes to the main reason of the accident. Therefore, the current number of accidents due to technical failure is assumed negligible, so 0 is the lower bound. The upper bound is more difficult to determine. Again, the assumption is made that AVs are not implemented if they are less safe, so it cannot be a large number. An upper bound of 80 fatalities is chosen, since it is assumed that more fatalities are unrealistic for implementation. The level '40 fatalities' is added to test for non-linearity.
Deliberate misuse fatalities	• 0 • 30 • 60	• 0 • 3 • 6	mis	The current number of deliberate misuse is unknown, but assumed negligible. The upper bound is arbitrarily set on 60 fatalities. Again, more fatalities are assumed unrealistic for implementation.

Average change in travel time for all road users	• - 25 % • No change • +25 %	• -25 • 0 • 25	tti	A 25 per cent change in travel time is assumed to have a significant effect in utility and realistic to respondents. Since it is unknown how travel time will develop, an increase and decrease of 25 per cent is used, with the level no change to test for non-linearity.
Average change in emissions of all road users	• - 25 % • No change • +25 %	• -25 • 0 • 25	emi	The change is emissions is assumed to be less certain than change in travel costs due to the uncertainty of the introduction of alternative fuels for cars. Therefore, a 25 per cent increase and decrease is used, with the level no change to test for non-linearity.

The theoretical analysis also argued that gender, age, level of education, modal choice, residential area, familiarity with ADAS, crash experience and technology- & sustainability enthusiasm are personal factors that can influence the social acceptance. All attributes are included in the survey. The next table will provide the attribute levels and questions that will be used in the survey, which are extracted from surveys from the KiM or the CBS database. The coding is found in Appendix I.

Table 5: Attribute levels Socio-demographic factors

Attribute	Attribute level	Sign	Question
Gender	• Male • Female	gen	What is your gender?
Age	• 18-30 • 30-40 • 40-50 • 50-60 • 60-70 • 70-80 • 80+	age	What is your year of birth?
Level of education	• Basisonderwijs • Voortgezet onderwijs • Middelbaar beroepsonderwijs • Hoger beroepsonderwijs, universiteit • Anders, nl..	edu	What is your level of education?
Residential area	• Rural area • Village • Suburbs • City • City centre	res	Which of the following regions best describes your state of residence?
Licence	• Yes • No		Are you in possession of a drivers' licence?
Modal choice	• (Almost) every day • 5-6 days per week • 3-4 days per week • 1-2 days per week • 1-3 days per month • 6-11 days per year • 1-5 days per year • Less than 1 day per year	mod	How often do you travel with one of the following modalities? [car as driver, car as passenger, train, bus/tram, bicycle]

Current usage of ADAS1	• Yes • No • I don't know	ada1	Are you familiar with one the following driving assistance systems? [adaptive cruise control, lane departure warning, blind spot detector, automated park assist]
Current usage of ADAS2	• Yes • No • I don't know	ada2	Have you used one the following driving assistance systems? [adaptive cruise control, lane departure warning, blind spot detector, automated park assist]
Past experience with minor accident	• Yes • No	acm	Have you been involved in a traffic accident where **no** one involved suffered serious injury?
Past experience with severe accident	• Yes • No	acs	Have you been involved in a traffic accident where one or more involved suffered serious injury?
Environmentalist	• Not at all • Hardly • Neutral • To some extent • Very	env	To what extent does the expression 'I am environmentally conscious' apply to you?
Technology lovers	• Not at all • Hardly • Neutral • To some extent • Very	tlo	To what extent does the expression 'I find technology fun and interesting' apply to you?

3.3. Respondents

The experiment is held among a sample of 535 adult Dutch citizens collected via the paid panel PanelClix. The data was checked on fraud data and socio-demographics of the respondents. Fraud data is defined as data of respondents that contained missing answers or their survey's completion time was below 3 minutes (which is not deemed humanly possible). The data of respondents that is categorized as fraud were excluded from further research. After deleting 25 fraud results, 510 complete surveys were retrieved. 50% was female and 50% was male, like in Dutch society (CBS, 2017). The distributions for age and level of education are as follows:

Table 6: Age and level of education distribution

Age	% of respondents	% in Dutch society (CBS, 2017)	Education	% of respondents	% in Dutch society (CBS.nl, 2013)
18-30	0,17	0,16	Primary	0,02	0,08
30-40	0,21	0,15	VO & MBO	0,62	0,63
40-50	0,16	0,18	HBO	0,28	0,19
50-60	0,20	0,19	Universiteit	0,08	0,10
60-70	0,22	0,16			
70-80	0,04	0,10			
80+	0,01	0,06			

Citizens above the age of 70 years old are slightly underrepresented. This could be explained by the fact that the survey was presented on a computer. Also, citizens with primary education are slightly underrepresented. Nevertheless, it is concluded that the respondents are a good representation of Dutch society.

3.4. Estimation Models

This paragraph will explain the models that are used to estimate parameters for the research questions, MNL model §3.4.1 and latent class choice model §3.4.2.

3.4.1. MNL Model

A well-known model in discrete choice modelling is the linear-additive multinomial logit model (MNL; Formula 1). This model is chosen since "this model allows for decision makers' heterogeneity and inconsistency, is econometrically rigorous and is flexible and practical in multi-attribute and multinomial choice situations" (Chorus, 2017; page 19).

The total utility U of alternative *i* is the systematic utility V plus an error term ε. The systematic utility depends on the sum of the taste parameters β times the attributes x. According to the Random Utility Maximization (RUM; formula 2) theory by Nobel price-winning Daniel McFadden, alternative *i* is chosen if it has a higher utility than alternative *j* (McFadden, 2000).. The model assumes that people are utility-maximizers and assumes an error term that is independent, identically distributed EV Type I with variance $\frac{\pi^2}{6}$, which allows inferring preferences and trade-offs from people's choices.

[1] $$U_i = V_i + \varepsilon_i = \sum_m \beta_m \cdot x_{im} + \varepsilon_i$$

[2] $$\sum_m \beta_m \cdot x_{im} + \varepsilon_i > \sum_m \beta_m \cdot x_{jm} + \varepsilon_j, \forall j \neq i$$

A basic MNL RUM model will be estimated for the first three research questions in which all parameters are treated as linear. The utility function is below, in which V1 is the utility of the automated driving system and V2 is the utility of the current transport system (which is thus fixed at 0). Besides, there is accounted for the panel structure of the data in the estimation of the parameters.

- V1 = B_ASC_1 + B_rex_1 * rex + B_aut_1 * aut + B_tti_1 * tti + B_emi_1 * emi + B_hum_1 * hum + B_tec_1 * tec + B_mis_1 * mis
- V2 = 0

Acceptance in Number of Fatalities

Possibly, respondents simply added the number of fatalities of all three categories to determine their acceptance of a specific design. In other words, the hypothesis is that respondents are indifferent to different causes of fatalities. To test this hypothesis, two statistical tests will be conducted: 1) is a model with three parameters for each category statistically better than a model with only one parameter for all three categories? And 2) are the parameters statistically different from each other?

For the first test, a model will be estimated with only one parameter for human error fatalities, technical failure fatalities and deliberate misuse fatalities. This part of the utility function looks as follows:

- B_fat_1 * (hum + tec + mis)

For the second test, the 95 per cent confidence interval is calculated for each of the three parameters by adding/subtracting 1.96 times the standard error (p=0.05). It is tested if either one of the other two parameters is in the confidence interval.

Quadratic Functions

Finally, the parameters will be tested on linearity. The hypothesis is that respondents do not treat the attribute levels linearly. They might accept a small decrease/increase of an attribute, but have a strong dislike/preference for a large decrease/increase of an attribute. For example, respondents may accept a low number of technical failure fatalities, but will strongly oppose to a large number of technical failure fatalities. All the social impacts are tested on linearity.

3.4.2. Latent Class Choice Model

To test for heterogeneity among citizens, a latent class choice model with a class membership function is used. This model assumes that there are classes of citizens that are homogeneous in their preferences. However, these classes cannot be observed since they are latent. This model identifies these classes and with observable characteristics like socio-demographics it can predict who belongs to which class. As such, the chance that an individual with certain socio-demographics votes for or against ADS can be estimated. Since other models, like mixed logit, are not capable of explaining the heterogeneity, the latent class choice model is chosen.

This model is described with formula 3. Herein, $P_n (i \mid \beta)$ is the Probability that decision-maker n chooses alternative i, conditional on the model parameters β. π_{ns} is the class membership probability and $P_n (i \mid \beta_s)$ is the probability of n choosing i, given that decision maker n belongs to class s (Cranenburgh, 2017). For the class membership probability (formula 4), class-specific constants δ_s and vector of parameters γ_s are estimated. The linear-additive function $g(\circ)$ gives the functional form of the utility for in the class allocation model. This function is based on z_n, which are observed variables like socio-demographics or context variables. All models will account for the panel structure of the data, i.e. they will account for the fact that the respondents make several choices.

[3] $$P_n(i \mid \beta) = \sum_{s=1}^{S} \pi_{ns} P_n(i \mid \beta_s)$$

[4] $$\pi_{ns} = \frac{e^{\delta_s + g(\gamma_s, z_n)}}{\sum_{l=1..S} e^{\delta_l + g(\gamma_l, z_n)}}$$

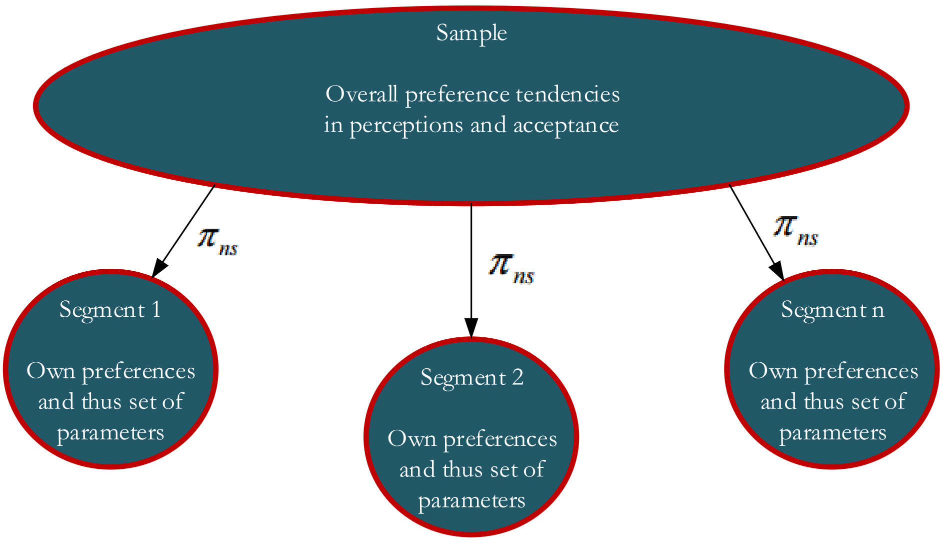


Figure 9: Latent class choice model with class membership

3.5. Experimental Design

To vary the attribute levels for respondents in different choice tasks, the software programme Ngene was used to make the experimental design. The following four considerations were made to write the Ngene syntax.

Firstly, in §3.2 it was assumed that ADS will not be implemented if ADS causes more fatalities than the current system. Collectively the three types of fatalities may not exceed 600 fatalities. So, when the attribute level of human error fatalities is '600', technical failure fatalities and deliberate misuse fatalities must be '0'. To comply with this restriction, conditions were added to the syntax. It also means that an orthogonal design is no longer possible, so an efficient design is used. More specifically, a D-efficient design is chosen since its takes into account both variances & covariances.

Secondly, it is assumed that more than 12 choice tasks (thus 12 rows) for each respondent would be too exhausting. With more choice tasks, the risk that respondents will take the final questions less seriously becomes too large. In that case, no information can be withdrawn from these questions.

Since an efficient design will be used, priors are needed. A small pilot study (N=21) was used to measure the priors (see Table 7). The respondents were recruited from the social circle of the author of this research.

Table 7: Priors

Prior	Name	Value
Alternative specific constant	B_ACS	3.74
Level of automation	B_aut	-0.430
Road exemption	B_rex	-0.021
Travel time	B_tti	-0.049
Emissions	B_emi	-0.050
Human error fatalities	B_hum	-0.0061
Technical failure fatalities	B_tec	-0.0065
Deliberate misuse fatalities	B_mis	-0.0271

Fourth and finally, the number of blocks has to be decided upon. To do so, four designs were tested. Design A is a design without the assumption that no more than 600 fatalities may occur. It is an orthogonal design and used as benchmark for the other designs. Design B, C and D are D-efficient designs with the abovementioned assumptions with increasing number of blocks (and thus rows).

