# Voice Assistants in Cars - Dream or Nightmare?

The effects of voice assistants on
trust, emotions and purchase intention

Author: Anna-Maja Wolf
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

**ABSTRACT**

Over the past years, voice assistants have become a constant companion in people's daily lives. Not only are they used during day-to-day activities as calling someone, ordering food, or asking for the weather, but they are additionally supporting people during their drive to work, friends, or holidays. Current literature found that negative emotions lead to an increase in a risk-supporting attitude and more accidents. This study examines the effects of voice assistants on trust and purchase intention and emotions as stress, anger, confusion, and concentration. A scenario-based 2 (gender: male vs. female) x 2 (humanity: human vs. computer) x 2 (time of recording: day vs. night) experiment was created with a total of 256 German participants between 18 and 50+ years. The multivariate analysis of variances showed that humanity influences the felt emotions and the purchase intention. Respondents who heard a human voice experienced a lower level of stress and anger and a higher possibility to concentrate. Furthermore, they were more likely to purchase a car having a human voice assistant. Additionally, the study revealed a significant interaction effect between gender, humanity, the time of recording, and stress, meaning that a higher stress level was found when the voice was an artificial voice taking into account the gender and recording time. These findings suggest that using a voice assistant in cars might be helpful and supporting and would decrease the negative emotions when a human voice is used instead of an artificial one.

**Graduation Committee members:**
First supervisor – dr. C. Herrando
Second supervisor – dr. E. Constantinides

**Keywords**
Artificial Intelligence, Voice Assistants, Emotions, Driving Behavior, Trust, Purchase Intention, Stress, Anger, Confusion, Concentration

# 1. INTRODUCTION

## 1.1 Academic Relevance

The emotional state drivers experience while driving has influences on their driving performance. Negative emotions as anger, frustration, or sadness lead to a decrease in their driving performance and endanger road safety for all other drivers (Braun, Schubert, Pfleging, & Alt, 2019). Therefore, it is crucial to increase the driver's emotions to ensure a safe drive for them and their travelers (Eyben et al., 2010). A helpful tool often used in cars to relieve the drive for the driver are in-voice care assistants, hereafter VA. The possibility of a vehicle interacting with the driver and serving as a virtual partner is seen as a great opportunity for current and future cars (Eyben et al., 2010). It will help increase driving safety (Eyben et al., 2010) as it serves as virtual assistance which adds to the security of driving (Bubb, 2003). Generally, the driving performance is influenced by the current emotional state of the driver as some abilities, like focus and attention or communication, are dependent on the present emotions (Eyben et al., 2010). When considering the driver's responsibility over his passengers and other road members, it is crucial to keep the driver in an emotional state that suits the driving task best (Eyben et al., 2010). Resulting also in a proper fit of the voice assistant, as it supports the driver in the performance of primary driving tasks (Bubb, 2003) and, therefore, indirectly influences the emotional state (Eyben et al., 2010). Primary driving tasks include all the tasks directly related to the drive, like the choice of the lane, the speed, the route, or the space between the own car and the front car (Bubb, 2003). The voice assistant's interaction is operated as a secondary driving task (Braun et al., 2019). A study conducted by Agrawal, Giripunje & Bajaj (2013) tried to recognize drivers' emotions during their drive using a technology system called Fuzzy Rule-Based System. The system had an accuracy of 90% (Agrawal, Giripunje, & Bajaj, 2013). This shows the possibilities of today's technologies to interact with the driver based on the previously identified emotional state. Furthermore, it can be said that if the voice of the car assistant matches the drivers' current emotional state, it increases the bond between the driver and the assistant and replaces the role of a co-driver (Eyben et al., 2010).

The assistant's voice plays an essential role as well as the characteristics it has, whether it is a male or a female voice or if it is human likely or not (Jonsson & Chen, 2007). The increasing usage of in-car assistants makes it curial to understand the driver's emotions and the influence the voice assistant has on them. It could be that the VA's characteristic influences the emotional state of the driver. As safety is the most important aspect in designing cars and implementing VAs, it is essential to know which impacts different characteristics of the VA might have on the driver.

## 1.2 Research Question

The research question of this thesis therefore is:

***Which effects have the use of voice assistants on the drivers' emotional state?***

# 2. THEORETICAL BACKGROUND

## 2.1 Artificial Intelligence in Voice Assistants

Artificial Intelligence (AI) is the expansion of computers to interact in humanoid processes like learning, analyzing, and interacting (Kok, Boers, Kosters, Van der Putten, & Poel, 2009). According to Poushneh (2021), AI is the intelligence to imitate and interact with a human attitude which is often digitally shown. Generally, there are three different bits of AI intelligence: mechanical, thinking, and feeling (Huang & Rust, 2021). Feeling AI is constructed for the bi-directional interaction implicating humans and/or additionally for interpreting hominid perceptions and emotions (Huang & Rust, 2021). The purpose is to interpret current and future consumer desires (Huang & Rust, 2021).

Voice Assistants are a sample of speech-empowered artificial intelligence (Poushneh, 2021). They allow customers to use various functions without the need to interact directly with other people. These functions can include navigation, listening to music, sending messages, managing smart home systems, calling somebody, ordering a meal, or a ride home (Poushneh, 2021). Voice Assistants have the ability to raise positive affirmations, decrease discouragement and contribute positively to social communication (Kachouie, Sedighadeli, Khosla, & Chu, 2014).

Current studies highlight that the use of in-voice assistants influences driving behavior. The interaction with a virtual assistant affects the personal attitude, performance, or behavior (Harris & Nass, 2011). Herby, differences in the different genders have been found. Female drivers perceived a VA helpful in low and high traffic locations and mentioned that the in-voice assistant had an increasingly positive effect on their driving performance (Jonsson & Chen, 2007). However, male drivers preferred the assistant more in low traffic locations than in high traffic situations as they were afraid of being distracted by it (Jonsson & Chen, 2007).

## 2.2 Trusting the Voice Assistant

Trust can be described as one's compliance to be dependent on something else due to the abilities of the opposite (Mcknight, Carter, Thatcher, & Clay, 2011) and plays a fundamental role in supporting users to overcome their consciousness of hazard and uncertainty when using and approving the latest technology (Li, Hess, & Valacich, 2008). Another definition of Mayer, Davis & Schoorman (1995) describes trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer, Davis, & Schoorman, 1995, p.712). Trust is a central element in every situation where skeptics exist or where unpredictable events are imaginable (Fukuyama, 1997). When deciding whether to trust new technology, users decide on the available knowledge and current recognition of the technology (Moore & Benbasat, 1991). Factors influencing trustworthiness are, for example, gender or speech quality (Schwieren, 2020). On the other hand, users are unlikely to trust the new technology and interact with it when they recognize a serious risk related with it (Agarwal & Prasad, 1997). Generally, trusting in a precise technology means that one believes the technology has the required abilities to behave and perform in the safest manner during a situation which could lead to negative outcomes (Mayer et al., 1995). When analyzing the trustworthiness of a new technology, a study from Lankton, McKnight, & Tripp (2015) distinguished between "system-like and human-like technology" (Lanktop, McKnight, & Tripp, 2015, p.881). Human-like technology is defined as seeing the technology as a human-being while measuring trust (Lankton et al., 2015). The trust is measured utilizing the human-like trust design consisting of benevolence, ability, and integrity which originally have been used to estimate interpersonal trust (Lankton et al., 2015). Benevolence can be defined as the acceptance that the agent will act in the best behavior for the trustor, without egocentric motives (Mayer et al., 1995). Ability

is the trust that the agent has the necessary skills and competencies to perform the tasks (Mayer et al., 1995). Lastly, integrity is the belief of the agent operating consistently and in line with given principles (McKnight et al., 2011). Additionally, when deciding to purchase a voice assistant, individuals are influenced not only by their trustworthiness but also by the expressed emotions and behavior of them (Poushneh, 2021). Based on the presented literature, the first two hypotheses have been developed.

*H1: Trust in a VA is higher when the VA has a human-likely voice than when the VA has a robot-likely voice.*

*H2: Trust in a VA is higher when the VA has a female voice than when the VA has a male voice.*

## 2.3 Emotional states

Kleinginna & Kleinginna (1981) have defined emotions as a multi-level set of interactions within various subjective and objective factors managed by human hormones. They can influence behavior, manage the cognitive procedure and allow experiences (Kleinginna & Kleinginna, 1981). Cabanac (2002) encloses that emotions are perceived in different categories as anger, happiness, sadness, or surprise. Psychologists have found that it is impassable to think or behave without involving the emotional system (Nass et al., 2005). Emotions affect various competencies and traits like perception and organization of memory (Bower, 1981), categorization and preference (Zajonc, 1984), strategic planning (LeDoux, 1992), focus and attention (Derryberry & Tucker, 1992) and, motivation and performance (Colquitt, LePine, & Noe, 2000). Furthermore, they are influenced by different acoustic characteristics like volume, speed, or frequency (Nass et al., 2005).

A study by Harris & Nass (2011) analyzed the influence VAs have on the driver's emotional state. The outcome showed that participants who were part of the reappraisal-down condition had more positive emotions than the other groups' participants.

A study from Braun et al. (2019) achieved the knowledge that a VA with an emotional state and the ability to interact with the driver is used best when improving emotions and increasing the driving performance. Furthermore, speech is an essential transporter of information emotionally (Nass et al., 2005).

To increase the approval and the use of VA, it is advised to include human capacities in the machines (Borau, Otterbring, Laporte, & Fosso Wamba, 2021). These capacities include verbal and non-verbal contact (Borau et al., 2021). This human-like interaction with the VA helps to increase trust and the relationship between the machine and the user (Borau et al., 2021). Resulting from the literature presented above, the third hypothesis is:

*H3: The driver's emotional state is better if the VA has a human-likely voice than when the VA has a robot-likely voice.*

Additionally, appointing a gender to the VA helps to increase the anticipated humanness (Borau et al., 2021). Hereby, differences between the genders have been found. Women, in general, are perceived more as having favorable human intentions (e.g., friendly, warm, and credulous) and the capacities to recognize and feel emotions, like anger, empathy, frustration (Borau et al., 2021). Multiple studies have shown that customers prioritize a female over a male voice because they are superior in transporting and experiencing emotions (Eyssel, De Ruiter, Kuchenbrandt, Bobinger, & Hegel, 2012; Otterbache & Talias, 2017; Stroessner & Benitez, 2019). Therefore, the hypothesis H4a and H4b concentrate on the influence gender has on the emotional state.

*H4a: The driver's emotional state is positively influenced if the VAs gender is female.*

*H4b: The driver's emotional state is negatively influenced if the VAs gender is male.*

## 2.4 Driving behavior under the influence of emotions

Driving appears in a miscellaneous environment involving bikers, pedestrians, other cars, traffic, and traffic signs. This reduces the opportunity for the driver to decrease frustrating situations (Harris & Nass, 2011). Ideally, drivers' frame of mind should support the competencies required for driving safely. These emotions are related to paying attention, good decision-making, and intuition (Agrawal, Giripunje, & Bajaj, 2013).

In reality, due to the increasing amount of vehicles during traffic hours, drivers experience an increase in frustration and negative emotions (Harris & Nass, 2011). Frustration functions as an entrance emotion, leading to a more aggressive driving attitude and anger (Harris & Nass, 2011). Harris & Nass (2011) define this as road rage behavior, which is an increasing problem in today's traffic and the reason for multiple accidents. Reasons for this behavior are, on the one hand, many adjacent hours of driving, and on the other hand, the outcome of stress aroused by heavy traffic (Alvarez, Lopez-de Ipiña, Daily, & Gilbert, 2012).

The personal characteristics traits were also considered to play an important role regarding the driving behavior and the emotional state (Deffenbacher, Lynch, Oetting & Yingling, 2001; Underwood, Chapman, Wright & Crundall, 1999; Zimasa, Jamson & Henson, 2017). Hereby, Zimasa et al. (2017) put up the hypothesis of whether happy drivers are safer drivers. Their study outcome accepted this hypothesis by finding out that drivers in a happy or neutral mood have a better and faster reaction to upcoming hazards than drivers in a sad mood. A study from Underwood et al. (1999) concluded that there was a strong positive relation between almost having an accident and the level of anger the driver feels during the drive. It was found that drivers who generally have a lower boundary for anger are more likely to show angry behavior in road traffic and adapt to a more unsafe and combative driving style (Deffenbacher et al., 2001).

Generally, it can be distinguished between different types of emotions that negatively influence the driving performance of the driver: anger, stress, confusion, and concentration.

### *2.4.1 Aggressiveness and Anger*

Research has shown that anger negatively influences driving behavior and leads to a more risk-supporting attitude such as speed, lane deviations, and collisions (Jeon, Walker, & Yim, 2014).

Aggressiveness and anger can be summarized in the term road rage. Based on the definition from Wells-Parker et al. (2002), road rage is seen as the impulsive performance in precarious activities that leads to an increase in the possibility of an accident. Additionally, less extreme appearances can be found in expressing anger verbally in inconspicuous ways as shouting through a locked window, talking to oneself, or expressing the emotions using functions as the headlights' flash (Wells-Parker et al., 2002). These milder forms appear most often in the daily traffic routines (Hennessy & Wiesenthal, 1997).

### *2.4.2 Stress*

Stress is defined as a feeling from a patient state to an annoying state activated by one or more factors in the environment (Healey & Picard, 2005). Most stress factors are either intellectual, personal, or perceptual (Healey & Picard, 2005). Stress usually develops through factors as full roads, rush hours, driving scrupulously behind other cars, taking an incoming call, or trying to be punctual at the final destination.

### *2.4.3 Confusion*

Confusion appears in complex traffic settings, as misleading traffic signs, a complication in the route, or through the performance of other drivers. It can also result from driving the route for the first time or being in a bigger city. Confusion is seen as a significant negative influence on the safety risk, as the driver is not aware of how to behave in the situation and whether one is responsible for the primary driving tasks (Wilson, Yang, Roady, Kuo, & Lenné, 2020).

### *2.4.4 Concentration*

Keeping concentration high while driving is a significant operation to ensure safe road traffic as distraction of the driver is one of the most crucial contributors to accidents (Fofanova & Vollrath, 2011). Almost 23% of all accidents happening on the road are caused through distraction by secondary tasks (Fofanova & Vollrath, 2011). Harbluk, Noy, Trbovich, et Eizenman (2007) conducted a study where drivers need to perform secondary no-hands tasks besides driving. They found out that drivers are mostly distracted by their environment and spend time looking outside of their vehicle. Furthermore, this distraction leads to a reduction in the monitoring of the mirrors or traffic lights (Harbluk et al., 2007).

Resulting from the literature hypothesis H5a, H5b, H5c and H5d are the following:

*H5a: Anger is lower when a VA with a human-likely voice is used during the drive than when a VA with a robot-likely voice is used.*

*H5b: Stress is lower when a VA with a human-likely voice is used during the drive than when a VA with a robot-likely voice is used.*

*H5c: Confusion is lower when a VA with a human-likely voice is used during the drive than when a VA with a robot-likely voice is used.*

*H5d: Concentration is higher when a VA with a human-likely voice is used during the drive than when a VA with a robot-likely voice is used.*

When focusing on VAs and their influence on driving performance, it can be said that based on a study from Eyben et al. (2010), participants have agreed on the positive influence of VAs in unknown and unpleased situations. Additionally, it was found out that the voice of the VA plays a significant role in the influence it has on the driver (Nass et al., 2005). In an experiment with a 2x2 study matrix (emotion of driver and emotion of in-voice assistant), the conclusion was drawn that the number of accidents can be reduced if the car assistant's voice fits the emotional state of the driver (Nass et al., 2005).

## 2.5 Purchase Intention

Purchase Intention can be defined as the purpose of one to decide which brand to choose to purchase from (Rezvani et al., 2012). Furthermore, it describes the feeling of purchasing a product and the loyalty to the brand, which results out of personal tendencies towards a product (Rezvani et al., 2012). Another definition describes purchase intention as the individual's consciousness to make a try to buy a product (Shabbir, Kirmani, Iqbal, & Khan, 2009). A previous study from Daneshvary and Schower (2000) found out that the intention to purchase is influenced by demographic factors as age, gender, or the country of origin (Wang, Li, Barnes, & Ahn, 2012). Other external factors influencing an individual's decision to purchase are the quality perception, the price, and the given value (Chang & Wildt, 1994), but also the customer's knowledge (Pires, Stanton, & Eckford, 2004). The higher the value of a product, the higher the likelihood a consumer is going to purchase it (Chang & Wildt, 1994). Regarding voice assistants, in today's society, most of the conversational agents are female (i.e. Alexa, Siri, Cortana). They do not only have female names but also speak with a female voice (Feine, Gnewuch, Morana, & Maedche, 2019). Over 60% of German inhabitants have used voice assistants (Tenzer, 2019), and a study from Adams (2019) revealed that almost 56% of her respondents are in the current usage of a voice assistant. Based on the literature, the sixth hypothesis is as followed:

*H6: Purchase Intention is higher when the voice assistant has a female voice than when the VA has a male voice.*

## 2.6 Conceptual Framework

Based on the literature and the formulated hypothesis, the following conceptual framework (Figure 1) will serve as a foundation estimating the positive influence the correct characteristics of voice assistants can have on the overall driving performance.

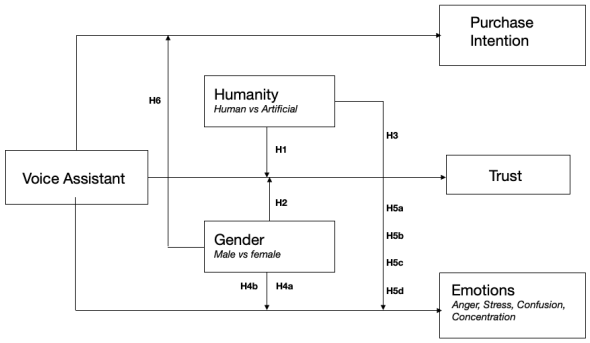



**Figure 1. Conceptual Framework**

# 3. METHODOLOGY

## 3.1 Research Design

A proper design method needs to be used to answer the research question and justify the hypotheses. Regarding the scope of the study, an experiment will be conducted with different experimental groups. The overall research question is answered using primary and secondary data. For exploiting primary research data is used that is "collected directly from researches for the purposes of their research objectives," (Chrysochou, 2017, p.411). which enables them to have adequate control over the collection of the data. The secondary research includes "sources and data that are already available" (Chrysochou, 2017, p.411). Keywords used for the secondary data collection were, e.g., Artificial Intelligence, Emotions, Driving behavior, Voice assistants, characteristics or differences, primarily focusing on the relationship between two or more of the keywords. Based on this research, a conceptual framework has

been created, which serves as the basis for the following study. The primary data will be collected using a quantitative experimental study consisting of a video and one survey. Surveys are a common method to explore data, as they add the opportunity for an accurate operationalization and analysis of the data (Chrysochou, 2017).

The experimental study will be conducted between subjects where the participants will see one out of eight possible videos, and afterward, a survey needs to be filled out. In total, there will be eight different stimuli in the end, as seen in figure 2.

The study focuses on whether different characteristics of VAs have different impacts on the driver's emotional state. Therefore, different VAs will be used: a female with a human voice, a female with a computer voice, a male with a human voice, and a male with a computer voice. To encourage different stress situations, each VA will be used once during the day and once during the night. To keep the limitations as small as possible and not get in trouble with legal issues, the researcher has taken the videos herself using a GoPro. The car used for all videos is a Renault Zoe (year of construction 2019), and the location of filming is Ahaus, Germany. The route driven is for all videos the same and can be found in the appendix (Appendix Part 1), but only parts of the route have been used for the video.