Table 8: Survey design: number of blocks

Survey Design	Rows (choice sets)	Blocks	Highest correlation	Highest SP-estimate	# of observation requiered
A	12	1	0.286	443	5316
B	12	1	0.416	458	5496
C	24	2	0.303	208	4992
D	36	3	0.431	141	5076

More blocks results in a lower SP-estimate, which indicates how many respondents are required before the parameter becomes statistically significant. By adding more blocks, more choice sets are added to the survey. Respondents will be distributed in blocks over these choice sets. To account for this effect, the total number of required observations is calculated by multiplying the SP-estimate with the number of choice sets. The design with the lowest required number of observation is the most efficient design.

Design C is chosen since it is the most efficient design. Moreover, the highest correlation of this design is lower than designs B and D. The syntax below is used and the full design is found in appendix II. Besides, the coding used to estimate priors and determine the experimental design is different from the coding used in the estimation of the parameters of the survey results. The different coding scheme is used to better interpret the results in chapter 4.

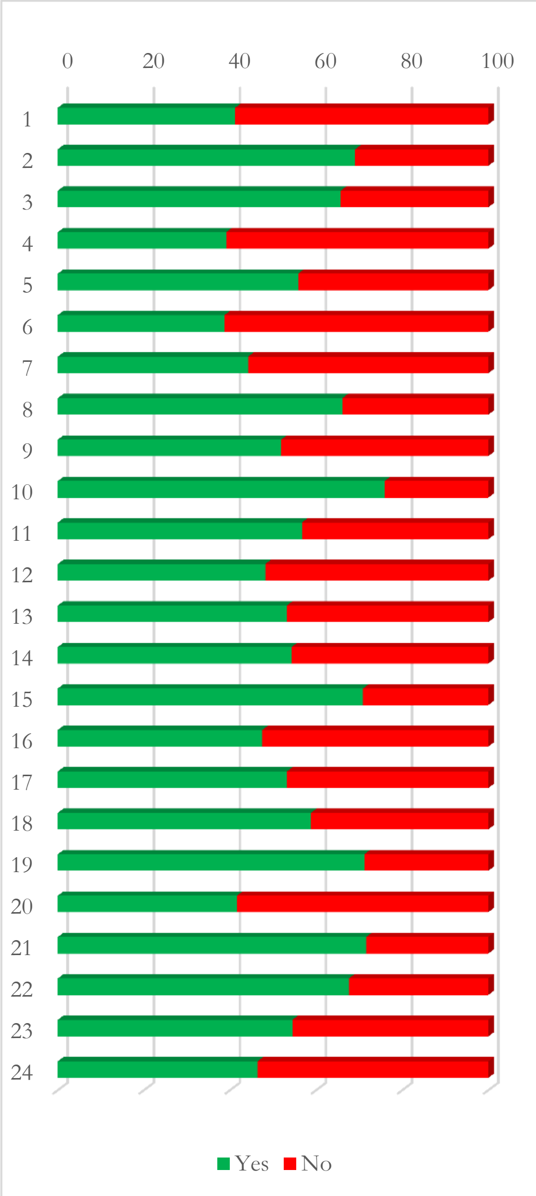
```
? AV acceptance
Design
;alts=ADS,CS
;rows=24
;block=2
;eff=(mnl,d)
;con
;cond:
if(ADS.Hum=600, ADS.Tec=0),
if(ADS.Hum=600, ADS.MIS=0),
if(ADS.Hum=500, if(ADS.Tec=80, ADS.mis=0))
;model:
U(ADS) =b0 + b1 * rex[120,80,50] + b2 * aut[3,4] + b3 * tti[-25,0,25] + b4 * emi[-25,0,25]
+ b5 * hum[300,400,500,600] + b6 * tec[80,40,0] + b7 * mis[60,30,0]
$
```

4. Results

This chapter will present and interpret the results of the survey. Below is a response overview per choice task expressed in percentages. The green beam shows the percentage of citizens that are in favour of the suggested future of ADS, the red beam shows the percentage of citizens that are against the suggested future. It provides a good overview of the different choices that citizens made along the survey. Most choice tasks are within 30 to 70 per cent acceptance, while exactly half of the choice tasks are within 40 to 60 per cent acceptance. Note that each respondent was presented only 12 choice tasks, but that the respondents were divided in blocks.

Table 9: Response overview per choice task

Choice task	rex	aut	tti	emi	hum	tec	mis
1	80	3	25	25	300	40	30
2	50	3	-25	25	300	40	0
3	80	3	0	-25	600	0	0
4	80	4	0	25	400	40	30
5	120	3	25	0	300	80	0
6	80	4	25	25	400	0	30
7	120	4	-25	25	300	40	60
8	120	4	-25	25	400	0	0
9	120	4	25	-25	500	0	30
10	50	3	0	0	300	0	30
11	50	3	-25	-25	500	40	60
12	80	4	0	0	500	40	30
13	50	4	-25	-25	300	80	60
14	50	3	25	-25	300	80	30
15	80	3	0	-25	500	40	0
16	120	4	0	-25	400	80	60
17	120	3	-25	25	400	80	0
18	120	3	25	0	300	0	60
19	50	4	-25	0	500	0	0
20	50	4	0	0	400	80	30
21	120	3	-25	0	500	0	60
22	50	4	25	-25	300	0	0
23	80	3	0	25	400	0	60
24	80	4	25	0	400	40	0



The MNL model is used to estimate parameters from the data to answer the first three research questions. The latent class choice model is used for the final research question. The structure of this chapter follows the four research questions.

4.1. The Social Acceptance of Automated Driving Systems

A MNL model is estimated for the first three research questions. It is tested if the parameters are linear and if the parameters for different types of fatalities are statistically different from one another. The model with the 'best' fit includes a quadratic function for human error fatalities. It has a statistically better model fit than a model with only linear parameters (likelihood ratio statistic (LRS) of 15, with a critical value of 3.841 (df=1; p=0.05)). Also, a model with three different parameters for types of fatalities is statistically better than a model with only one parameter for all fatalities (LRS of 132, while the critical value is 5.991 (df =2; p=0.05)). The parameters human error fatalities, technical failure fatalities and deliberate misuse fatalities are also statistically different from one another (p=0.05).

Therefore, a MNL model with the following utility function is used to answer the first four research questions:

- V1 = B_ASC + B_rex * rex + B_aut * aut + B_tti * tti + B_emi * emi + B_huq * hum * hum + B_tec * tec + B_mis * mis
- V2 = 0
-

Estimation Results

Number of estimated parameters	8
Sample size	6120
Excluded observations	0
Init log likelihood	-4242.061
Final log likelihood	-4048.934
Likelihood ratio test for the init. model	386.253
Rho-square for the init. model	0.046
Rho-square-bar for the init. model	0.044
Akaike Information Criterion	8113.869
Bayesian Information Criterion	8167.623
Final gradiant norm	+2.357e-004
Diagnostic	Trust region algorithm with simple bounds (CGT2000): Convergence reached
Iterations	14
Data processing time	00:00
Run time	00:02
Nbr of threads	4

Table 10: All parameters of the MNL model to determine the social acceptance

Parameter	Name	Value	Standard Error	t-test	p-value
Alternative specific constant	B_ACS_1	1.54	0.121	12.66	0.00
Level of automation	B_aut_1	-0.385	0.058	-6.66	0.00
Road exemption	B_rex_1	-0.046	0.033	-1.34	0.16*
Travel time	B_tti_1	-0.015	0.001	-11.24	0.00
Emissions	B_emi_1	-0.016	0.002	-10.38	0.00
Human error fatalities quadratic	B_huq_1	-0.00028	3.77 E-05	-7.43	0.00
Technical failure fatalities	B_tec_1	-0.101	0.009	-11.67	0.00
Deliberate misuse fatalities	B_mis_1	-0.137	0.015	-9.14	0.00

The first research question is: *What percentage of citizens thinks automated driving systems are socially accepted?*

This question relates to the intrinsic preference for ADS and thus the alternative specific constant (ASC). The ASC shows the utility of ADS if all attribute levels are 0. However, this would indicate that ADS causes no traffic fatalities at all which is not very realistic. Therefore a design is formulated that has the same social impacts as the current system and minimal mixed traffic. The result is called the 'least-impact' design: automation level 3 only allowed on highways and no social impacts (0% change in travel time and emissions, 600 human error fatalities and 0 technical- and deliberate misuse fatalities). The utility of the least-impact design can be interpreted as the intrinsic preference for ADS. In other words, this utility indicates if citizens would accept the simplest form of ADS with no social benefits in safety, accessibility or environmental impact.

The ASC is 1.54 and the quadratic human error fatalities parameter is -0.00028. Therefore, the utility of the 'least-impact design' is 1.54 + -0.00028 * 60 * 60 = 0.53. The utility of the current transport system is fixed at 0. The acceptance percentage can be calculated with the MNL model:

$$\frac{e^{V1}}{e^{V1} + e^{V2}} = \frac{e^{0.53}}{e^{0.53} + e^{0}} = 63\%$$

This indicates that the 'least-impact' design is socially accepted by 63% of the respondents, while 37% are against this design of ADS. It could be interpreted that a majority is positive towards ADS. It does not mean that 63% would use AVs. This percentage does not account for socio-demographics like income and current modal choice. Therefore, the percentage that intends to use AVs will likely be lower. This coincides with the estimation between 40% and 60% in the literature review of Becker & Axhausen (2016).

The parameter value for level of automation (-0.385) indicates the utility difference between automation levels 3 and 4. The preference for lower automation levels coincides with literature (Becker & Axhausen, 2016). Road exemption is not statistically significant. A mixed logit (ML) model is also estimated to exclude the possibility that heterogeneity caused this factor to not become significant (see Appendix V). The results mean that citizens have preference for road exemption. Since this is the first study that explores preferences for road exemption that is known to the author, more research should be conducted to validate this result.

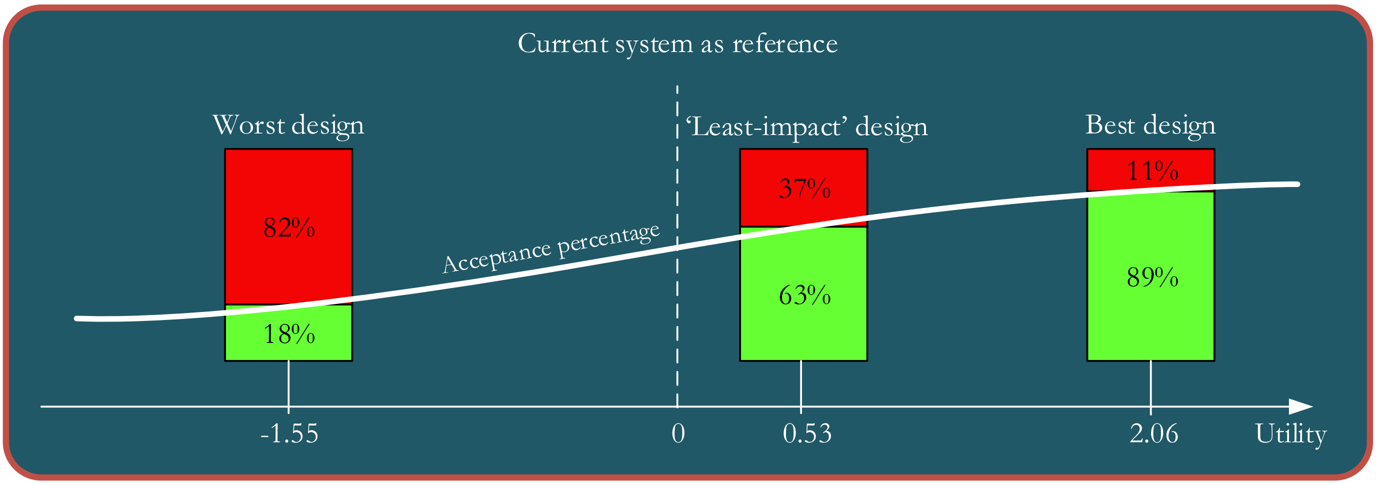


Figure 10: Different designs, their utility and social acceptance

The estimated model allows predicting the utility and therefore social acceptance of every other design. The best possible design in the survey has automation level 3, -25% travel time and emissions, 300 human error fatalities, and no other fatalities. It has a utility of 2.06 which means that this design is accepted by 89% of citizens, while 11% are against this design. The worst design in the survey has automation level 4, +25% travel time and emissions, 500 human error fatalities, 40 technical failure fatalities and 60 deliberate misuse fatalities (note that it is assumed that the total number of fatalities must not exceed 600). It has a utility of -1.55 which means that this design is accepted by 18% of citizens, while 82% are against this design.