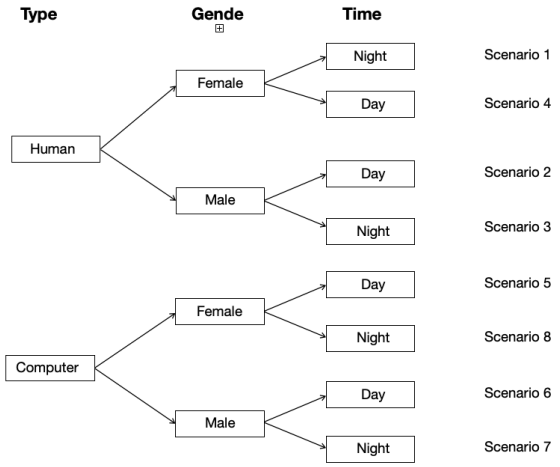



**Figure 2. Eight stimuli**

## 3.2 Designing questions

The questions used in the survey are based on existing literature (see Appendix Part 2). The whole questionnaire that was sent out can be found in the appendix (Appendix Part 3). Furthermore, a link to all eight videos, published on YouTube, can be seen in the appendix (Appendix Part 4).

At the beginning of the questions, the respondents were asked to answer five questions regarding their individual being. These questions contained information on the participant's age, the gender they identify with, the current living location, how long the participant has their driver's license, and the amount of car rides they do per week. These questions have been answered by the participants using the multiple-choice option with pre-selected answers. It continued with a question about ones' driving behavior, where the option was to select multiple answers as the individual choices were not mutually exclusive. When designing a survey, it is crucial to review the individual multiple choice options to ensure that the answers are mutually exclusive and exhaustive (Neuman, 2014). Lastly, the participants indicated how much they agree or disagree with the statement 'I am an experienced driver.'.

After answering these questions, the participants were randomly assigned to one out of the eight different videos. In total, 256 respondents have been used for further analyses. Hereby it was ensured that each stimulus had more than 30 respondents. In the end, scenario two and three got 34 participants, scenario eight 33, scenario four 32, scenario five, six, and seven each got 31 participants, and scenario one got 30 participants. A comprehended overview of the distribution among the scenarios can be found in table 1. The videos have an approximate length of three minutes and showed the researcher driving around. During the video, the participants got different commands and support from a voice assistant. The human VAs have been recorded by persons from the researchers' personal circle. It was ensured that the voices express stereotypical attributes of their gender and have a clear pronunciation. The website notevibes.com served as the base for the computer voices. Hereby, the persona 'Markus' has been chosen as the most convenient male voice, whether the persona 'Annika' has been used as the female assistant.

In the second half of the survey, the respondents were asked to answer statements on a five-point Likert scale from (1) strongly disagree to (5) strongly agree (McKnight, Choudhury, & Kacmar, 2002). The Likert scale can be described as "one of the most fundamental and frequently used psychometric tools" (Joshi, Kale, Chandel, & Pal, 2015, p.369). It was ensured that all questions of each item are unidimensional, meaning that "all the items in a scale or index fit together, or measure a single construct "(Neuman, 2014, p.125). The variables with the belonging items can be seen in the appendix (Appendix Part 2). The first nine statements measure the trustworthiness of the voice assistant regarding benevolence, ability, and integrity. Participants were asked to state whether, for example, the VA was concerned about their well-being, the VA is qualified in helping the driver, or if the pieces of advice have been helpful. The following thirteen statements concerned the feelings arising by the participant when hearing the voice of the voice assistant. The emotions checked were stress, anger, confusion, and concentration and were chosen to achieve a better understanding of the driver's emotional state during the drive. Lastly, participants were asked if they can imagine buying a car with a voice assistant and which functions are important to them when deciding for a car. Hereby, again multiple options were given as a parking aid, a camera, a navigation system, or a voice assistant. Due to missing mutual exclusivity, the participants were able to choose multiple options or indicate that they have other or no wishes for their car.

To ensure the attention of the participants, in the end they have been asked to mention whether they heard a male or a female voice and what the time of recording was (day vs. night).

To simplify the results for the quantitative analysis, open questions have been averted. The questionnaire consists of a total of 17 questions, excluding the video, where each question either represented an exact question or a question block of comparable statements. Question blocks were entirely used in the interest of Likert-Scales. The duration to finish the survey is approximately seven minutes and could be done at any time in a period of three weeks, completed from every device with an internet connection (Computer, Smartphone, Tablet, etc.)

### *3.2.1 Pre-test*

To ensure the study and the questions are clearly formulated, a pre-test was conducted, where five people watched the pre-recorded videos and the surveys before. After the received feedback, grammar and spelling mistakes have been removed, and two of the questions have been reformulated to avoid misunderstanding. The responses were deleted afterward, and

the five participants did not fill out the survey when it was online.

## 3.3 Data collection

Due to the ongoing worldwide pandemic, the collection of the data was completely online. An online survey allowed the researcher to reach a broader audience and get an extensive understanding as the majority of the target group are internet users. The survey was distributed via different social media channels, e.g., What's App, Instagram, and Facebook. To avoid confusion and misunderstanding by the participants, the instruction language was German, as the videos have been recorded in a German city. This allowed the researcher to gain a diversified sample regarding the age groups as it enables participation for older people who do not speak English fluently. The data which was conducted from the survey was treated completely anonymously to provide confidentiality. The BMS ethics committee of the University of Twente approved the survey first to ensure no personal affection. All responses were collected voluntarily with the possibility to withdraw the participation at the end after being informed about the general aim of the study.

## 3.4 Sampling

In total, 303 participants responded to the survey. This is a representative sample size which raises the verification of the findings and allows sub-groups to be compared meaningfully (Hartley, 2014). In this survey, the sub-groups focused especially on comparing different cases regarding gender, age groups, and living locations. Nevertheless, some exclusion criteria have been made. To participate in the survey, a sufficient level of German was required as well as the availability of a driver's license. Furthermore, participants needed to be above 18 and are able to see and hear clearly.

However, due to surveys being incomplete and answers with no use, some of the responses needed to be removed. More specifically, all responses have been checked regarding the manipulation check at the end and the duration. Completed surveys with a duration below three minutes and a misunderstanding of the gender of the voice assistant have been removed. Moreover, all responses have been checked on their usefulness based on the potential impact on the research question. In the end, out of the 303 responses, only 256 have been used for further analysis.

The sample differs in the age groups from 18-50+, with the majority being between 18-21 years (36,33%). Additionally, the duration of being in charge of a driver's license supports this, as 38,28% of the participants possess their driver's license for three to five years. Most of the participants identified with the female gender (60,55%), but 0,78% wanted to keep their gender to themselves. As the survey was designed only for German-speaking participants, the living locations were based on the 16 different federal states. The majority of the participants (71,48%) currently live in North-Rhine Westphalia, followed by 15,63% living in Lower Saxony. Regarding the amount of driving, 30,86% of the participants estimate that they drive daily, 21,09% drive between six to four times per week, 20,03% between three to two times per week. Lastly, 17,58% drive less than once a week, and 10,16% drive around once a week. The majority of the participants said they drive with very high consciousness and attention (41,83%). In the end, the respondents had been asked to mention which functions are important to them when buying a car, for example, a navigation system, a camera, or a voice assistant. Hereby, they had the possibility of choosing more than one answer. 23.2% indicated that a navigation system is essential to them, closely followed by a parc distance control (21.7%). A more detailed overview of the sample characteristics can be found in the appendix (Appendix Part 5).

## 3.5 Measurement Model Validation

To ensure an equal spread of gender in the scenarios a Chi-Square Test was conducted, which did not show significance ($X^2$ (16, N=256) = 12,67, p-value = 0,621). Out of the 259 respondents 155 (60,55%) identified as female, 99 (38,7%) as male and 2 participants (0,8% ) did not want to give information about their gender. The distribution within the scenarios had a wider spread, but it was ensured that the scenarios were seen by both genders. An overview of the full gender distribution can be seen in Table 1.

**Table 1. Gender distribution among the scenarios**

| Scenario* | N | % | % Female** | % Male** | % No information** |
|---|---|---|---|---|---|
| **1** | 30 | 11.7 | 50.0 | 50.0 | 0.0 |
| **2** | 34 | 13.3 | 64.7 | 32.4 | 2.9 |
| **3** | 34 | 13.3 | 64.7 | 35.3 | 0.0 |
| **4** | 32 | 12.5 | 65.6 | 31.3 | 3.1 |
| **5** | 31 | 12.1 | 74.2 | 25.8 | 0.0 |
| **6** | 31 | 12.1 | 51.6 | 48.4 | 0.0 |
| **7** | 31 | 12.1 | 54.8 | 45.2 | 0.0 |
| **8** | 33 | 12.9 | 57.6 | 42.4 | 0.0 |
| **Total** | **256** | **100.0** | | | |

* Scenario 1=female, human, night
Scenario 2=male, human, day
Scenario 3=male, human, night
Scenario 4=female, human, day
Scenario 5=female, computer, day
Scenario 6=male, computer, day
Scenario 7=male, computer, night
Scenario 8=female, computer, night

** % within each scenario

Furthermore, the two manipulation questions were checked using once an independent sample t-test and once a cross table. The first independent t-test for gender (male vs. female) showed that those participants who had a female voice indicated that they heard a female voice and vice versa with the male voice ($t(254) = -30.51$, $p < .001$). The test showed that participants who heard a female voice were able to record that they heard a female voice ($M = 1.48$, $SD = 0.50$) and not a male voice ($M = 3.34$, $SD = 0.489$). The second independent t-test for the humanity of the voice assistants showed significance as well ($t(254) = -5.65$, $p < .001$). Even if it was harder for participants to decide whether the voice was a human voice ($M = 2.13$, $SD = 0.99$) or an artificial voice ($M = 2.85$, $SD = 0.99$), the majority indicated it right. The cross table was used to measure the right indication of the recorded time. It was seen that all of the participants reported correctly whether the video was recorded during the day or the night. The cross table can be found in the appendix (Appendix Part 6).

To ensure the validity and reliability of the scales and to see whether the items match the variable, a confirmatory factor analysis was conducted. Furthermore, the average variance extracted (AVE), the composite reliability (CR), and Cronbach's alpha were calculated. To make sure all items measure in the same direction, two items were recoded firstly. For the variable trust, results showed that two items were loaded in a different factor. The items 'Der Sprachassistent hat

mir geholfen' ('The VA has helped me') and 'Die Tipps des Sprachassistenten waren konsequent' ('The advice of the VA were consistent') were removed for further analysis. The AVE of trust showed a value below 0.5. To be consistent, the values need to be above 0.5. Consequently, the lowest item 'Der Sprachassistent ist um mein Wohlbefinden besorgt' ('The VA is concerned about my well-being') was removed from the data for further analysis. For the second variable, emotions, the factor analysis reported that the individual emotions are seen as belonging to different factors. Therefore, in the following analysis, the emotions will be seen separately from each other. All variables were loaded correctly for the factor purchase intention, and no items needed to be removed. The Kaiser-Meyer-Olkin value was significant with a value of 0.83 and is used to check sampling adequacy (Rasheed & Abadi, 2014).

The next step was to calculate Cronbach's alpha to regulate the internal consistency of the variables. All of the dependent variables showed a value of above 0.7 except for the variable confusion. Furthermore, the average variance extracted (AVE) has been measured, which measures the degree of differences and expresses consistency with a value higher than 0.5 (Alarcón, Sánchez, & De Olavide, 2015). All variables showed consistency with values higher than 0.5 except for confusion again. As a consequence confusion with the items 'Die Anweisungen des Sprachassistenten waren schwer zu verstehen' ('The tips of the VA were hard to understand'), 'Die Anweisungen des Sprachassistenten haben mich irritiert' (The tips of the VA have irritated me') and 'Ich war unsicher, ob ich den Anweisungen vertrauen konnte' ('I was unsure whether I can trust the advices') had been removed from the sample for further analysis.

Finally, the composite reliability (CR) index was calculated, which showed reliability with values above 0.7. As all of the items show values above 0.7, it can be concluded that the dependent variables fulfill the requirements of reliability. A detailed overview of the analyses is shown in table 2.

# 4. ANALYSIS

## 4.1 Main effects

To check the hypotheses and answer the research question, the collected data has been analyzed using a multivariate analysis of variances (MANOVA) and an analysis of variances, short ANOVA. The analysis exposed several significant effects.

## 4.2 Gender

The multivariate analysis of variance revealed no significant effects between the gender of the voice assistant and the dependent variables. Whether the voice assistant had a female or a male voice had no significant effect on their trust perception ($F(1.248) = 0.82$, $p = 0.37$). Participants who had listened to a female voice ($M = 2.88$, $SD = 0.04$) trusted her as much as participants who had a male voice ($M = 2.94$, $SD = 0.04$). The same outcome was found for all of the dependent emotion variables anger, stress, and concentration. First, the level of anger towards the voice assistant was found out to be independent of the gender of those ($F(1.248) = 0.13$, $p = 0.72$). Respondents assigned to the female voice experienced the same level of anger ($M = 2.31$, $SD = 0.08$) as those who were assigned to the male voice ($M = 2.35$, $SD = 0.08$). Second, the stress level of the participants was not significant in relation to the gender of the voice assistant ($F(1.248) = 1.15$, $p = 0.29$). Participants whose voice assistant had a female voice were stressed ($M = 2.62$, $SD = 0.08$) on the same level as participants whose voice assistant had a male voice ($M = 74$, $SD = 0.08$). Lastly, the gender of the voice assistant had no significant influence on the level of concentration ($F(1.248) = 0.03$, $p = 0.86$). Respondents were able to concentrate on the same level whether the voice was female ($M = 3.14$, $SD = 0.07$) or male ($M = 3.16$, $SD = 0.08$). The multivariate analysis of variance furthermore revealed that there are also no significant effects of gender on the variable of purchase intention ($F(1.248) = 0.21$, $p = 0.65$).

**Table 3. MANOVA effects – Gender**

| Dependent variable | Sum of sq. | df | Mean sq. | F | Sig. |
|---|---|---|---|---|---|
| Trust | 0.16 | 1 | 0.16 | 0.82 | 0.37 |
| Anger | 0.11 | 1 | 0.11 | 0.13 | 0.72 |
| Stress | 0.90 | 1 | 0.90 | 1.15 | 0.29 |
| Concentration | 0.02 | 1 | 0.20 | 0.30 | 0.86 |
| Purchase Intention | 0.04 | 1 | 0.04 | 0.21 | 0.65 |

**Table 4. Mean and standard deviations – Gender**

| | Female | | Male | |
|---|---|---|---|---|
| **Dependent variable** | **M** | **SD** | **M** | **SD** |
| Trust | 2.88 | 0.04 | 2.94 | 0.04 |
| Anger | 2.31 | 0.08 | 2.35 | 0.08 |
| Stress | 2.62 | 0.08 | 2.74 | 0.08 |
| Concentration | 3.14 | 0.07 | 3.16 | 0.08 |
| Purchase Intention | 2.68 | 0.04 | 2.65 | 0.04 |

## 4.3 Humanity

For the independent variable humanity all dependent variables showed significant effects except for the variable trust ($F(1.248) = 0.68$, $p = 0.41$). For the respondents there is no difference in their trust perception whether the voice assistant had a human voice ($M = 2.94$, $SD = 0.04$) or an artificial voice ($M = 2.90$, $SD = 0.04$). In contrast to that, the multivariate analysis of variances has found effects of the humanity on the variable of anger ($F(1.248) = 28.44$, $p < 0.01$). Respondents who were confronted with a human voice experienced a lower level of anger ($M = 2.01$, $SD = 0.07$), than respondents who were confronted with an artificial voice ($M = 2.64$, $SD = 0.09$). Similar results were found for the level of stress ($F(1.248) = 21.95$, $p < 0.01$). People confronted with a human voice experienced a lower level of stress ($M = 2.41$, $SD = 0.07$), than people who were confronted with an artificial voice ($M = 2.95$, $SD = 0.09$). The emotion concentration showed comparable results ($F(1.248) = 5.88$, $p = 0.01$). When hearing a human voice, participants were able to concentrate better ($M = 3.30$, $SD = 0.07$), than when hearing an artificial voice ($M = 3.02$, $SD = 0.08$). The multivariate analysis of variances furthermore, reports a significant effect of the humanity of the voice assistant on the variable purchase intention ($F(1.248) = 7.47$, $p = 0.01$). Respondents are more likely to buy a car with a voice assistant if the VA has a human voice ($M = 2.74$, $SD = 0.04$) in contrast to if it has an artificial voice ($M = 2.60$, $SD = 0.04$).

**Table 5. MANOVA – Humanity**

| Dependent variable | Sum of sq. | df | Mean sq. | F | Sig. |
|---|---|---|---|---|---|
| Trust | 0.13 | 1 | 0.13 | 0.68 | 0.41 |
| Anger | 22.65 | 1 | 22.65 | 28.44 | <0.01 |
| Stress | 17.28 | 1 | 17.28 | 21.95 | <0.01 |
| Concentration | 3.95 | 1 | 3.95 | 5.88 | 0.01 |
| Purchase Intention | 1.35 | 1 | 1.35 | 7.47 | 0.01 |

**Table 2. Measurement Model Validation**

| Variable | Item | Cronbach's alpha | Composite reliability (CR) | Average variance extracted (AVE) | λ (factor loadings) |
|---|---|---|---|---|---|
| Trust | Q8_Trust_2 | | | | 0.73 |
| | Q8_Trust_3 | | | | 0.70 |
| | Q9_Trust_1 | 0.85 | 0.87 | 0.52 | 0.76 |
| | Q9_Trust_2 | | | | 0.82 |
| | Q9Trust_3 | | | | 0.62 |
| | Q10_Trust_1 | | | | 0.71 |
| Anger | Q11_Emotions_Anger_1 | | | | 0.80 |
| | Q11_Emotions_Anger_2 | 0.87 | 0.80 | 0.57 | 0.81 |
| | Q11_Emotions_Anger_3 | | | | 0.63 |
| Stress | Q12_Emotions_Stress_1 | | | | 0.78 |
| | Q12_Emotions_Stress_2 | 0.82 | 0.82 | 0.54 | 0.67 |
| | Q12_Emotions_Stress_3 | | | | 0.68 |
| | Q12_Emotions_Stress_4 | | | | 0.81 |
| Concentration | Q14_Emotions_Concentration_1 | | | | 0.82 |
| | Q14_Emotions_Concentration_2 | 0.75 | 0.75 | 0.50 | 0.58 |
| | Q14_Emotions_Concentration_3 | | | | 0.60 |
| Purchase Intention | Q15_PurchaseIntention_1 | | | | 0.88 |
| | Q15_PurchaseIntention_2 | 0.87 | 0.86 | 0.69 | 0.68 |
| | Q15_PurchaseIntention_3 | | | | 0.88 |

**Table 6. Mean and Standard Deviation – Humanity**

| | Human | | Artificial | |
|---|---|---|---|---|
| **Dependent variable** | **M** | **SD** | **M** | **SD** |
| Trust | 2.94 | 0.04 | 2.90 | 0.04 |
| Anger | 2.01 | 0.07 | 2.64 | 0.09 |
| Stress | 2.41 | 0.07 | 2.95 | 0.09 |
| Concentration | 3.30 | 0.07 | 3.02 | 0.08 |
| Purchase Intention | 2.74 | 0.04 | 2.60 | 0.04 |

## 4.4 Time of recording

The multivariate analysis of variances revealed no significant effects between the time of recording and the dependent variables. There has been no significant difference found for the perception of trust ($F(1.248) = 0.28$, $p = 0.59$) regarding the time of recording. Participants experience the same level of trust independently of the recording of the videos ($M = 2.93$, $SD = .04$), ($M = 2.90$, $SD = 0.04$). Furthermore, the multivariate analysis of variances found no significant effects of the time of recording on the variables anger ($F(1.248) = 0.02$, $p = 0.90$), stress ($F(1.248) = 0.13$, $p = 0.72$) and concentration ($F(1.248) = 1.95$, $p = 0.16$). Similar results were revealed for the dependent variable purchase intention ($F(1.248) = 0.053$, $p = 0.82$).