The acceptance percentage follows an s-curve (see Figure 10). The maximum slope is at the centre where both ADS and the current system have a utility of 0. Since the least-impact design has a utility higher than 0, it indicates that negative social impacts have a larger effect on the social acceptance than positive social impacts of the same magnitude. The effect is that an increase in social acceptance is more difficult to realize than a decrease in social acceptance.

4.2. The Influence of Safety, Accessibility and Environmental Impact

The second research question is: *How is the social acceptance influenced by safety, accessibility and environmental impact?*

The attributes (parameters) for accessibility and environmental impact are travel time (-0.015) and emissions (-0.016). The parameters indicate the change in utility by 1% increase in travel time/emissions if the other attribute levels do not change. For example, 1% increase in travel time means a 0.015 decrease in utility.

The safety attributes (parameters) are human error fatalities (-0.00028; quadratic), technical failure fatalities (-0.101) and deliberate misuse fatalities (-0.137). They are scaled 1:10. For technical failure fatalities and deliberate misuse fatalities they represent the change in utility per 10 extra fatalities. For human error fatalities the parameter represents 0.0252 utility decrease per 10 extra fatalities for the chosen range.

To compare different social impacts the utility contributions of the attributes are calculated. The utility contribution is the value of each parameter multiplied by the range in attribute levels. It shows the importance of the attributes in the decision-making by citizens. They are shown in Table 11, categorized by importance.

Table 11: Utility Contributions

Attribute	Utility contribution
Travel time	0.75
Human error fatalities	0.76
Emissions	0.80
Technical failure fatalities	0.81
Deliberate misuse fatalities	0.82

Travel time is the least important attribute according to the utility contributions. Especially the fatalities caused by AVs are important. Technical failure fatalities and deliberate misuse fatalities are the highest ranked in the utility contributions. The results coincide with the found safety worries in literature (Kyriakidis et al., 2015). The next paragraph will elaborate on differences between the different parameters for safety.

4.3. Valuation of Traffic Fatalities Caused by Automated Vehicles and Current Traffic Fatalities

The third research question is: *Are traffic fatalities caused by automated vehicles valued differently than current traffic fatalities?*

The previous research question already emphasized the importance of safety. This question will elaborate on different causes of accidents. An expected yet interesting result is that citizens have a stronger dislike for fatalities caused by automated driving than human error fatalities. The strongest dislike is for deliberate misuse fatalities. Technical failure fatalities are weighed as much as 4 human error fatalities. Roughly said, every 4 humans saved annually because of less human error can only be replaced by 1 technical failure fatality to reach the same social acceptance. Between human error fatalities and deliberate misuse fatalities is a factor 5.5. These results coincide with literature. Kyriakidis et al. (2015) found that people are very concerned about misuse of AVs, a little bit more than about technical failure.

The difference between types of fatalities can partly be explained by the regret theory, in which losses loom larger than gains of an equal magnitude (Chorus, 2017; Loomes & Sugden, 1983). A decrease in human error fatalities is a gain compared to the current system; an increase in fatalities caused by AVs is a loss compared to the current system. As following from the regret theory, an increase of a certain amount of fatalities caused by AVs loom larger than a decrease of the same amount of human error fatalities. The result that deliberate misuse fatalities are even worse than technical failure fatalities can be partly explained by the theory of Pearsall & Hanks (2001) on safety and security.

Although it was hypothesized that citizens would have a stronger dislike for fatalities caused by AVs than current traffic fatalities, the magnitude is larger than expected. If it is assumed that ADS could prevent 20% of human error fatalities of the current system, thus 120 fatalities, then still ADS can only result in 22 deliberate misuse fatalities to reach the same social acceptance. Although such absolute comparisons are not very realistic, it does show that fatalities caused by AVs are hardly accepted.

It is concluded that traffic fatalities caused by ADS are valued to be more important than human error fatalities. Especially deliberate misuse fatalities are considered to be much more important than current traffic fatalities.

4.4. The Heterogeneity in Social Acceptance among Citizens

The Latent Class Choice model is estimated to answer the final research question: *Is there heterogeneity in the social acceptance among citizens?*

The latent class choice model has the same utility function as the MNL model for each class except for a linear parameter for human error fatalities since the quadratic parameter did not become significant. Also, it has to be decided which number of classes fits the data 'best'. Therefore, models with an increased number of classes are estimated, while their final loglikelihood (LL), Bayesian Information Criterion (BIC) and Rho^2 are compared. Possibly, a local maximum LL is found, so 14 different starting values for the class membership parameters are used. The best results are shown in the table below.

Table 12: Number of Classes in Latent Class Choice Models

Name	LL	BIC	Rho^2
Basic RUM model	-4051	8162	0.045
2 class model	-3499	7120	0.175
3 class model	-3316	6893	0.218
4 class model	-3243	6791	0.236
5 class model	-3179	6743	0.250
6 class model	-3150	6761	0.258

By adding more parameters, it is logical that the LL and Rho^2 improve. However, the size of improvement gets smaller by the addition of more classes. The BIC is the key criterion for deciding how many classes fit the data best, since it takes the added parameters into consideration (Cranenburgh, 2017). For this criterion, the 5 class model fits best, since it has the lowest BIC.

In Appendix VI the results of the 5 class model can be found. Unfortunately, the results were not suited for interpretation. Having 510 respondents might not be sufficient for this extensive model. The fourth class only contains approximately 50 respondents, which are not many respondents to measure a MNL model. Therefore, models with fewer classes are explored. Appendix VI shows the loglikelihood function for different number of classes. There is a big bend in this function at the 3 class model, meaning that loglikelihood does not improve substantially after this point. The 3 class model also showed the best results for interpretation. Therefore, this research will interpret the results of the 3 class model.

Estimation Results

Number of estimated parameters	30
Sample size	6120
Excluded observations	0
Init log likelihood	-4242.061
Final log likelihood	-3315.817
Likelihood ratio test for the init. model	1852.488
Rho-square for the init. model	0.218
Rho-square-bar for the init. model	0.211
Akaike Information Criterion	6691.633
Bayesian Information Criterion	6893.213
Final gradiant norm	+2.858e-003
Diagnostic	Trust region algorithm with simple bounds (CGT2000): Convergence reached
Iterations	32
Data processing time	00:00
Run time	01:47
Nbr of threads	4

The class membership parameters and corresponding probabilities are added to determine if the classes can be interpreted by socio-demographic or other personal factors. All factors from §2.4.6 are modeled in the 3 class LC model. Modal choice, residential area, crash experience and technology- & sustainability enthusiasm did not become statistically significant. The classes that are formed by the latent class model cannot be explained by these socio-demographics. In other words, citizens have a probability to be assigned to either class regardless of their modal choice, residential area, crash experience and technology- & sustainability enthusiasm.

Gender and familiarity with ADAS did become statistically significant and will be interpreted from the 3 class model. The significant parameters to answer the fifth research question are highlighted in red in Table 13. The results show that citizens can be segmented to three classes: automated driving enthusiasts, central mass and the risk-averse class. Each of the classes will be described according to the parameters. The full results of the 3 class model can be found in Appendix VIII.

Table 13: 3 Class LC Model Parameters

Parameter	Automation enthusiasts	Central mass	Risk-averse class
Alternative specific constant (ASC)	4.67	2.46	-0.455
Level of automation (aut)	0.485	-0.836	0.638
Road exemption (rex)	-0.145	-0.088	0.430
Travel time (tti)	-0.449	-0.550	-0.603
Emissions (emi)	-0.630	-0.555	-0.284
Human error fatalities (hum)	-0.0045	-0.0028	-0.0048
Technical failure fatalities (tec)	-0.0116	-0.0139	-0.0273
Deliberate misuse fatalities (mis)	-0.0145	-0.0183	-0.0430
Class Membership parameters			
Class (s)	0 (fixed)	0.494	-0.663
Gender (gen)	0 (fixed)	0.240	0.420
ADAS familiarity (ada1)	0 (fixed)	0.290	-0.026

Table 14: Class Membership Probabilities

Probabilities		Automation enthusiasts	Central mass	Risk-averse class
Class membership probabilities		32%	52%	16%
Gender (gen)	Probability that a male belongs to a class	38%	49%	13%
	Probability that a female belongs to a class	26%	54%	20%
ADAS familiarity (ada1)	Probability that someone who is **not** familiar with ADAS belongs to a class	36%	45%	19%
	Probability that someone who is familiar with ADAS belongs to a class	27%	59%	14%

Class 1: Automated Driving Enthusiasts

The first class has the highest social acceptance of ADS. The ASC is relatively high and overall they have the lowest penalty for fatalities of all classes. Even so, they still value fatalities caused by AVs to be worse than human error fatalities. The 'least-impact' design has a utility of 1.98 and is therefore socially accepted by 88% of this class. An interesting result is that the parameter for level of automation becomes positive for this class. It means that this class prefers automation level 4 over level 3. In the MNL model the average preference of citizens is the other way around. It explains the significant higher Rho^2 for the latent class model in comparison with the MNL model. Altogether this class has the highest social acceptance towards ADS. 32% of all citizens can be assigned to this class.

Males have a higher chance to be assigned to this class. On average, males have a higher social acceptance of ADS. Interestingly, this class has a lower than average chance to be familiar with ADAS. It might be possible that citizens who are familiar with ADAS know the limitations of these systems and perceive a higher risk with higher automation levels.

In conclusion, this class distinguishes itself for having a high social acceptance, especially in number of fatalities. Members of this class are also more likely to be male. Finally, they prefer high automation levels. Therefore, this class is categorized as automated driving enthusiasts.

Class 2: The Central Mass

Class 2 can be seen as the average class. This class show similar results to the MNL model of the whole sample. Their 'least-impact' design has a utility of 0.77 and is therefore socially accepted by 68% of this class. In contrary of class 1, they prefer automation level 3 over level 4. Even more so, they have a strong dislike for automation level 4, almost twice as strong as the preference of class 1 for automation level 4. Class 2 also has a slightly stronger dislike for fatalities by AVs than class 1.

This class is also the largest class with more than half of all citizens assigned to it. Citizens who are familiar with ADAS have a higher chance to be assigned to this class. Becker & Axhausen (2016) found a positive correlation between having higher automation levels in current vehicles (mostly ADAS) and intention-to-use and WTP for AVs. This research found that on average this is the case, but not for the automated driving enthusiasts. The social acceptance of this class is very similar to the average citizen as measured with the MNL model. Therefore, this class is named the central mass.

Class 3: Risk-averse

The final class has a very low social acceptance of ADS. Members of this class will vote against ADS in most cases. The 'least-impact' design has a utility of 0.78 and is socially accepted by only 5% of this class. They also have the highest penalty for fatalities, which categorizes them as risk-averse. What is interesting is that their parameter for emissions did not become significant. It might be possible that this class solely looked at number of fatalities and that environmental impact did not matter.