**Table 7. MANOVA – Time of recording**

| **Dependent variable** | **Sum of sq.** | **df** | **Mean sq.** | **F** | **Sig.** |
|---|---|---|---|---|---|
| Trust | 0.06 | 1 | 0.06 | 0.28 | 0.59 |
| Anger | 0.01 | 1 | 0.01 | 0.02 | 0.90 |
| Stress | 0.10 | 1 | 0.10 | 0.13 | 0.72 |
| Concentration | 1.31 | 1 | 1.31 | 1.95 | 0.16 |
| Purchase Intention | 0.01 | 1 | 0.01 | 0.05 | 0.82 |

**Table 8. Mean and Standard Deviation – Time of recording**

| | Day | | Night | |
|---|---|---|---|---|
| **Dependent variable** | **M** | **SD** | **M** | **SD** |
| Trust | 2.93 | 0.04 | 2.90 | 0.04 |
| Anger | 2.33 | 0.08 | 2.34 | 0.08 |
| Stress | 2.70 | 0.08 | 2.66 | 0.08 |
| Concentration | 3.23 | 0.08 | 3.10 | 0.08 |
| Purchase Intention | 2.67 | 0.04 | 2.66 | 0.04 |

## 4.5 Interaction effects

To analyze the data further and check the hypothesis and the research question, multivariate analysis of variances (MANOVA) was used to determine if there are interaction effects between the five dependent variables. The analysis revealed that gender, time of recording, and humanity significantly affect the variable stress ($F(2.148) = 4.18$, $p = 0.04$). Respondents indicated a higher level of stress when the gender of the voice was a female, artificial voice recorded during the day ($M = 2.84$, $SD = 0.17$), in contrast to when the voice was human ($M = 2.40$, $SD = 0.15$). Similar results were found when the recording time was during the night. Participants felt more stressed when the voice was a female, artificial voice during the night ($M = 3.12$, $SD = 0.16$) than a female, human voice during the day ($M = 2.15$, $SD = 0.16$). Comparable results were revealed when the voice of the assistant was male. People indicated a higher stress level when the voice was a male, artificial voice and the recording time was during the day ($M = 3.10$, $SD = 0.19$) as when the voice was a male, but human voice with the same recording time ($M = 2.50$, $SD = 0.14$). Lastly, the multivariate analysis of variances revealed a lower stress level for a human, male voice recorded during the night ($M = 2.62$, $SD = 0.14$) than for an artificial, male voice recorded during the night ($M = 2.77$, $SD = 0.19$).

**Table 9. MANOVA – Interaction effect Gender, Humanity & Time of recording**

| **Dependent variable** | **Sum of sq.** | **df** | **Mean sq.** | **F** | **Sig.** |
|---|---|---|---|---|---|
| Trust | 0.01 | 1 | 0.01 | 0.04 | 0.85 |
| Anger | 1.19 | 1 | 1.19 | 1.49 | 0.22 |
| Stress | 3.29 | 1 | 3.29 | 4.18 | 0.04 |
| Concentration | 0.53 | 1 | 0.53 | 0.79 | 0.38 |
| Purchase Intention | 0.13 | 1 | 0.13 | 0.72 | 0.40 |

## 4.6 Duration of driver's license

To check the correlation between the dependent variables and some characteristics of the participants, the data has been analyzed using an analysis of variances (ANOVA). The analysis focused on the correlation between the duration of being in charge of a driver's license and the dependent variables. The analysis of variances showed significant effects of the duration on the variable of stress ($F(5, 250) = 2.15$, $p = 0.03$). Participants who have a driver's license longer than ten years showed a lower level of stress ($M = 2.05$, $SD = 0.76$) than participants which are in charge of their driver's license for less than a year ($M = 2.61$, $SD = 1.14$).

**Table 10. ANOVA – Duration of driver's license**

| **Dependent variable** | **Sum of sq.** | **df** | **Mean sq.** | **F** | **Sig.** |
|---|---|---|---|---|---|
| Trust | 1.22 | 5 | 0.24 | 1.25 | 0.29 |
| Anger | 11.03 | 5 | 2.21 | 2.53 | 0.03 |
| Stress | 9.01 | 5 | 1.80 | 2.15 | 0.06 |
| Concentration | 2.17 | 5 | 0.43 | 0.63 | 0.68 |
| Purchase Intention | 1.38 | 5 | 0.28 | 1.50 | 0.19 |

**Table 11. Mean and Standard Deviation – Duration of driver's license**

| **Duration** | **N** | **M** | **SD** |
|---|---|---|---|
| < 1 year | 6 | 2.61 | 1.14 |
| 1 – 2 years | 21 | 2.57 | 0.99 |
| 3 – 5 years | 98 | 2.47 | 1.00 |
| 6 – 10 years | 37 | 2.25 | 1.06 |
| > 10 years | 94 | 2.05 | 0.76 |
| **Total** | **256** | | |

# 5. DISCUSSION AND IMPLICATIONS

The driving performance is highly influenced by external factors (Alvarez, Lopez-de Ipiña, Daily, & Gilbert, 2012). A possibility to reduce accidents and stabilize the emotional state of the driver would be by using voice assistants, which act as virtual partners (Eyben et al., 2010). Hereby, the right fit of the voice assistant is crucial to support the driver during one's

drive (Nass et al., 2005). The overall aim of the study was to find out if the use of a voice assistant might support the driver taken into account the different characteristics.

## 5.1 Gender

The experiment revealed no significant differences for the gender of the voice assistant. The respondents indicated the same level of trust perception independently of the gender the VA had. Even if current scientific literature suggests that a female voice is prioritized over a male voice (Eyssel et al., 2012), this could not be supported. There was no evidence found for the statement that women are perceived more as having a favorable voice (Borau et al., 2021). The hypothesis H2 ('Trust in a VA is higher when the VA has a female voice.') has therefore been rejected. Furthermore, the hypotheses H4a (‚The driver's emotional state is positively influenced if the VAs gender is female. ') and H4b ('The driver's emotional state is negatively influenced if the VAs gender is male. ') have been rejected as well. For the driver, the emotions of anger and stress and the ability to concentrate are not influenced by the gender of the voice assistant. Concerning the purchase intention, no significant results have been found. Respondents show the same willingness to pay the product independently of the gender of the voice assistant. Hypothesis H6 ('Purchase Intention is higher when the voice assistant has a female voice.') is therefore rejected. Although all hypotheses have been rejected and no significant effects were found, the analysis revealed interaction effects between the gender, humanity, and time of recording for the dependent variable stress. Respondents indicated a higher level of stress when the voice was a male, artificial voice, and the recording time was during the day ($M = 3.10$, $SD = 0.19$) as when the voice was a female, artificial voice with the same recording time ($M = 3.12$, $SD = 0.16$). This results also have been found for the human-likely voice. As current literature already suggest, women are seen as more capable of recognizing emotions like anger, empathy, and frustration (Borau et al., 2021), Therefore, using a female voice can help to reduce those negative emotions, resulting in fewer accidents (Eyben et al., 2010).

## 5.2 Humanity

According to the study results, the humanity of the voice assistant has a significant effect on two of the three variables. The results revealed that a human voice has a positive influence on the emotional state and the purchase intention of the driver. Participants indicated a lower level of stress and anger and a better possibility to concentrate when a human voice was used. This supports hypothesis H3 ('The driver's emotional state is better if the VA has a human-likely voice. '). More specifically, this supports the hypothesis H5a ('Anger is lower when a VA with a human-likely voice is used during the drive than when a VA with a robot-likely voice is used. '). A lower level of anger allows the driver to concentrate more on the road traffic and decreases the risk of accidents (Eyben et al., 2010). The hypothesis H5b ('Stress is lower when a VA with a human likely voice is used during the drive than when a VA with a robot likely voice is used. ') is supported as well. As a high level of stress leads to a loss of attention and focus, which results in lower driving performance, reducing the stress level helps to keep the driver's attention and increase the driving performance again (Eyben et al., 2010). Similar results have been revealed for the hypothesis H5d ('Concentration is higher when a VA with a human likely voice is used during the drive than when a VA with a robot likely voice is used. ') that has been accepted. To ensure a safe on-road behavior keeping concentration high is essential, as a lower level of concentration leads to an increase in accidents (Fofanova & Vollrath, 2011). Furthermore, the study revealed significant interaction effects between gender, humanity, time of recording and the variable stress. For both genders the stress level was higher when the voice was a robot-like voice than when it was a human vocie, taken into account the time of recording. Generally, this means that using a VA with a human voice helps to decrease negative emotions, which results in experiencing more positive emotions, which will lead to fewer accidents, as happier drivers are generally safer drivers (Zimasa et al., 2017). Lastly, for hypothesis H1 ('Trust in a VA is higher when the VA has a human likely voice. ') no significant effects have been found; therefore, the hypothesis has been rejected.

## 5.3 Time of Recording

Additionally, the videos were recorded at different times to check if the time of recording influences the three dependent variables. The conducted study found no significant differences between a recording by day and one by night. Furthermore, the participants did not indicate a difference in their trust perception, emotional state, or willingness to purchase the product.
An overview of all the hypothesis and their significance can be seen in the appendix (Appendix Part 7)..

## 5.4 Practical Implications

Based on the findings analyzed before, voice assistants' usage might help decrease accidents in nowadays traffic. Out of that, some main implications can be drawn. The main focus should lie on the humanity of the voice assistant. Today's VAs are usually spoken by an artificial voice. However, the conducted study revealed clear disproval of this usage, meaning that the level of anthropomorphism of the VA should be as human-like as possible. An assistant with a human voice leads to a lower level of stress and anger, but on the other hand, to a higher ability to concentrate. Current scientific literature shows that anger and stress negatively influence driving performance, resulting in a more risk-supporting attitude (Jeon et al., 2014). The findings of this study help to improve the development of future voice assistants and increase revenue for future cars. People have a higher purchase intention if the VA inside the car has a human-like voice compared to a computer voice. As the results suggest, gender makes no difference, but companies should make use of the humanity of voice assistants. This will lead to a decrease in negative emotions, a higher purchase intention, and a decrease in road accidents in the long term

# 6. CONCLUSION

The study was conducted to find an answer for the research question presented in the beginning:

*Which effects have the use of voice assistants on the drivers' emotional state?*

A total of eight hypotheses were tested in this relation. By testing the hypotheses, it became visible that different voice assistants have different emotional effects. This paper suggests that using voice assistants to support the driver is helpful if used with the right criteria. An essential characteristic is that the voice assistant should have a human or human-likely voice. The resulting effects were a lower level of stress and anger and a higher level of concentration compared to an artificial voice. Furthermore, the study did not

reveal any significant differences between the gender. No moderate impacts have been found for the variable of trust. However, an interaction effect between gender, humanity, and the time of recording on the variable of stress revealed significance. Concluding the study, the outcomes clearly reported that the implementation of a human voice shows significant effects on stress, anger and confusion, and the willingness to purchase the product. Therefore, the research question can be answered that the use of voice assistants might have aspects depending taken into account a human voice.