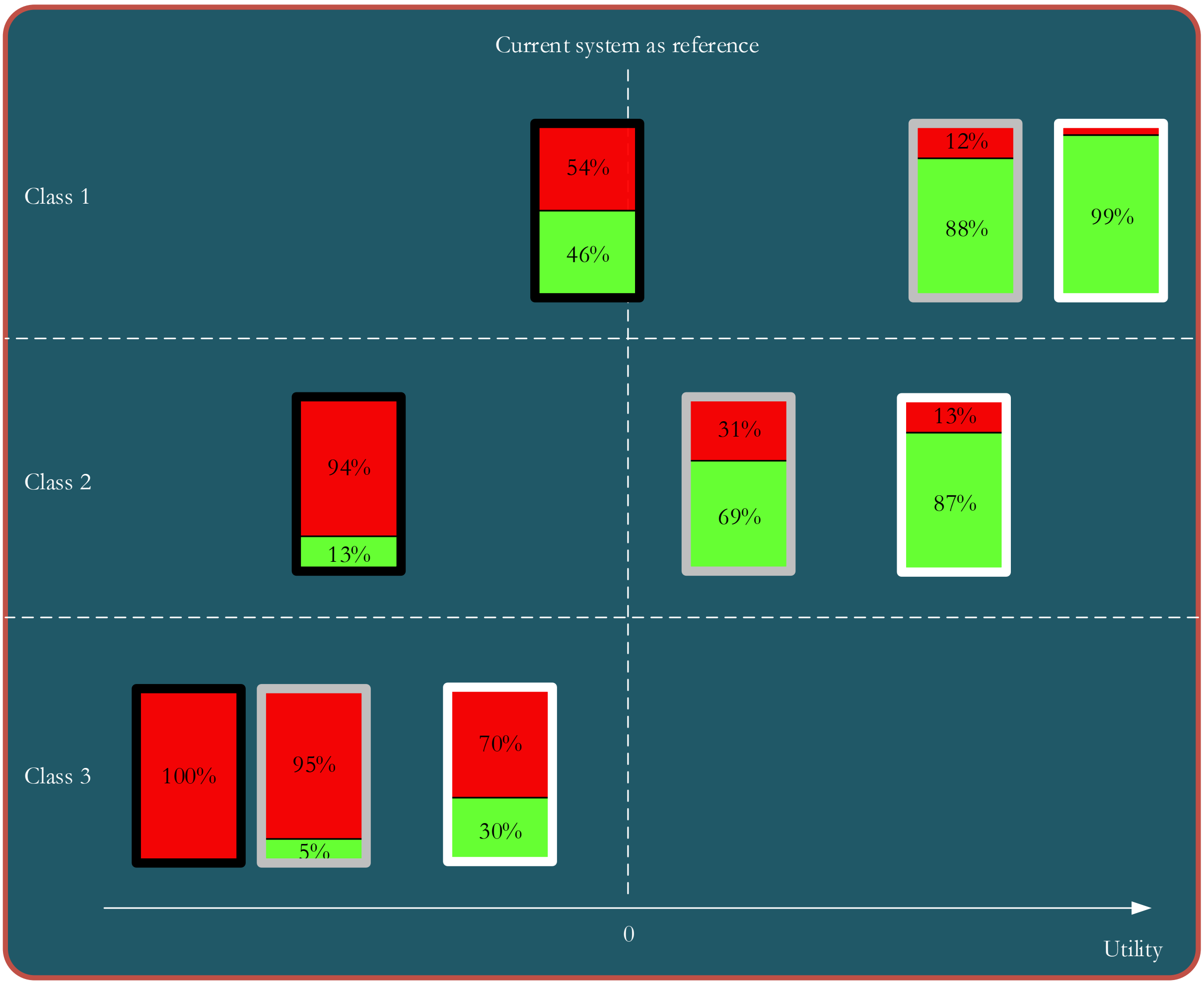


Figure 11: Social Acceptance of worst-, least-impact-, and best Design for different Classes

Finally, this is the smallest class with only 16% of citizens assigned to it. However, small groups should never be underestimated in society. Females have a relatively high chance to be assigned to this class, which coincides with the findings of Kyriakidis et al. (2015).

Class Overview

Figure 11 shows the social acceptance of the worst design (black), least-impact design (grey) and best design (white) for all three classes. This figure very clearly shows the heterogeneity among citizens in the social acceptance of ADS: Three classes with a very different preference for the same design. Both class 1 and class 2 have a higher social acceptance for the 'least-impact' design than the current transport system.

Especially the differences in dislike for fatalities are deemed interesting. Table 15 shows the factors between the different types of fatalities. The parameter for deliberate misuse fatalities of the risk-averse class is approximately 3 times larger than for automated driving enthusiasts.

Table 15: Differences in dislike for fatalities between classes

	ADS Enthusiasts	Central mass	Risk-averse
Human error versus technical failure	2.7	5	5.5
Human error versus deliberate misuse	3.4	6.5	9.8

It is concluded that there is a large heterogeneity among citizens in the social acceptance of ADS. This research shows that citizens can be assigned to three classes: automated driving enthusiasts, the central mass and risk-averse class.

5. Conclusion

Automated driving systems become a growing interest for policy makers, engineers and society in general. This research argued that the implementation of automated driving systems can lead to wicked problems for which responsible innovation can be used to bridge the gap between normative- and technical issues. To contribute to responsible innovation, this research conducted an empirical experiment to measure the social acceptance of ADS. This chapter will formulate the conclusions of this research together with recommendations for policy makers, researchers and engineers.

1. *What percentage of citizens thinks automated driving systems are socially accepted?*

To determine the social acceptance of different ADS designs, the current transport system was used as a reference. Also, a 'least-impact' design is formulated, which entails no social impacts, automation level 3 and only allows automated driving on highways. The 'least-impact' design was socially accepted by 63% of all citizens. It is interpreted that citizens are rather positive towards ADS. Even so, negative social impacts can strongly reduce the social acceptance.

Citizens generally prefer automation level 3 over automation level 4. The preference for lower automation levels coincides with literature (Becker & Axhausen, 2016). Most citizens do not seem ready to completely trust AVs. The contradiction is that the common understanding in research is that automation level 3 can be very dangerous (De Winter et al., 2014). From a safety perspective, it is possible that only automation level 4 will be allowed on public roads. However, level 4 is more challenging and could therefore delay the implementation.

In conclusion, citizens have a high social acceptance for ADS. The results show that ADS do not have to be safer than the current system to reach a higher social acceptance: the 'least-impact' design is not safer, yet it is accepted by 63% of all citizens. With the insights of this research, different designs in automation levels, safety, accessibility and environmental impact can be tested. Researchers and technology developers can predict the social impacts of a certain design and determine if their expected design is more socially accepted than the current system.

2. *How is the social acceptance influenced by safety, accessibility and environmental impact?*

According to this research, travel time is the least important attribute that was taken into consideration. Deliberate misuse fatalities were the most important, followed by technical failure fatalities, environmental impact and human error fatalities.

It is concluded that the social acceptance is mostly influenced by safety, particularly fatalities caused by AVs. Even so, the differences are not substantial: the social acceptance is influenced by all three social impacts and they should all be taken into consideration for the implementation of ADS.

3. *Are traffic fatalities caused by automated vehicles valued differently than current traffic fatalities?*

It is concluded that a technical failure fatalities weighs as much as 4 human error fatalities. Between deliberate misuse fatalities and human error fatalities is a factor 5.5. Consistently throughout the modelling of the results, this effect became statistically significant. The results coincide with literature (Kyriakidis et al., 2015).

The magnitude of the difference between these factors is larger than expected, which might causes problems for the implementation of ADS. The results imply that AVs have to cause hardly any fatalities to reach social acceptance even if they cause much less human error fatalities. Given these results, it is questionable if the Netherlands should be a test-bed for ADS. Although safety measures are in place, these experiments still propose risks to citizens, which are currently not accepted. It also makes the ethical acceptability of such experiments questionable at least. Only a couple of accidents might cause societal pressure to stop experiments and delay the implementation of ADS for many years.

The paradox is that ADS can be safer in the future, but that a potentially risky transition phase is needed to reach this future. Citizens prefer a future that includes ADS, but do not want an unsafe learning period. More on this discrepancy is in the recommendation.

4. *Is there heterogeneity in the social acceptance among citizens?*

Finally, it is concluded that there is a large heterogeneity in the social acceptance of ADS. This research has shown that citizens can be segmented into three classes: automated driving enthusiasts, the central mass and risk-averse class.

32% of citizens can be assigned to automated driving enthusiasts and these citizens have a preference for automation level 4. Interestingly, this is contradictory to the average citizen. This class has relatively low penalties for fatalities. Even so, technical failure fatalities (deliberate misuse fatalities) still weigh as much as 2.7 (3.4) human error fatalities. Automated driving enthusiasts are more likely to be male and to be less familiar with ADAS.

The central mass shows similar results as the average citizen. This class contains 52% of all citizens who distinguish themselves by a large dislike for automation level 4. They weigh technical failure fatalities (deliberate misuse fatalities) as much as 5 (6.5) human error fatalities. Also, they are more likely to be familiar with ADAS. Possibly they know the limitations of ADAS and therefore prefer a slow increase in automation levels.

Finally, the risk-averse class has a low social acceptance and every fatality is strongly penalized. Members of this class weigh technical failure fatalities (deliberate misuse fatalities) as much as 5.5 (9.8) human error fatalities. They are more likely to be female and it is also the smallest class with only 16% assigned to it.

The progress that has to be made in safety is much larger for the risk-averse class than for automated driving enthusiasts. Assuming that some fatalities caused by AVs are inevitable, automated driving enthusiasts would accept the overall safety record of ADS approximately three times more likely than the risk-averse class. It might imply a difference of years or even decades when ADS are socially accepted by the two classes. Also, some automated driving enthusiasts are likely to use AVs, thus proposing risks to society. Although these users might accept these risks, these same risks are not accepted by the risk-averse class or even the central mass. Therefore, the heterogeneity is critical for the implementation of ADS and is discussed in the recommendations.

Recommendations

As following from responsible innovation and Value-Sensitive design framework and in particular its focus on empirical research as one of its necessary elements, an effort is made to provide more insights in the social acceptance of automated driving systems. The main finding throughout this research is that citizens have a relatively high social acceptance of ADS. Even without benefits, citizens generally prefer ADS over the current transport system. It is therefore recommended to policy makers to have a positive approach to ADS and to possibly allow more AVs on the road. That being said, there are a couple of considerations for policy makers based on discrepancies between technical- and normative issues: 1) High social acceptance versus strong dislike for fatalities caused by AVs; 2) Citizens who are enthusiastic about ADS versus citizens who are risk-averse.

High social acceptance versus strong dislike for fatalities caused by AVs

The discrepancy is that citizens have a high social acceptance for ADS, but only if AVs would be very safe. However, AVs are not safe until they learn how to drive, much like people who get their driving licence. So, policy makers have a moral obligation to promote the implementation of ADS to reach a safer future, while they also have a moral obligation to prevent unpredictable risks in ADS experiments. How the learning curve is achieved is therefore considered a critical factor in the implementation of ADS. Too little experiments means AVs will never learn how to drive and too many experiments means too much unpredictable risks.

It is therefore recommended to be extra cautious during the transition phase especially with experiments for private vehicles. Current experiments are largely based on professional users, like freight transport and public transport. Risks can be easier contained for these users, for example by using segregated lanes (WE pods) or supervision by trained professionals (truck platooning). For private vehicles it might be possible to experiment in traffic jams. Low speeds in traffic jams ensure minimal impact in case of an accident. Either way, the risk of fatalities caused by AVs should be minimized.

Also, policy makers should decide if AVs may use cooperate driving, i.e. be connected to other vehicles, infrastructure or a cloud. Shladover (2016) argues that automated driving is not possible without cooperate driving. However, these connections make AVs vulnerable for deliberate misuse which is strongly penalized by citizens. The experiment by Bridle (2017) showed that with only a package of salt he could alter the road markings read by AVs. Imagine what a hacker can do by altering 3D-maps or other environmental information of AVs. The report of Roland Berger (2017) indicates that in the Netherlands, as well as globally, research in cooperate driving is lacking. If ADS cannot exist without cooperate driving, more knowledge on this subject is needed. It is therefore recommended to conduct more research into cooperate driving. Possibly, this should lead to restrictions for technology producers like TomTom to safely deploy cooperate driving technologies.

Citizens who are enthusiastic about ADS versus citizens who are risk-averse

This research has found large heterogeneity among citizens, especially in the acceptance of fatalities caused by AVs. The Vienna Convention (1986) states that a human driver has to be in control of its vehicle and its environment. However, it is likely that some automated driving enthusiasts will use the automated driving features since they prefer high automation levels. Consequently, they will not be in complete control of the vehicle and dangerous situations may arise. When lower priced vehicle will obtain automated driving capabilities, the technology will be available to a larger public and more enthusiasts will be able to experiment with the technology. In turn, this might inflict an unaccepted situation, especially for citizens who are risk-averse.

Therefore, a passive approach to AVs cannot be made. Currently, there are no limitations to automated driving technologies in vehicles, so it is recommended to review the licencing of such vehicles. The Ministry of Infrastructure and the Environment and the RDW have to actively engage with technology producers whose technologies are available to consumers. Even if these vehicles are licensed, restrictions should be made to the technology. For example, make sure automated driving technologies are switched off above the speed of 50 km/h; or make sure the technology cannot be switched on uninterruptedly for long periods of time. For private vehicles, it is recommended to stick to direct human control for the transition phase. As the technology progresses, direct human control may be widened to meaningful human control, thus including some kind of delegation.