# 7. LIMITATIONS AND FURTHER RESEARCH

## 7.1 Limitations

The study adds new information to the current scientific literature. In addition, some significant results could have been revealed, which can be used for further research. However, there are some limitations worth mentioning.

The first one concerns the living location of the participants. The majority of respondents come from the same region in Germany. Therefore, cultural differences are not entirely covered in this research. It might be possible that inside of Germany, the different regions have different attitudes towards using voice assistants, which could lead to a change in the results. The second limitation concerns the nationality of the participants. Due to several crucial reasons, the study was only available in German. Participants needed to be able to speak German fluently. This forbids discovering cultural differences between different countries. A third limitation affects the design of the survey. Participants were not aware of the fact that different stimuli have been used with different characteristics of the voice assistant. As today's technology and artificial voices are similar to human voices, respondents struggled to indicate if the voice was a human voice ($M = 2.13$, $SD = 0.99$) or an artificial voice ($M = 2.85$, $SD = 0.99$). As a result, around 1/5 of the participants made the wrong choice. Even if it was not visible in the results, the manipulation check reported differences. Furthermore, the second manipulation question could not have been analyzed statistically due to missing answer choices. Even if all of the participants were right about the time of recording, the result is not statistically significant.

This leads to another limitation. As the videos were only recordings, respondents did not have a real-life experience of how voice assistants affect their driving behavior. Coombs (2020), mentioned that an essential aspect for consumers to judge the situation is depending on how they are affected by it. In the case of the survey, the respondents experienced the voice assistant only via a pre-recorded video. Therefore, results might change if the voice assistant would be used during a real-life experiment. The last limitation concerns the reliance of the scales of Beldad, van Laar, & Hegner (2018) , where two variables needed to be extracted, and Schweizer, Kotouc, Wagner, & Rudolph (2006), where all three variables needed to be extracted from the study. In this specific case relying on the scales of other researchers prohibited the study from discovering more explorative and robust research findings.

## 7.2 Further research

As mentioned in the limitations section, the study only focused on German participants. To test the obtained results further, the experiment should be repeated within different countries, regions, educational levels and cultural communities. This ensures that the findings from this research also show validity and significance in a further research.

Generally, the usage of voice assistants during the drive should be analyzed further using real life experiments. This paper is used as a starting point in this field. However, besides the analyzed dependent variables trust, emotions and purchase intention other variables as the emotional state or the driving performance should be analyzed. By testing the usage of voice assistants over a long time, an increase or decrease in the number of accidents might become visible, to support the findings and the current literature even further. Regarding the humanity of voice assistants, the study revealed several significant aspects. People generally perform better, when with humans, than with robots. Redesigning future robots, making them look more like humans helps to increase trust and the likelihood of buying the product. It furthermore, reduces negative emotions concerning the environment leading to a higher satisfaction with the product.

# 8. ACKNOWLEDGMENTS

First of all, I would like to thank my first supervisor, Dr. Carolina Herrando, for all of the support she gave me while writing my thesis. She provided me with extensive feedback and helped as a guide during the whole development process. She has been an enormous help during the last months and supported me in writing my thesis in the best possible way. Even if all communication was online due to the current COVID-19 pandemic, she supported me in the best possible way. Next to this, I also want to thank all of my participants who participated in my survey. But I want to thank especially those who helped me spread the survey by actively forwarding it. Without them, there would have been no qualifying data for me to analyze. Finally, I would like to thank my thesis circle, which was always willing to answer questions fast. With the knowledge of having people in the same situation, the development of the thesis became a lot easier.

Lastly, I would like to show my appreciation for my friends and family, especially my boyfriend, who supported and motivated me throughout the process and provided me with essential inputs.

## 9. REFERENCES


Adams, R.J. (2019). "Alexa, how can we increase trust in you?": An investigation of trust in smart home voice assistants. Retrieved June 26, 2021 from: https://essay.utwente.nl/78515/

Agarwal, R., & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision Sciences, 28*(3), 557-582. doi:10.1111/j.1540-5915.1997.tb01322.x

Agrawal, U., Giripunje, S., & Bajaj, P. (2013). Emotion and gesture recognition with soft computing tool for drivers assistance system in human centered transportation. *2013 IEEE International Conference on Systems, Man, and Cybernetics.* doi:10.1109/SMC.2013.785

Alarcón, D., Sánchez, J. A., & De Olavide, U. (2015). Assessing convergent and discriminant validity in the ADHD-R IV rating scale: user-written commands for Average Variance Extracted (AVE), Composite Reliability (CR), and Heterotrait-Monotrait ratio of correlations (HTMT). *Spanish STATA Meeting 2015.*

Alvarez, I., Lopez-de Ipiña, K., Daily, S. B., & Gilbert, J. E. (2012). Emotional adaptive vehicle user interfaces: moderating negative effects of failed technology interactions while driving. *Workshop of Automotive Natural Interfaces, together with International Conference on Automotive User Interfaces.*

Beldad, A. D., van Laar, E., & Hegner, S. M. (2018). Should the shady steal thunder? The effects of crisis communication timing, pre-crisis reputation valence, and crisis type on post-crisis organizational trust and purchase intention. *Journal of Contingencies and Crisis Management, 26*(1), 150-163. doi:10.1111/1468-5973.12172

Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing, 38*(7), 1052-1068. doi:10.1002/mar.21480

Bower, G. H. (1981). Mood and memory. *American psychologist, 36*(2), 129-148. doi:10.1037/0003-066X.36.2.129

Braun, M., Schubert, J., Pfleging, B., & Alt, F. (2019). Improving driver emotions with affective strategies. *Multimodal Technologies and Interaction, 3*(1), 1-19. doi:10.3390/mti3010021

Bubb, H. (2003). *Fahrerassistenz - primär ein beitrag zum komfort oder für die sicherheit?/Driver assistance - firstly a contribution to primary safety or rather to comfort?* Düsseldorf, DE: VDI-Berichte

Cabanac, M. (2002). What is emotion? *Behavioural Processes, 60*(2), 69-83. doi:10.1016/S0376-6357(02)00078-5

Chang, T.-Z., & Wildt, A. R. (1994). Price, product information, and purchase intention: An empirical study. *Journal of the Academy of Marketing Science, 22*(1), 16-27. doi:10.1177/0092070394221002

Chrysochou, P. (2017). Consumer behavior research methods. *Consumer Perception of Product Risks and Benefits*, 409-428. doi:10.1007/978-3-319-50530-5_22

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: a meta-analytic path analysis of 20 years of research, *Journal of Applied Psychology, 85*(5), 678-707. doi: 10.1037/0021-9010.85.5.678

Daneshvary, R., & Schwer, R. K. (2000). The association endorsement and consumers' intention to purchase. *Journal of Consumer Marketing, 17,* 203-213. doi:10.1108/07363760010328987

Deffenbacher, J. L., Lynch, R. S., Oetting, E. R., & Yingling, D. A. (2001). Driving anger: Correlates and a test of state-trait theory. *Personality and Individual Differences, 31*(8), 1321-1331. doi:10.1016/S0191-8869(00)00226-9

Derryberry, D., & Tucker, D. M. (1992). Neural mechanisms of emotion. *Journal of consulting and clinical psychology, 60*(3), 329-338. doi:10.1037/0022-006X.60.3.329

Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B., & Nguyen-Thien, N. (2010). Emotion on the Road: Necessity, acceptance, and feasibility of affective computing in the car. *Advances in Human-Computer Interaction vol. 2010*, 1-17. doi:10.1155/2010/263593

Eyssel, F., De Ruiter, L., Kuchenbrandt, D., Bobinger, S., & Hegel, F. (2012). "If you sound like me, you must be more human": On the interplay of robot and user features on human-robot acceptance and anthropomorphism. *7th ACM/IEEE International Conference on Human-Robot Interaction (HRI),* 125-126. doi:10.1145/2157689.2157717

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). Gender bias in chatbot design. *International Workshop on Chatbot Research and Design,* 79-93. doi:10.1007/978-3-030-39540-7_6

Fofanova, J., & Vollrath, M. (2011). Distraction while driving: The case of older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(6), 638-648. doi:/10.1016/j.trf.2011.08.005

Fukuyama, F. (1997). Trust: The social virtues and the creation of prosperity. *International Journal on World Peace, 14*(1), 84-87. Retrieved June 26, 2021, from http://www.jstor.org/stable/20752121

Harbluk, J. L., Noy, Y. I., Trbovich, P. L., & Eizenman, M. (2007). An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis & Prevention, 39*(2), 372-379. doi: /10.1016/j.aap.2006.08.013

Harris, H., & Nass, C. (2011). Emotion regulation for frustrating driving contexts. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Vancouver*. doi:10.1145/1978942.1979050

Hartley, J. (2014). Some thoughts on Likert-type scales. *International Journal of Clinical and Health Psychology, 14*(1), 83-86. doi:10.1016/S1697-2600(14)70040-7

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems, 6*(2), 156-166. doi:10.1109/TITS.2005.848368

Hennessy, D. A., & Wiesenthal, D. L. (1997). The relationship between traffic congestion, driver stress and direct versus indirect coping behaviours. *Ergonomics, 40*(3), 348-361. doi:10.1080/001401397188198

Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science, 49*(1), 30-50. doi:10.1007/s11747-020-00749-9

Jackson, S. A., & Marsh, H. (1996). Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of Sport & Exercise Psychology, 18*(1), 17-35. doi:10.1123/jsep.18.1.17

Jeon, M., Walker, B. N., & Yim, J.-B. (2014). Effects of specific emotions on subjective judgment, driving performance, and perceived workload. *Transportation Research Part F: Traffic Psychology and Behaviour, 24*, 197-209. doi: 10.1016/j.trf.2014.04.003

Jonsson, I.-M., & Chen, F. (2007). In-vehicle information S system used in complex and low traffic situations: Impact on driving performance and attitude. *Universal Access in Human-Computer Interaction, Ambient Interaction, Berlin, Heidelberg,* 421-430. doi:0.1007/978-3-540-73281-5_45

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *Current Journal of Applied Science and Technology*, *7*(4), 396-403. doi:10.9734/BJAST/2015/14975

Kachouie, R., Sedighadeli, S., Khosla, R., & Chu, M.-T. (2014). Socially assistive robots in elderly care: A mixed-method systematic literature review. *International Journal of Human–Computer Interaction, 30*(5), 369-393. doi:10.1080/10447318.2013.873278

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion, 5*(4), 345-379. doi:10.1007/BF00992553

Kok, J. N., Boers, E. J., Kosters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence, 1*, 270-299. Retrieved June 26, 2021 from: http://www.eolss.net/sample-chapters/c15/e6-44.pdf

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems, 16*(10), 880-918. doi: 10.17705/1jais.00411

LeDoux, J. E. (1992). Brain mechanisms of emotion and emotional learning. *Current Opinion in Neurobiology, 2*(2), 191-197. doi:10.1016/0959-4388(92)90011-9.