Also, the social acceptance is partly based on incomplete or even false information. So even if two individuals have the same norms and values, they might have a different social acceptance of ADS based on different information they received. Therefore policy makers can use campaigns to provide information to citizens about ADS. If policy makers and engineers believe that social impacts are positive and philosophers believe ADS is ethically acceptable, campaigns can be used to make sure citizens are informed about the social impacts and way of implementation. For example, automated driving enthusiast can be made aware of the risks of automated driving thus making sure they use the technology more responsibly; or the risk-averse class and central mass can be made aware of overall safety benefits, changing their acceptance of fatalities caused by AVs. Either way, these campaigns might ensure citizens that their norms and values are embedded in the design of ADS.

Automated driving systems are set to be one of the more disruptive technologies for the coming future. The way we see and experience transport is very likely to change in the coming decades. If we can align research, development and design to the social acceptance and ethical acceptability of the innovation process and its marketable products, and if we can align the corresponding actors in this field, we can make sure that automated driving systems become a responsible innovation. This research suggests that a small step towards this goal can be made by following these recommendations:

- Policy makers should have a positive and active approach towards ADS and continue with ADS experiments, primarily for professional users;
- Technology producers and policy makers should intensify research into cooperate driving;
- Policy makers (especially RDW) should review the licensing of AVs;
- General: use information campaigns about the risks and benefits of ADS.

6. Reflection

The final chapter will reflect on the methodology and the results of this research.

Firstly, in making the survey there was a struggle to inform respondents sufficiently to have reasoned answers, but not too much to make them biased. It is well known that not all respondents have a certain level of understanding of ADS, which could result in meaningless answers. Overall, the results showed similarity with other publications which validates the findings.

Also, a lot of technical- and normative issues relate to travel behaviour in mixed traffic. AVs are still so new that hardly any experience exists in the communication and cooperation between automated and non-automated traffic. Consequently, citizens may not have an opinion on road exemption or level of automation, simply because they have no experience with the effects of such design requirements. Currently, the SWOV conducts a research into possible problems in mixed traffic in real-life situations. More research into mixed traffic factors can shed a brighter light on the social acceptance and ethical acceptability by providing more insights in interaction and responsibilities of road users.

Thirdly, the social acceptance as measured in this research is the current status-quo. This acceptance is likely to change based on different factors like time. Therefore, more research is needed how the acceptance will change and how rapidly this change will come. Moreover, one specific issue is that ADS will mostly likely be used by future generations. These generations are not included in the social acceptance of this research, but should be included in the ethical acceptability of ADS. Preferably, the social acceptance is measured frequently in the future and linked to the ethical acceptability by ethics scholars.

Finally, terrorists found a new terrifying method in their attacks by driving cars into pedestrians. Terrorists could also use AVs as 'killing machines'. This research has therefore focused on deliberate misuse fatalities and the social acceptance thereof. However, ADS could also prevent terrorist attacks like these. The truck that was used in the Christmas fair in Berlin was stopped prematurely because of an automated brake system (The Telegraph, 2016). If AVs are programmed to stop if they are about to inevitably hit something which cannot be overruled by humans, lives could actually be saved. This aspect of ADS is not taken into consideration for this research.

7. References

Abraham, H., Lee, C., Brady, S., Fitzgerald, C., Mehler, B., Reimer, B., & Coughlin, J. (2017). Autonomous Vehicles, Trust, and Driving Alternatives: a Survey of Consumer Preferences. *TRB Annual Meeting*, 1–14.

Anderson, J. M., Kalra, N., Stanley, K. D., Sorensen, P., Samaras, C., & Oluwatola, O. A. (2014). *Autonomous Vehicle Technology - A Guide for Policymakers*. *Rand*. http://doi.org/10.7249/RR443-2

auto-en-vervoer.infonu.nl. (2015). Tien redenen waardoor verkeersongevallen worden veroorzaakt.

Bakermans, B. A. (2016). *Truck platooning: Enablers, Barriers, Potential and Impacts*. Delft University of Technology.

Bansal, P., Kockelman, K. M., & Signh, A. (2016). ASSESSING PUBLIC OPINIONS OF AND INTEREST IN NEW VEHICLE 2 TECHNOLOGIES: AN AUSTIN PERSPECTIVE. The University of Texas at Austin.

Bax, A. (2012). *Historie en toekomst: ruim honderd jaar verkeersveiligheidsbeleid en -kennis in Nederland*. *SWOV*. Leidschendam.

Becker, F., & Axhausen, K. W. (2016). Literature Review on Behavioral Experiments for Autonomous Vehicles. *TRB Annual Meeting*.

Behere, S., Törngren, M., & Chen, D. J. (2013). A reference architecture for cooperative driving. *Journal of Systems Architecture*, *59*(10 PART C), 1095–1112. http://doi.org/10.1016/j.sysarc.2013.05.014

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The Social Dilemma of Autonomous Vehicles. *Science*, *352*(6293).

Bos, A., Groningen, H. van, Mom, G., & Vinne, V. van der. (1996). Het Paardloze Voertuig: De auto in Nederland een Eeuw geleden. Kluwer Bv.

Bridle, J. (2017). Autonomous Trap 001.

CBS. (2017). Bevolking; geslacht, leeftijd, burgerlijke staat en regio, 1 januari.

CBS.nl. (2013). Beroepsbevolking; behaalde onderwijs naar persoonskenmerken 2001-2012.

CBS.nl. (2017). Overledenen; belangrijke doodsoorzaken (korte lijst), leeftijd, geslacht. Den Haag/Heerlen.

Chorus, C. G. (2017). *SEN1221 Statistical analysis of choice behavior*. *Textbook*.

Correia, G. H. de A., & van Arem, B. (2016). Solving the User Optimum Privately Owned Automated Vehicles Assignment Problem (UO-POAVAP): A model to explore the impacts of self-driving vehicles on urban mobility. *Transportation Research Part B: Methodological*, *87*, 64–88. http://doi.org/10.1016/j.trb.2016.03.002

Cranenburgh, S. van. (2017). SEN1721; Latent Class Discrete Choice Models for Travel Behaviour Research. Delft: TU Delft.

Cuppen, E. (2012). Diversity and constructive conflict in stakeholder dialogue: Considerations for design and methods. *Policy Sciences*, *45*(1), 23–46. http://doi.org/10.1007/s11077-011-9141-7

Davies, A. (2017). The Numbers Don't Lie: Self-Driving Cars Are Getting Good. *Wired*.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340.

De Financiele Telegraaf. (2017). Intel koopt Mobileye voor $15 miljard.

De Vlieger, I., De Keukeleere, D., & Kretzschmar, J. G. (2000). Environmental e!ects of driving behaviour and congestion related to passenger cars. *Atmospheric Environment*, *34*, 4649–4655. http://doi.org/10.1016/S1352-2310(00)00217-X

De Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(PB), 196–217. http://doi.org/10.1016/j.trf.2014.06.016

European Commission. (2015). Reducing CO2 emissions from passenger cars. *European Commision*. Retrieved from http://ec.europa.eu/clima/policies/transport/vehicles/cars_en#tab-0-0

European Union. Declaration of Amsterdam: Cooperation in the field of connected and automated driving (2016). Amsterdam: European Union. Retrieved from http://english.eu2016.nl/binaries/eu2016-en/documents/publications/2016/04/14/declaration-of-amsterdam/2016-04-08-declaration-of-amsterdam-final-format-3.pdf

EuroScientist. (2017). Responsible Science. EuroScientist journal.

Fish, M. J. (2014). Tolerance, Acceptance, Understanding.

Foot, P. (1978). The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Review*, *5*.

Friedman, B., Kahn Jr., P. H., & Borning, A. (2006). Value Sensitive Design and Information Systems (PREPRINT). *Human-Computer Interaction and Management Information Systems: Foundations*, 1–27. http://doi.org/10.1145/242485.242493

GfK. (2016). Duurzaamheid bij consument in de lift.

Green, D., Jacowitz, K. E., Kahneman, D., & McFadden, D. (1998). Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics*, *20*(2), 85–116. http://doi.org/10.1016/S0928-7655(97)00031-6

Greenemeier, L. (2016). Deadly Tesla Crash Exposes Confusion over Automated Driving. *Scientific American*. Retrieved from http://www.scientificamerican.com/article/deadly-tesla-crash-exposes-confusion-over-automated-driving/?print=true on 3rd Oct. 2016

Hagenzieker, M. P. " That bollard could have been a child ". On traffic safety and human behaviour in traffic, Inaugural Speech, Delft University of Technology (2015).

Hensher, D. A., & Li, Z. (2013). Referendum voting in road pricing reform: A review of the evidence. *Transport Policy*, *25*, 186–197. http://doi.org/10.1016/j.tranpol.2012.11.012

Jamson, A. H., Merat, N., Carsten, O. M. J., & Lai, F. C. H. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. *Transportation Research Part C: Emerging Technologies*, *30*, 116–125. http://doi.org/10.1016/j.trc.2013.02.008

Kasteleijn, N. (2017). TomTom wil zelfrijdende auto's straks de weg wijzen. NOS.

KiM. (2017). *Paths to a self-driving future*. The Hague: Netherlands Institute for Transport Policy Analysis.

Kyriakidis, M., Happee, R., & De Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 127–140. http://doi.org/10.1016/j.trf.2015.04.014

Lavasani, M., Asgari, H., Jin, X., & Pinjari, A. (2017). Investigating Willingness to Pay for Autonomous Vehicles and the Likelihood of Residential Relocation. *TRB Annual Meeting*, (813).

Lavieri, P. S., Garikapati, V. M., Bhat, C. R., Pendyala, R. M., Astroza, S., & Dias, F. F. (2017). Modeling Individual Preferences for Ownership and Sharing of Autonomous Vehicle Technologies. *TRB Annual Meeting*.

Loomes, G., & Sugden, R. (1983). Regret Theory and Measurable Utility. *Economic Letters*. University of Newcastly upon Tyne: North-Holland Publishing Company.

Loukas, G. (2015). 3 – Cyber-Physical Attacks on Implants and Vehicles. *Cyber-Physical Attacks*, 59–104. http://doi.org/10.1016/B978-0-12-801290-1.00003-5

Madden, T., Ellen, P., & Ajzen, I. (1992). A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action. *Personality and Social Psychology Bulletin*, *18*(1), 3–9. http://doi.org/0803973233

McFadden, D. L. (2000). Economic Choices. Berkeley: Department of Economics, University of California.

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(PB), 274–282. http://doi.org/10.1016/j.trf.2014.09.005

Milieudefensie. (n.d.). Luchtvervuiling en de gevolgen voor je gezondheid.

Ministerie van Infrastructuur en Milieu. (2017). Nationale Markt- Markt en Capaciteitsanalyse 2017 (NMCA) Hoofdrapport, *2017*, 1–62.

Mouter, N., & Chorus, C. (2016). Trade-off reistijd, wegenbelasting, verkeersveiligheid.

Nihlén Fahlquist, J. (2006). Responsibility Ascriptions and Vision Zero. *Accident Analysis & Prevention*, *38*(6), 1113–1118.

NOS. (2016). Video: Tesla reageert razendsnel op snelweg. Retrieved from http://nos.nl/artikel/2150431-video-tesla-reageert-razendsnel-op-snelweg.html

NTSB. (2017). *Collision between a Car Operation with Automated Vehicle Control Systems and a Tractor-Semitrailer Truck*. Williston, Florida.

nu.nl. (2017). Nieuwe Audi A8 rijdt bijna geheel zelfstandig.

NWO. (2015). Responsible innovation programme to focus on all top sectors from 2016 onwards.

Oatman-Stanford, H. (2014). Murder Machines: Why Cars Will Kill 30,000 Americans This Year.

Oomen, J. (2005). Verbeterde parkshuttle rijdt weer tussen rotterdam en capelle aan den ijssel. *De Ingenieur*, *20*(November), 2005.

Pearsall, J., & Hanks, P. (2001). *The new Oxford dictionary of English*. Oxford: Oxford University Press.

Poel, van de. (2016). Responsible innovation. *Delft University of Technology, SPM9239*. http://doi.org/10.1080/23299460.2015.1059010

Presidency of the Council of the European Union. (2014). Rome Declaration on Responsible Research and Innovation in Europe, (November), 1–2.

Roland Berger. (2017). Automotive Disruption Radar: Tracking disruption signals in the automotive industry, (April).

RVO. (2017). Cijfers elektrisch vervoer.

SAE International. (2014). Automated Driving, Levels of Driving Automation are defined in new SAE

International Standard J3016.

Santoni de Sio, F. (2016). *Ethics and Self-driving Cars; A White Paper on Responsible Innovation in Automated Driving Systems*.

Shladover, S. (2016). Progress toward automated driving - Halmstad Colloquium.

STAD. (2017). Project Partners: Spatial and Transport impacts of Automated Driving. Retrieved from http://scholar.google.com/scholar?start=80&q=author:Tazim+author:Jamal&hl=en&as_sdt=0,5#5

STAD Event. (n.d.). Automated Driving: a silver bullet for urban regions?

SWOV. (2013). *SWOV-Factsheet, Fietsers.*

SWOV. (2016). *Verkeersdoden in Nederland. SWOV-Factsheet.*

SWOV. (2017a). Dutch road safety in international perspective.

SWOV. (2017b). Verkeersdoden in Nederland. Den Haag.

Taebi, B. (2016). Bridging the Gap between Social Acceptance and Ethical Acceptability. *Risk Analysis*, *forthcomin*(3), 1–21. http://doi.org/10.1111/risa.12734

Telegraaf. (2017). Tesla die op vrachtwagen A1 botste reed zelf.

The Guardian. (2016). Two years until self-driving cars are on the road – is Elon Musk right? *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/jun/02/self-driving-car-elon-musk-tech-predictions-tesla-google

The Telegraph. (2016). Advanced braking system in truck used in Berlin Christmas market terrorist attack prevented even worse carnage.

Timmermans, H. (1980). Functionele hierarchie en ruimtelijk keuzegedrag van consumenten: een theoretische uiteenzetting en een empirische toepassing van een ruimtelijk preferentie model. *K.N.A.G. Geografisch Tijdschrift*, *14*(3), 182–193.

TNO. (2016). Economische Wegwijzer; Alarmerende stijging fileschade, 2015–2016.

United Nations. (2017). UNTC Chapter XXVII; Environment; Paris agreement.

Van Asselt Marjolein, B. A., & Rijkens-klomp, N. (2002). A look in the mirror : reflection on participation in Integrated Assessment from a methodological perspective, *12*, 167–184.

van den Berg, V. A. C., & Verhoef, E. T. (2016). Autonomous cars and dynamic bottleneck congestion: The effects on capacity, value of time and preference heterogeneity. *Transportation Research Part B: Methodological*, *94*, 43–60. http://doi.org/10.1016/j.trb.2016.08.018

van den Hoven, J. (n.d.). Responsible Innovation in brief. Delft University of Technology.

van Loon, R. J., & Martens, M. H. (2015). Automated Driving and its Effect on the Safety Ecosystem: How do Compatibility Issues Affect the Transition Period? *Procedia Manufacturing*, *3*(Ahfe), 3280–3285. http://doi.org/10.1016/j.promfg.2015.07.401

Van Wee, B., Annema, J. A., & Banister, D. (2013). The transport system and transport policy: an introduction, 399.

Verkeersinformatiedienst. (2016). Filedruk neemt in 2016 met 10 % toe.

Vienna Convention. (1986). *CONVENTION ON ROAD TRAFFIC*. Vienna: UNECE.

Vinne, V. van der. (2007). De trage verbreiding van de auto in Nederland 1896 - 1939. Amsterdam: De Bataafsche Leew.

von Schomberg, R. (2011). Prospects for Technology Assessment in a framework of responsible research and innovation. *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methode*, 39–61.

Wees, K. Van, & Brookhuis, K. A. (2005). Product liability for ADAS; legal and human factors perspectives. *European Journal of Transport and Infrastructure Research*, *5*, 357–372.

WEpods.nl. (n.d.). Homepage - GeneDB. Retrieved from http://www.genedb.org/Homepage

Zakharenko, R. (2015). Self-Driving Cars Will Change Cities. *SSRN Electronic Journal*, *61*(October 2015), 26–37. http://doi.org/10.2139/ssrn.2683312

APPENDICES

I. Coding of Socio-demographics

Attribute	Attribute level	Coding						
Gender	Male	-1						
	Female	1						
Age	18-30	0						
	31-40	1						
	41-50	2						
	51-60	3						
	61-70	4						
	71-80	5						
	80+	6						
Level of education		edu2	edu3	edu4				
	Basisonderwijs	0	0	0				
	Voortgezet onderwijs	1	0	0				
	MBO	0	1	0				
	HBO, universiteit	0	0	1				
Residential area		res2	res3	res4	res5			
	Rural area	0	0	0	0			
	Village	1	0	0	0			
	Suburbs	0	1	0	0			
	City	0	0	1	0			
	City centre	0	0	0	1			
Licence	Yes	0						
	no	1						
		mod2	mod3	mod4	mod5	mod6	mod7	mod8
Modal choice	(almost) every day	0	0	0	0	0	0	0
	5-6 days per week	1	0	0	0	0	0	0
	3-4 days per week	0	1	0	0	0	0	0
	1-2 days per week	0	0	1	0	0	0	0
	1-3 days per month	0	0	0	1	0	0	0
	6-11 days per year	0	0	0	0	1	0	0
	1-5 days per year	0	0	0	0	0	1	0
	Less than 1 day per year	0	0	0	0	0	0	1
Usage ADAS1	Yes	1						
	No	0						
Usage ADAS2	Yes	1						
	No	0						
Accident experience1	Yes	1						
	No	0						
Accident experience2	Yes	1						
	No	0						

Environmentalist	Not at all	0
	Hardly	1
	Neutral	2
	To some extent	3
	Very	4
Technology lovers	Not at all	0
	Hardly	1
	Neutral	2
	To some extent	3
	Very	4

II. Ngene design

Design								
Choice situation	ads.rex	ads.aut	ads.tc	ads.emi	ads.hum	ads.tec	ads.mis	Block
1	80	3	-25	-25	300	80	30	1
2	50	3	0	0	300	40	0	1
3	120	3	25	25	300	0	60	1
4	80	4	-25	25	500	0	0	2
5	50	3	25	25	400	80	60	2
6	120	3	-25	-25	400	40	30	2
7	80	3	0	0	400	0	0	2
8	120	4	0	-25	300	0	60	2
9	80	4	25	0	300	80	30	2
10	50	4	-25	25	300	40	0	2
11	120	4	0	25	400	40	0	1
12	80	4	25	-25	400	0	60	1
13	50	4	-25	0	400	80	30	1
14	80	3	25	-25	500	40	0	1
15	50	3	-25	0	500	0	60	1
16	120	3	0	25	500	80	30	1
17	50	3	25	25	500	40	30	2
18	120	3	-25	-25	500	0	0	2
19	120	4	-25	0	300	0	30	1
20	80	4	0	25	300	80	0	1
21	50	4	25	-25	300	40	60	1
22	50	3	-25	-25	400	80	0	2
23	120	3	0	0	400	40	60	2
24	80	3	25	25	400	0	30	2

III. Introduction to Survey, English Version

Page 1

Dear sir/madam,

Thank you for your participation in this survey.

De technology for self-driving vehicles is rapidly advancing. It may be possible that in the coming 5 to 20 years self-driving vehicles will be on the public roads. Self-driving vehicles can have a major impact, among others in the field of traffic safety, accessibility and the environment. Therefore, it is very important that we as society think how we want that this future will look like. The goal of this survey is to collect as many opinions of the self-driving car to create a desirable future for everyone. Your feedback is therefore highly appreciated!

This survey is meant for both car users as other road users with the age of 18 and up.

It will take you about 5-10 minutes. Your answers will be handled anonymously. In this study, no individual answers are reported, only averages.

Thank you in advance and good luck!

Bart Overakker

For questions: bartoverakker@gmail.com

Page 2

This research will outline some possible future prospects. They will include self-driving vehicles that will drive semi- or fully automated on certain roads.

Semi-automated: The car is self-driving, but the driver must pay attention. He/she may not conduct other tasks like reading or sleeping, since the driver must be able to intervene in an emergency situation.

Fully automated: The car is self-driving and the driver may conduct other tasks like reading or sleeping. The technology is thus fully trusted.

Highway: 100-130 km/h, separated roadways and no same-level crossings.

Regional roads: 50-80 km/h, non-separated roadways and same-level crossing with other car traffic.

Inside build environment: 30-50 km/h, same-level crossing with cars, cyclist and pedestrians.

Page 3

The self-driving vehicle affects all road users. Two examples:

- When self-driving vehicles will cause more congestion on highways, the travel costs and emissions of all road users will increase;
- When self-driving vehicles are better in prevention accidents than human drivers, than the traffic safety will increase for all road users, including cyclists, pedestrians and other car users.

In this study, the following effects are distinguished:

Travel costs: The average travel costs for all road users

Emissions: The average emissions of all road users

Victims of human error: The number of annual traffic deaths in the Netherlands due to human error, like driving under influence

Victims of technical failure of the self-driving vehicle: The number of annual traffic deaths in the Netherlands due to a software or hardware failure of a self-driving vehicle

Victims of deliberate misuse of the self-driving vehicle: The number of annual traffic deaths in the Netherlands due to deliberate hacking with bad intentions of a self-driving vehicle

Page 4

We ask you to vote 12 times for or against a possible future with the self-driving vehicle.

Are you for, than we ask you to imagine this future will become a reality.

Are you against, than the future will not become reality and the situation stays as it is.

The self-driving vehicle also influences other factors and some future prospects might some seem realistic to you. However, we do ask you kindly to ignore other factors and try to make a considered decision. There are no 'good' or 'wrong' answers, we just want to know what the preferences are of Dutch society with respect to self-driving vehicles.

IV. Complete Survey Dutch Version

Geachte heer/mevrouw,

Dank voor uw deelname aan dit onderzoek.

De technologie voor zelfrijdende auto's ontwikkelt zich in hoog tempo. Het is daarom mogelijk dat er in de komende 5 tot 20 jaar zelfrijdende auto's op de openbare weg gaan rijden. Volgens veel mensen kunnen de zelfrijdende auto's grote effecten hebben, onder andere op het gebied van verkeersveiligheid, bereikbaarheid en milieu. Het is daarom erg belangrijk dat wij als samenleving nadenken hoe wij willen dat deze toekomst eruit gaat zien. Het doel van deze vragenlijst is om erachter te komen hoe u en de rest van de samenleving over zelfrijdende auto's denken. Uw deelname wordt daarom zeer gewaardeerd!

Deze vragenlijst is onderdeel van een MSc Thesis aan de Technische Universiteit Delft en is bedoeld voor zowel autogebruikers als andere weggebruikers van 18 jaar of ouder.

Het invullen van de vragen kost u ongeveer 10 minuten. Uw antwoorden worden volledig anoniem verwerkt. In deze studie worden geen individuele antwoorden gerapporteerd, alleen gemiddelden.

Alvast bedankt en succes met het invullen!

Bart Overakker

Voor vragen: bartoverakker@gmail.com

In dit onderzoek wordt een aantal toekomstbeelden geschetst met mogelijke effecten van zelfrijdende auto's. Wij vragen of u vóór of tegen deze mogelijke toekomst bent. Met deze toekomstbeelden worden zelfrijdende auto's bedoeld die half- of volledig geautomatiseerd rijden op bepaalde wegen. De volgende wegen worden onderscheiden:

Snelweg	**Provinciale weg**	**Binnen de bebouwde kom**
100 - 130 km/h	50 - 80 km/h	30 - 50 km/h
Gescheiden rijrichtingen en geen gelijkvloerse kruisingen	Ongescheiden rijrichtingen en gelijkvloerse kruisingen met autoverkeer	Ongescheiden rijrichtingen en gelijkvloerse kruisingen met auto's, fietsers en voetgangers



Het volgende wordt bedoeld met half- of volledig geautomatiseerd:

Half geautomatiseerd	**Volledig geautomatiseerd**
De auto rijdt zelfstandig, maar de bestuurder moet blijven opletten. Hij/zij mag geen andere taken op zich nemen, zoals slapen of lezen, want de bestuurder moet kunnen ingrijpen in het geval van een noodsituatie.	De auto rijdt zelfstandig en de bestuurder kan andere taken op zich nemen dan het besturen van de auto, zoals slapen of lezen. Er wordt dus volledig vertrouwd op de technologie.