Levenstein, S., Prantera, C., Varvo, V., Scribano, M. L., Berto, E., Luzi, C., & Andreoli, A. (1993). Development of the perceived stress questionnaire: A new tool for psychosomatic research. *Journal of Psychosomatic Research, 37*(1), 19-32. doi: 10.1016/0022-3999(93)90120-5

Li, X., Hess, T. J., & Valacich, J. S. (2008). Why do we trust new technology? A study of initial trust formation with organizational information systems. *The Journal of Strategic Information Systems, 17*(1), 39-71. doi: 10.1016/j.jsis.2008.01.00

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review, 20*(3), 709-734. doi:10.2307/258792

McDonald, L., Glendon, A. I., & Sparks, B. (2011). Measuring consumers' emotional reactions to company crises: Scale development and implications. *NA - Advances in Consumer Research Volume 39,* 333-340. Retrieved from June 26, 2021 from: https://www.acrwebsite.org/volumes/1009627/volumes/v39/NA-39

McKnight, D., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research, 13*(3), 334-359. doi: 10.1287/isre.13.3.334.81

Mcknight, H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems, 2*(2), 1-25. doi:10.1145/1985347.1985353

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information systems research, 2*(3), 192-222. doi:10.1287/isre.2.3.192

Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., & Takayama, L. (2005). Improving automotive safety by pairing driver emotion and car voice emotion. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*. doi: 10.1145/1056808.1057070

Neuman, W. L. (2014). *Basics of social research* (3th ed.). London, UK: Pearson.

Pires, G., Stanton, J., & Eckford, A. (2004). Influences on the perceived risk of purchasing online. *Journal of Consumer Behaviour: An International Research Review, 4*(2), 118-131. doi:10.1002/cb.163

Poushneh, A. (2021). Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *Journal of Retailing and Consumer Services, 58*, 102-283. doi:10.1016/j.jretconser.2020.102283

Rasheed, F. A., & Abadi, M. F. (2014). Impact of service quality, trust and perceived value on customer loyalty in malaysia services industries. *Procedia-Social and Behavioral Sciences, 164*, 298-304. doi: 10.1016/j.sbspro.2014.11.080

Rezvani, S., Dehkordi, G. J., Rahman, M. S., Fouladivanda, F., Habibi, M., & Eghtebasi, S. (2012). A conceptual study on the country of origin effect on consumer purchase intention. *Asian Social Science, 8*(12), 205-215. doi:10.5539/ass.v8n12p205

Schweizer, M., Kotouc, A., Wagner, T., & Rudolph, T. (2006). Scale Development for consumer confusion. *NA - Advances in Consumer Research Volume 33,* 84-190. Retrieved June 26, 2021 from: http://www.acrwebsite.org/volumes/12287/volumes/v33/NA-33

Schwieren, J. (2020) Conversing with automated cars: Exploring voice assistants for trustful user interactions with fully automated vehicles. Retrieved June 26, 2021 from: https://essay.utwente.nl/81628/

Shabbir, M. S., Kirmani, M. S., Iqbal, J., & Khan, B. (2009). COO and Brand Names effect on consumer behavior and purchase intention in pakistan. *Interdisciplinary Journal of Contemporary Research in Business, 1*(3), 84-95. Retrieved June 26, 2021 from: https://www.researchgate.net/publication/341755928_COO_and_Brand_Name_s_effect_on_Consumer_Behaviour_and_Purchase_Intention_in_Pakistan

Stroessner, S. J., & Benitez, J. (2019). The social perception of humanoid and non-humanoid robots: Effects of gendered and machinelike features. *International Journal of Social Robotics, 11*(2), 305-315. doi: 10.1007/s12369-018-0502-7

Underwood, G., Chapman, P., Wright, S., & Crundall, D. (1999). Anger while driving. Transportation research part f: *Traffic Psychology and Behaviour, 2*(1), 55-68. doi:10.1016/S1369-8478(99)00006-6

Wang, C. L., Li, D., Barnes, B. R., & Ahn, J. (2012). Country image, product image and consumer purchase intention: Evidence from an emerging economy. *International Business Review, 21*(6), 1041-1051. doi:10.1016/j.ibusrev.2011.11.010

Wells-Parker, E., Ceminsky, J., Hallberg, V., Snow, R. W., Dunaway, G., Guiling, S., Anderson, B. (2002). An exploratory study of the relationship between road rage and crash experience in a representative sample of US drivers. *Accident Analysis & Prevention, 34*(3), 271-278. doi: 10.1016/S0001-4575(01)00021-5

Wilson, K. M., Yang, S., Roady, T., Kuo, J., & Lenné, M. G. (2020). Driver trust & mode confusion in an on-road study of level-2 automated vehicle technology. *Safety Science, 130*. doi:10.1016/j.ssci.2020.104845

Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist, 39*(2), 117–123. doi:10.1037/0003-066X.39.2.117

Zimasa, T., Jamson, S., & Henson, B. (2017). Are happy drivers safer drivers? Evidence from hazard response times and eye tracking data. Transportation research part f: Traffic. *Psychology and Behaviour, 46,* 14-23. doi:10.1016/j.trf.2016.12.005

# APPENDIX

**Part 1**: Route driven for the videos
**Part 2:** Variables with correlating items
**Part 3**: Survey
**Part 4**: Video Links
**Part 5:** Cross table analysis
**Part 6**: Demographic data about Respondents
**Part 7:** Overview of the hypotheses

## Part 1: Route driven for the videos


## Part 2: Variables with correlating items

| Variable | Item | Literature |
|---|---|---|
| **Trust** | The VA is concerned about my well-being. | Benevolence (Beldad, van Laar, & Hegner, 2018) |
| | The VA would not do anything the brings me in danger. | |
| | The VA does everything to support me. | |
| | The VA knows what to do. | Ability (Beldad et al., 2018) |
| | I am pretty sure about the abilities of the VA. | |
| | The VA is qualified in helping me. | |
| | The VA knows what is best for me. | Integrity (Beldad et al., 2018) |
| | The VA has helped me. | |
| | The advices of the VA were consistent. | |
| **Emotions** | The VA made me angry. | Anger (McDonald, Glendon, & Sparks, 2011) |
| | The VA made me enraged. | |

| | | |
|---|---|---|
| | I am outraged by the tips from the VA. | |
| | The VA frustrated me. | Stress (Levenstein et al., 1993) |
| | Through the tips from the VA I do not feel rested. | |
| | The tips from the VA have put me under pressure. | |
| | The VA bothered me. | |
| | The tips of the VA were hard to understand. | Confusion (Schweizer, Kotouc, Wagner, & Rudolph, 2006) |
| | The tips of the VA have irritated me. | |
| | I was unsure whether I can trust the advices. | |
| | Through the tips I was very focused. | Concentration (Jackson & Marsh, 1996) |
| | The tips helped me to know what to do. | |
| | The VA's tips helped me to have a higher concentration on hazardous situations. | |
| **Purchase Intention** | If I would have the choice, it is very likely, that I will buy a car with a VA. | (Beldad et al., 2018) |
| | If I would have the choice, it is very likely that I would buy a car with the exact same VA. | |
| | If I would have the choice, it is very unlikely that I will buy a car with a VA. | |

# Part 3: Survey

Sehr geehrte*r Teilnehmer*in,
du wurdest hiermit eingeladen, an einer Forschungsstudie mit dem
Titel „Emotions triggered by Artifical Intelligence“ teilzunehmen. Diese Studie wird von Anna-Maja Wolf von der Fakultät für Verhaltens-, Management- und Sozialwissenschaften der University of Twente geleitet und durchgeführt.
Ziel dieser Studie ist es, Einblicke in das emotionale Verhalten von Autofahrern zu bekommen. Die Teilnahme an diesem Experiment dauert ungefähr 10 Minuten. Falls ihr die Seite ausversehen aktualisieren solltet, werdet ihr an den Anfang der Studie zurück geführt. Probiert dies also wenn möglich bitte zu vermeiden.
Die Teilnahme an dieser Studie ist anonym und es können keine persönlichen Informationen auf dich zurückgeführt werden. Daher kann und wird deine Identität in keiner Veröffentlichung, die basierend auf den gesammelten Daten dieser Studie entsteht, offengelegt. Deine Daten werden außerdem nur zur Analyse von Gesamtergebnissen verwendet. Das heißt, dass niemals einzelne

Antworten einer einzelnen Person isoliert und unabhängig betrachtet oder verwendet werden. Deine Daten werden des Weiteren ausschließlich für diese Studie und nicht für andere Zwecke verwendet.
Die Teilnahme an dieser Studie ist völlig freiwillig. Du kannst dich jeder Zeit dazu entscheiden, nicht mehr an dieser Studie teilzunehmen, durch das Schließen des Fensters oder deine Einwilligung zur Teilnahme am Ende widerrufen. Wenn du dich dazu entscheidest, nicht an dieser Studie teilzunehmen oder dich von dieser zurückziehen willst, wirst du hierfür nicht bestraft.
Unter allen Teilnehmern wird ein 20€ Gutschein verlost, welchen du in einem Onlineshop deiner Wahl einlösen kannst.
Bei möglichen Fragen wende dich gerne an Anna-Maja Wolf (a.wolf-1@student.utwente.nl)