In dit onderzoek worden de volgende effecten onderscheiden, die gelden voor alle weggebruikers:

Effect	**Uitleg**
Reistijd	De gemiddelde reistijd voor alle weggebruikers. Denk hierbij aan de tijd die u kwijt bent aan fietsen, autorijden, openbaar vervoer, parkeren enzovoort.
Uitstoot	De gemiddelde uitstoot van uitlaatgassen van alle weggebruikers. Denk hierbij aan CO_2, NO_X, fijnstof enzovoort.
Verkeersdoden menselijke falen	Het aantal verkeersdoden per jaar in Nederland door menselijke falen, zoals afleiding en rijden onder invloed. In 2016 telde Nederland ruim 600 verkeersdoden met menselijk falen als oorzaak.
Verkeersdoden technisch falen zelfrijdende auto	Het aantal verkeersdoden per jaar in Nederland door een fout in de software of hardware van een zelfrijdende auto. Dit is een fout van de autofabrikant/ technologieleverancier.
Verkeersdoden opzettelijk misbruik zelfrijdende auto	Het aantal verkeersdoden per jaar in Nederland door het hacken van een zelfrijdende auto. Dit gebeurt opzettelijk door iemand met kwade bedoelingen.

Op de volgende pagina's worden 12 verschillende toekomstbeelden met zelfrijdende auto's geschetst. Wij vragen u zich uit te spreken of u voor of tegen dit toekomstbeeld bent. U kunt antwoorden alsof het een referendum is.

Wij wijzen u er nogmaals op dat zelfrijdende auto's invloed kunnen hebben op alle weggebruikers. Twee voorbeelden:

- Als zelfrijdende auto's zorgen voor meer of minder files op de snelwegen, dan zullen de reistijd en uitstoot van alle snelweggebruikers toenemen of afnemen;
- Als zelfrijdende auto's beter of slechter zijn in het voorkomen van ongelukken dan menselijke bestuurders, dan zal de veiligheid toenemen of afnemen voor alle weggebruikers, inclusief voetgangers, fietsers en autogebruikers.

Let op: Zelfrijdende auto's hebben ook invloed op andere factoren en sommige toekomstbeelden lijken u wellicht onlogisch. Wij vragen u toch vriendelijk om andere factoren te negeren en zo goed mogelijk een keuze te maken of u voor of tegen het voorgestelde toekomstbeeld bent. Er zijn geen 'goede' of 'foute' keuzes; wij willen alleen maar weten welke voorkeuren de Nederlandse bevolking heeft op het gebied van zelfrijdende auto's.

Wij vragen u of u voor of tegen het volgende toekomstbeeld bent.

In dit toekomstbeeld rijden zelfrijdende auto's uitsluitend...

...op de snelweg...

...en op de provinciale weg...

...en binnen de bebouwde kom...

...half geautomatiseerd.

Het effect van dit toekomstbeeld met zelfrijdende auto is...

De gemiddelde reistijd voor alle weggebruikers	25 % minder reistijd
De gemiddelde uitstoot voor alle weggebruikers	25 % meer uitstoot
Verkeersdoden per jaar menselijk falen	300 verkeersdoden minder
Verkeersdoden per jaar technisch falen zelfrijdende auto	40 verkeersdoden meer
Verkeersdoden per jaar opzettelijk misbruik zelfrijdende auto	Blijven 0 verkeersdoden

V. Mixed Logit Model Results

Below are the results of two mixed logit models. The first model explores the heterogeneity in the alternative specific constant. The second model explores the heterogeneity in the attribute road exemption.

The first model shows a large heterogeneity in the ASC. The Rho^2 statistic grows significantly from 0.045 to 0.155. It motivates to continue with the Latent Class Choice Model.

Number of draws	5
Number of estimated parameters	9
Sample size	6120
Number of individuals	510
Init log likelihood	-4242.061
Final log likelihood	-3585.331
Likelihood ratio test for the init. model	1313.460
Rho-square for the init. model	0.155
Rho-square-bar for the init. model	0.153
Final gradiant norm	+2.044e-002
Diagnostic	Convergence reached
Iterations	187
Run time	05:35

Table 16: Mixed Logit Model: ASC Results

Parameter	Name	Value	Standard Error	t-test	p-value
Alternative specific constant	B_ASC	2.57	0.227	11.28	0.00
Level of automation	B_aut	-0.499	0.076	-6.54	0.00
Road exemption	B_rex	-0.062	0.043	-1.42	0.16*
Travel time	B_tti	-0.497	0.044	-11.30	0.00
Emissions	B_emi	-0.512	0.050	-10.23	0.00
Human error fatalities	B_hum	-0.0029	0.000	-7.34	0.00
Technical failure fatalities	B_tec	-0.013	0.001	-11.68	0.00
Deliberate misuse fatalities	B_mis	-0.017	0.002	-8.97	0.00
Sigma ASC	SIGMA1	1.56	0.107	-14.62	0.00

Below are the results of the mixed logit model with an estimated sigma for road exemption. The sigma is not statistically significant. It means that road exemption did not play a significant role in the decision problem of citizens.

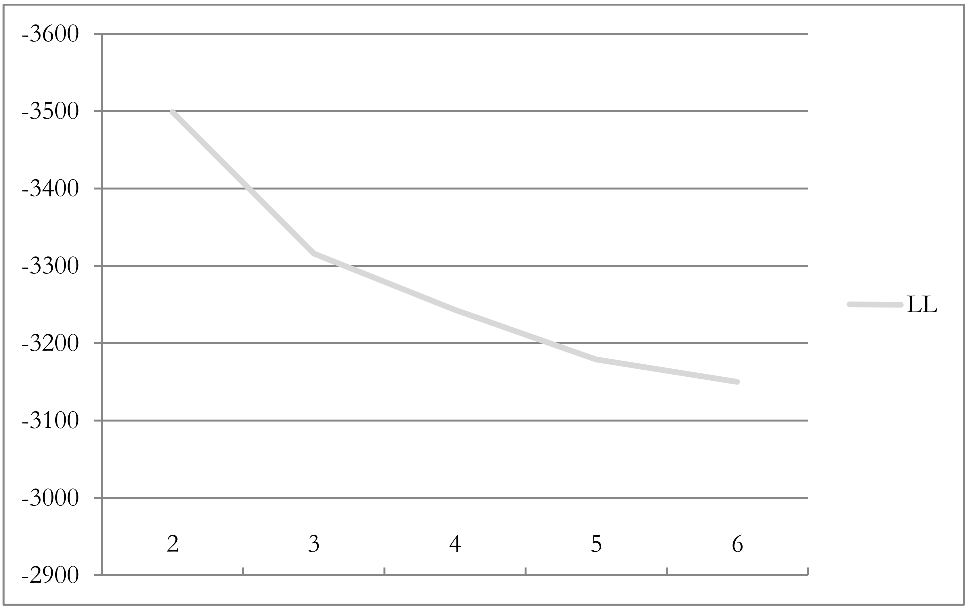
Number of draws	5
Number of estimated parameters	9
Sample size	6120
Number of individuals	510
Init log likelihood	-4242.061
Final log likelihood	-4049.619
Likelihood ratio test for the init. model	384.884
Rho-square for the init. model	0.045
Rho-square-bar for the init. model	0.043
Final gradiant norm	+3.568e-002
Diagnostic	Convergence reached
Iterations	12
Run time	00:22

Table 17: Mixed Logit Model: Road Exemption Results

Parameter	Name	Value	Standard Error	t-test	p-value
Alternative specific constant	B_ACS	1.92	0.198	9.70	0.00
Level of automation	B_aut	-0.371	0.053	-6.98	0.00
Road exemption	B_rex	-0.044	0.036	-1.23	0.22*
Travel time	B_tti	-0.373	0.037	-10.09	0.00
Emissions	B_emi	-0.385	0.040	-9.61	0.00
Human error fatalities	B_hum	-0.0022	0.0004	-5.59	0.00
Technical failure fatalities	B_tec	-0.0098	0.0009	-10.61	0.00
Deliberate misuse fatalities	B_mis	-0.0134	0.0018	-7.63	0.00
Sigma road exemption	SIGMA2	0.044	0.045	0.99	0.32*

VI. LC: Loglikelihood Function Compared to Number of Classes

This graph shows the loglikelihood of latent class discrete choice models with different number of classes (x-axis). There is a big bend at the 3 class model. It shows that the loglikehood improves much less after the 3 class model.



VII. 5 Class Latent Class Choice Model Results

Below are the results from the estimated 5 class latent choice model.

Number of estimated parameters	44
Sample size	6120
Excluded observations	0
Init log likelihood	-4242.061
Final log likelihood	-3179.451
Likelihood ratio test for the init. model	2125.220
Rho-square for the init. model	0.250
Rho-square-bar for the init. model	0.240
Akaike Information Criterion	6446.902
Bayesian Information Criterion	6742.552

Final gradiant norm	+2.231e-002
Diagnostic	Trust region algorithm with simple bounds (CGT2000): Convergence reached
Iterations	164
Data processing time	00:00
Run time	11:59
Nbr of threads	4

Parameter	Name	Value	Standard Error	t-test	p-value
ASC	B_ASC_1	4.42	0.703	6.29	0.00
	B_ASC_2	4.08	0.767	5.32	0.00
	B_ASC_3	3.75	0.713	5.26	0.00
	B_ASC_4	5.76	1.98	2.91	0.00
	B_ASC_5	-0.11	2.11	-0.05	0.96*
Level of automation	B_aut_1	-0.283	0.209	-1.35	0.18*
	B_aut_2	-0.162	0.165	-0.98	0.33*
	B_aut_3	0.077	0.199	0.39	0.70*
	B_aut_4	-4.160	0.819	-5.08	0.00
	B_aut_5	0.840	0.623	1.35	0.18*
Road exemption	B_rex_1	-0.083	0.155	-0.54	0.59*
	B_rex_2	0.500	0.186	2.68	0.01
	B_rex_3	-0.174	0.139	-1.26	0.21*
	B_rex_4	-0.733	0.315	-2.33	0.02
	B_rex_5	0.578	0.568	1.02	0.31*
Travel time	B_tti_1	-0.590	0.137	-4.32	0.00
	B_tti_2	-1.650	0.252	-6.55	0.00
	B_tti_3	-0.099	0.114	-0.87	0.38*
	B_tti_4	-0.168	0.280	-0.60	0.55*
	B_tti_5	-0.760	0.439	-1.73	0.08*
Human error fatalities	B_hum_1	-0.0053	0.0012	-4.55	0.00
	B_hum_2	-0.0012	0.0019	-0.62	0.53*
	B_hum_3	-0.0024	0.0014	-1.70	0.09*
	B_hum_4	-00041	0.0028	-1.46	0.15*
	B_hum_5	-0.0049	0.0027	-1.81	0.07*
Technical failure fatalities	B_tec_1	-0.028	0.0061	-4.57	0.00
	B_tec_2	-0.005	0.0028	-1.82	0.07*
	B_tec_3	-0.007	0.0036	-2.06	0.04
	B_tec_4	-0.015	0.0053	-2.45	0.01
	B_tec_5	-0.023	0.0174	-1.32	0.19*
Deliberate misuse fatalities	B_mis_1	-0.036	0.0089	-4.02	0.00
	B_mis_2	-0.001	0.0067	0.09	0.93*
	B_mis_3	-0.009	0.0054	-1.69	0.09*
	B_mis_4	-0.015	0.0132	-1.15	0.25*
	B_mis_5	-0.043	0.0321	-1.35	0.18*
Class membership parameter	s2	-0.343	0.365	-0.94	0.35*
	s3	0.251	0.335	0.75	0.45*
	s4	-0.900	0.379	-2.38	0.02
	s5	-0.733	0.249	-2.95	0.00

Below are the results per class (significant parameters in red). The class membership probability is calculated as follows:

$$\pi_{ns} = \frac{e^{\delta n}}{e^{\delta 1} + e^{\delta 2} + e^{\delta 3} + e^{\delta 4} + + e^{\delta 5}}$$

Table 18: 5 Class LC model parameters

Parameter	Class 1	Class 2	Class 3	Class 4	Class 5
ASC	4.42	4.08	3.75	5.76	-0.110
Level of automation	-0.283	-0.162	0.077	-4.16	0.844
Road exemption	-0.083	0.500	-0.174	-0.73	0.578
Travel time	-0.590	-1.65	-0.099	-0.168	-0.760
Emissions	-0.309	-1.70	-0.193	-0.655	-0.686
Human error fatalities	-0.0053	-0.0012	-0.0024	-0.0041	-0.0049
Technical failure fatalities	-0.0278	-0.0051	-0.0073	-0.0154	-0.0230
Deliberate misuse fatalities	-0.0359	0.0006	-0.0091	-0.0153	-0.0434
Class membership parameter (δ)	0 (fixed)	-0.343	-0.251	-0.900	-0.733
Class membership probability (π_{ns})	26%	26%	26%	10%	10%

To convert these results to useful information, classes must be identified that make sense in reality. Unfortunately, class 3 and class 5 hardly contain any information. For class 3, only the ASC and technical failure fatalities parameter become significant. For class 5, none of the parameters become significant. Although the other classes have more significant parameters, still no clear categorization of classes can be made. Moreover, no quadratic parameters became significant for any latent class choice model.

VIII. 3 Class Latent Class Choice Model Results

Number of estimated parameters	30
Sample size	6120
Excluded observations	0
Init log likelihood	-4242.061
Final log likelihood	-3315.817
Likelihood ratio test for the init. model	1852.488
Rho-square for the init. model	0.218
Rho-square-bar for the init. model	0.211
Akaike Information Criterion	6691.633
Bayesian Information Criterion	6893.213
Final gradiant norm	+2.858e-003
Diagnostic	Trust region algorithm with simple bounds (CGT2000): Convergence reached
Iterations	32
Data processing time	00:00
Run time	01:47
Nbr of threads	4

Parameter	Name	Value	Standard Error	t-test	p-value
ASC	B_ASC_1	4.67	1.09	4.28	0.00
	B_ASC_2	2.46	0.266	9.26	0.00
	B_ASC_3	-0.456	1.05	-0.44	0.66*
Level of automation	B_aut_1	0.485	0.239	2.03	0.04
	B_aut_2	-0.836	0.120	-6.95	0.00
	B_aut_3	0.638	0.503	1.27	0.20*
Road exemption	B_rex_1	-0.145	0.163	-0.89	0.37*
	B_rex_2	-0.088	0.060	-1.46	0.15*
	B_rex_3	0.437	0.309	1.41	0.16*
Travel time	B_tti_1	-0.449	0.187	-2.40	0.02
	B_tti_2	-0.550	0.059	-9.31	0.00
	B_tti_3	-0.603	0.263	-2.29	0.02
Human error fatalities	B_hum_1	-0.0045	0.0020	-2.24	0.03
	B_hum_2	-0.0028	0.0005	-5.28	0.00
	B_hum_3	-0.0048	0.0017	-2.80	0.01
Technical failure fatalities	B_tec_1	-0.0116	0.0044	-2.64	0.00
	B_tec_2	-0.0139	0.0016	-8.94	0.00
	B_tec_3	-0.0271	0.0084	-3.25	0.00
Deliberate misuse fatalities	B_mis_1	-0.0145	0.0066	-2.18	0.03
	B_mis_2	-0.0183	0.0026	-7.13	0.00
	B_mis_2	-0.0430	0.0196	-2.20	0.00
Class membership parameter	s2	0.494	0.198	2.50	0.01
	s3	-0.663	0.231	-2.87	0.00
Gender	gen2	0.240	0.121	1.98	0.05
	gen3	0.420	0.158	2.66	0.01
ADAS familiarity	ada2	0.290	0.148	1.96	0.05
	ada3	-0.026	0.182	-0.14	0.89*

IX. Scientific Article

Automated driving systems (ADS) can improve traffic safety, improve accessibility and reduce environmental impact (Shladover, 2016). On the contrary, on May 7th 2016, a fatal accident with a Tesla on autopilot in U.S. Florida was a harsh reminder that the technology is still in its testing phase (Greenemeier, 2016). In complex technical systems like ADS technical failure may occur, which forms a serious threat to human well-being. Moreover, studies have shown that citizens are very concerned about deliberate misuse of ADS (Kyriakidis, Happee, & De Winter, 2015), e.g. people purposely abusing ADS to cause damage or even hurt someone. Therefore, according to many experts, the implementation of ADS does not only entail technical issues but also normative issues.

To bridge the gap between technical- and normative issues, responsible innovation can be applied (Santoni de Sio, 2016). Friedman et. al. (2006) propose value-sensitive design to achieve responsible innovation via technical-, empirical- and conceptual research. As following from responsible innovation and value-sensitive design and in particular its focus on empirical research as one of its necessary elements, in this research an effort is made to use empirical research methods to provide more insights in normative issues of automated driving systems (ADS). The focus lies on social acceptance, particularly with respect to traffic safety. Also accessibility, environmental impact and heterogeneity among citizens will be analysed. Social acceptance is defined as "a person's assent to the reality of a situation, recognizing a process or condition (often a negative or uncomfortable situation) without attempting to change it, protest, or exit" (Fish, 2014, page 1). The following questions will be answered in this research:

What is the social acceptance of automated driving systems from the perspective of safety, accessibility and environmental impact and what is the corresponding heterogeneity?

1. *What percentage of citizens thinks automated driving systems are socially accepted?*
2. *How is the social acceptance influenced by safety, accessibility and environmental impact?*
3. *Are traffic fatalities caused by automated vehicles valued differently than current traffic fatalities?*
4. *Is there heterogeneity in the social acceptance among citizens?*

A survey is chosen as research method since it is a relatively inexpensive, flexible method to achieve extensive information about characteristics of a population. After an extensive theoretical analysis, seven attributes were identified that possibly influence the social acceptance: level of automation, road exemption, travel time, emissions, human error fatalities, technical failure fatalities and deliberate misuse fatalities. After the experiment was fine-tuned by a pilot study, it was held among a representative sample of 510 Dutch adults during the spring of 2017. The respondents had to state if they were in favour or against ADS for each of the twelve hypothetical futures that were presented to them. In these hypothetical futures, the attributes were systematically varied and described as a change to the current situation.

Using a MNL RUM model, the results show that 63% of all citizens prefer ADS over the current system. It is therefore concluded that citizens have a high social acceptance and thus are rather positive towards ADS. Also, citizens prefer a system where human drivers are still in control and can intervene in case necessary.

Next, it is concluded that the social acceptance is mostly influenced by fatalities caused by automated vehicles (AVs), while travel time is the least important attribute. However, the differences in influence of the attributes were not substantial. Safety, accessibility and environmental impact are all important for the social acceptance. It is also concluded that a technical failure fatalities weighs as much as 4 human error fatalities. Between deliberate misuse fatalities and human error fatalities is a factor 5.5. Consistently throughout the modelling of the results, this effect became statistically significant. The results coincide with literature (Kyriakidis et al., 2015).

The magnitude of the difference between these factors is larger than expected, which might causes problems for the implementation of ADS. The results imply that AVs have to cause hardly any fatalities to reach social acceptance even if they cause much less human error fatalities. Given these results, it is questionable if the Netherlands should be a test-bed for ADS. Although safety measures are in place, these experiments still propose risks to citizens, which are currently not accepted. It also makes the ethical acceptability of such experiments questionable at least. Only a couple of accidents might cause societal pressure to stop experiments and delay the implementation of ADS for many years.

The paradox is that ADS can be safer in the future, but that a potentially risky transition phase is needed to reach this future. Citizens prefer a future that includes ADS, but do not want an unsafe learning period. More on this discrepancy is in the recommendation.

Finally, it is concluded that there is a large heterogeneity in the social acceptance of ADS. This research has shown that citizens can be segmented into three classes: automated driving enthusiasts, the central mass and risk-averse class.

32% of citizens can be assigned to automated driving enthusiasts and these citizens have a preference for automation level 4. Interestingly, this is contradictory to the average citizen. This class has relatively low penalties for fatalities. Even so, technical failure fatalities (deliberate misuse fatalities) still weigh as much as 2.7 (3.4) human error fatalities. Automated driving enthusiasts are more likely to be male and to be less familiar with ADAS.

The central mass shows similar results as the average citizen. This class contains 52% of all citizens who distinguish themselves by a large dislike for automation level 4. They weigh technical failure fatalities (deliberate misuse fatalities) as much as 5 (6.5) human error fatalities. Also, they are more likely to be familiar with ADAS. Possibly they know the limitations of ADAS and therefore prefer a slow increase in automation levels.

Finally, the risk-averse class has a low social acceptance and every fatality is strongly penalized. Members of this class weigh technical failure fatalities (deliberate misuse fatalities) as much as 5.5 (9.8) human error fatalities. They are more likely to be female and it is also the smallest class with only 16% assigned to it.

The progress that has to be made in safety is much larger for the risk-averse class than for automated driving enthusiasts. Assuming that some fatalities caused by AVs are inevitable, automated driving enthusiasts would accept the overall safety record of ADS approximately three times more likely than the risk-averse class. It might imply a difference of years or even decades when ADS are socially accepted by the two classes. Also, some automated driving enthusiasts are likely to use AVs, thus proposing risks to society. Although these users might accept these risks, these same risks are not accepted by the risk-averse class or even the central mass. Therefore, the heterogeneity is critical for the implementation of ADS and is discussed in the recommendations.

In conclusion, primarily two discrepancies are identified that are critical for the implementation of ADS: 1) High social acceptance versus strong dislike for fatalities caused by AVs; 2) Citizens who are enthusiastic about ADS versus citizens who are risk-averse. They lead to the following recommendations:

It is recommended to be extra cautious during the transition phase especially with experiments for private vehicles. Current experiments are largely based on professional users, like freight transport and public transport. Risks can be easier contained for these users, for example by using segregated lanes (WE pods) or supervision by trained professionals (truck platooning). For private vehicles it might be possible to experiment in traffic jams. Low speeds in traffic jams ensure minimal impact in case of an accident. Either way, the risk of fatalities caused by AVs should be minimized.

Also, policy makers should decide if AVs may use cooperate driving, i.e. be connected to other vehicles, infrastructure or a cloud. Shladover (2016) argues that automated driving is not possible without cooperate driving. However, these connections make AVs vulnerable for deliberate misuse which is strongly penalized by citizens. The experiment by Bridle (2017) showed that with only a package of salt he could alter the road markings read by AVs. Imagine what a hacker can do by altering 3D-maps or other environmental information of AVs. The report of Roland Berger (2017) indicates that in the Netherlands, as well as globally, research in cooperate driving is lacking. If ADS cannot exist without cooperate driving, more knowledge on this subject is needed. It is therefore recommended to conduct more research into cooperate driving. Possibly, this should lead to restrictions for technology producers like TomTom to safely deploy cooperate driving technologies.

This research has found large heterogeneity among citizens, especially in the acceptance of fatalities caused by AVs. The Vienna Convention (1986) states that a human driver has to be in control of its vehicle and its environment. However, it is likely that some automated driving enthusiasts will use the automated driving features since they prefer high automation levels. Consequently, they will not be in complete control of the vehicle and dangerous situations may arise. When lower priced vehicle will obtain automated driving capabilities, the technology will be available to a larger public and more enthusiasts will be able to experiment with the technology. In turn, this might inflict an unaccepted situation, especially for citizens who are risk-averse.

Therefore, a passive approach to AVs cannot be made. Currently, there are no limitations to automated driving technologies in vehicles, so it is recommended to review the licencing of such vehicles. The Ministry of Infrastructure and the Environment and the RDW have to actively engage with technology producers whose technologies are available to consumers. Even if these vehicles are licensed, restrictions should be made to the technology. For example, make sure automated driving technologies are switched off above the speed of 50 km/h; or make sure the technology cannot be switched on uninterruptedly for long periods of time. For private vehicles, it is recommended to stick to direct human control for the transition phase. As the technology progresses, direct human control may be widened to meaningful human control, thus including some kind of delegation.

Also, the social acceptance is partly based on incomplete or even false information. So even if two individuals have the same norms and values, they might have a different social acceptance of ADS based on different information they received. Therefore policy makers can use campaigns to provide information to citizens about ADS. If policy makers and engineers believe that social impacts are positive and philosophers believe ADS is ethically acceptable, campaigns can be used to make sure citizens are informed about the social impacts and way of implementation. For example, automated driving enthusiast can be made aware of the risks of automated driving thus making sure they use the technology more responsibly; or the risk-averse class and central mass can be made aware of overall safety benefits, changing their acceptance of fatalities caused by AVs. Either way, these campaigns might ensure citizens that their norms and values are embedded in the design of ADS.

Automated driving systems are set to be one of the more disruptive technologies for the coming future. The way we see and experience transport is very likely to change in the coming decades. If we can align research, development and design to the social acceptance and ethical acceptability of the innovation process and its marketable products, and if we can align the corresponding actors in this field, we can make sure that automated driving systems become a responsible innovation.