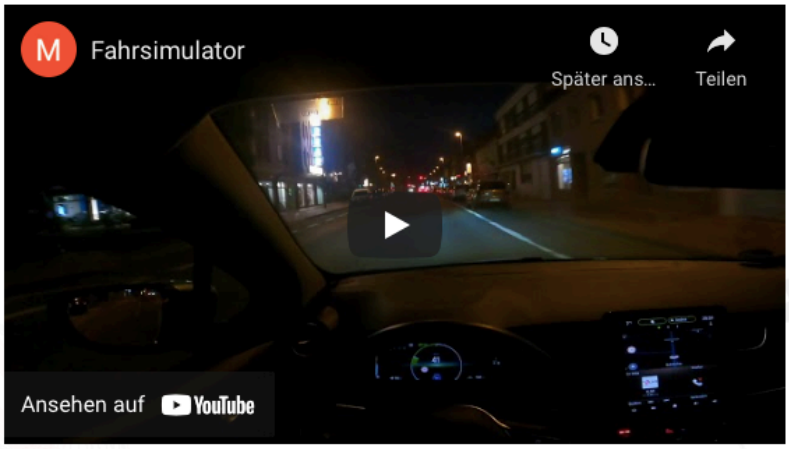
Einverständniserklärung
Ich bestätige hiermit, dass ich die oben genannten Informationen gelesen, oder sie mir vorlesen lassen habe, und diese danach verstanden habe. Ich wurde darüber informiert, dass ich zu jeder Zeit Fragen bezüglich dieser Studie stellen kann und diese zu meiner Zufriedenheit beantwortet werden. Ichhabe mich freiwillig dazu entschieden an dieser Studie teilzunehmen, und verstehe, dass ich mich jederzeit ohne Angabe von Gründen von der Teilnahme dieser Studie zurückziehen kann. Ich habe verstanden, dass meine Daten anonym interpretiert werden und diese ausschließlich für die Studie verwendet werden.

- o Ich erkläre mich hiermit mit den oben genannten Bedingungen einverstanden.

| | |
|---|---|
| *Frage 1 – Alter* | Wie alt bist du?<br>o Zwischen 18 und 21<br>o Zwischen 22 und 25<br>o Zwischen 25 und 34<br>o Zwischen 35 und 50<br>o Älter als 50 |
| *Frage 2 – Geschlecht* | Welches Geschlecht hast du?<br>o Männlich<br>o Weiblich<br>o Diverse |
| *Frage 3 – Wohnort* | Aus welchem Bundesland kommst du?<br>o Baden Württemberg<br>o Bayern<br>o Berlin<br>o Brandenburg<br>o Bremen<br>o Hamburg<br>o Hessen<br>o Mecklenburg-Vorpommern<br>o Niedersachsen<br>o Nordrhein Westfalen<br>o Rheinland-Pfalz<br>o Saarland<br>o Sachsen<br>o Sachsen-Anhalt<br>o Schleswig-Holstein<br>o Thüringen |
| *Frage 4 – Führerschein* | Wie lange besitzt du schon deinen Führerschein?<br>o Weniger als ein Jahr<br>o Zwischen einem und zwei Jahren<br>o Zwischen drei und fünf Jahren<br>o Zwischen sechs und zehn Jahren<br>o Länger als zehn Jahre II |
| *Frage 5 – Fahrverhalten* | Wie oft in der Woche fährst du Auto?<br>o Jeden Tag<br>o Zwei bis drei-mal die Woche<br>o Vier bis sechs-mal die Woche<br>o Einmal die Woche<br>o Weniger |

| | |
|---|---|
| *Frage 6 – Fahrverhalten* | Beschreibe deinen Fahrstil.<br>○ Ich bin ein ruhiger Fahrer.<br>○ Ich fahre sehr umsichtig und bedacht.<br>○ Ich lasse mich sehr leicht aus der Ruhe bringen beim Fahren.<br>○ Ich bin ein sehr ungeduldiger Fahrer.<br>○ Ich bin meistens gestresst beim Fahren. |
| *Frage 7 – Fahrerfahrung* | Lies dir bitte folgende Aussage durch, und gib an ob die Aussage zutrifft oder eher nicht.<br>Ich bin ein erfahrender Fahrer.<br>○ Stimme ich überhaupt nicht zu.<br>○ Stimme ich nicht zu.<br>○ Neutral<br>○ Stimme ich zu.<br>○ Stimme ich voll und ganz zu. |

Bitte sieh und höre dir das folgende Video so genau es geht an und klicke danach auf weiter.




Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Der Sprachassistent ist um mein Wohlbefinden besorgt. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent würde nichts tun, was mich in Gefahr bringt. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent tut alles, um mir zu helfen. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Der Sprachassistent weiß, was er tun muss. | ○ | ○ | ○ | ○ | ○ |
| Ich bin sehr sicher, über die Fähigkeiten des Sprachassistenten. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent ist nicht qualifiziert mir zu helfen. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Der Sprachassistent weiß, was gut für mich ist. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent hat mir geholfen. | ○ | ○ | ○ | ○ | ○ |
| Die Tipps des Sprachassistenten waren konsequent. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Der Sprachassistent hat mich sauer gemacht. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent hat mich verärgert. | ○ | ○ | ○ | ○ | ○ |
| Ich bin empört über die Anweisungen den Sprachassistenten. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Der Sprachassistent hat mich frustriert. | ○ | ○ | ○ | ○ | ○ |
| Durch die Anweisungen des Sprachassistenten fühle ich mich nicht erholt. | ○ | ○ | ○ | ○ | ○ |
| Die Anweisungen des Sprachassistenten haben mich unter Druck gesetzt. | ○ | ○ | ○ | ○ | ○ |
| Der Sprachassistent hat mich aufgeregt. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Die Anweisungen des Sprachassistenten waren schwer zu verstehen. | ○ | ○ | ○ | ○ | ○ |
| Die Anweisungen des Sprachassistenten haben mich irritiert. | ○ | ○ | ○ | ○ | ○ |
| Ich war unsicher, ob ich den Anweisungen vertrauen konnte. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Durch die Anweisungen des Sprachassistenten war ich sehr fokussiert. | ○ | ○ | ○ | ○ | ○ |
| Durch die Anweisungen des Sprachassistenten wusste ich, was ich machen sollte. | ○ | ○ | ○ | ○ | ○ |
| Durch die Anweisung des Sprachassistenten, habe ich eine erhöhte Aufmerksamkeit für gefährliche Situationen gehabt. | ○ | ○ | ○ | ○ | ○ |

Lies dir bitte folgende Aussagen durch, und gib an ob die Aussagen zutreffen oder eher nicht.

| | Stimme ich überhaupt nicht zu | Stimme ich nicht zu | Neutral | Stimme ich zu | Stimme ich voll und ganz zu |
|---|---|---|---|---|---|
| Wenn ich die Möglichkeit habe, ist es sehr wahrscheinlich, dass ich ein Auto mit einem Sprachassistenten kaufen werde. | ○ | ○ | ○ | ○ | ○ |
| Wenn ich die Möglichkeit habe, ist es sehr wahrscheinlich, dass ich ein Auto mit demselben Sprachassistenten kaufen würde. | ○ | ○ | ○ | ○ | ○ |
| Wenn ich die Möglichkeit habe, ist es sehr wahrscheinlich, dass ich ein Auto ohne Sprachassistenten kaufen werde in Zukunft. | ○ | ○ | ○ | ○ | ○ |

Wenn ich darüber nachdenke ein Auto zu kaufen, ist mir folgendes wichtig (mehrfach ankreuzen möglich)

- ☐ Eine Einparkhilfe (vorne und/oder hinten)
- ☐ Eine Kamera
- ☐ Einen Sprachassistenten, ähnlich wie den im Video
- ☐ Eine Sitz- und/oder Lenkradheizung
- ☐ Ein Navigationsgerät
- ☐ Ein Automatikgetriebe
- ☐ Ich habe keine Ansprüche an mein Auto
- ☐ Ich habe andere Ansprüche an mein Auto

| | |
|---|---|
| *Frage 17 – Manipulation Check* | Die Stimme des Sprachassistenten war<br>o Eine menschliche, weibliche Stimme<br>o Eine künstliche, weibliche Stimme<br>o Eine menschliche, männliche Stimme<br>o Eine künstliche, männliche Stimme |
| *Frage 18 – Manipulation Check* | Wann wurde das Video aufgenommen?<br>o Tagsüber<br>o Nachts |

**Das war's schon! Vielen Dank, dass du mitgemacht hast.**
Ich möchte dich hiermit darüber aufklären, dass dir der wahre Zweck des Experiments zu Beginn vorenthalten wurde. Das eigentlich Ziel der Studie ist, festzustellen, ob unterschiedliche Charakteristiken von Sprachassistenten unterschiedliche Auswirkungen auf den emotionalen Zustand des Fahrers haben könne.

Wenn du nach dieser Aufklärung deine Teilnahme an der Studie widerrufen willst, wähle dies bitte hier aus. Wenn du weiterhin teilnehmen möchtest, teile mir dies bitte ebenfalls mit, indem die passende Option unten auswählst.

<u>Einverständniserklärung</u>

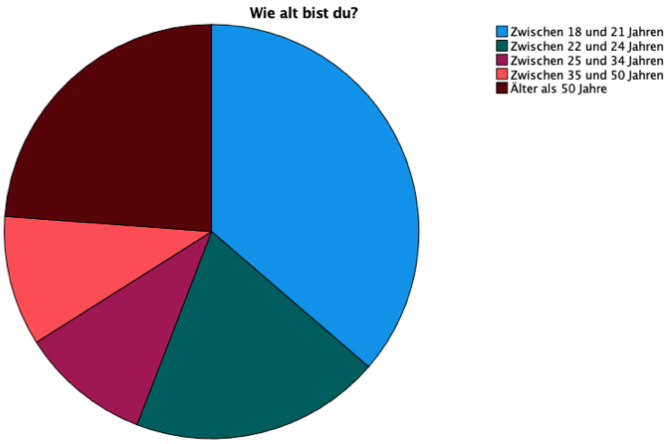
- o Ich bin damit einverstanden, dass meine Antworten weiterhin verwendet werden dürfen.
- o Ich bin nicht damit einverstanden, dass meine Antworten weiterhin verwendet werden dürfen.

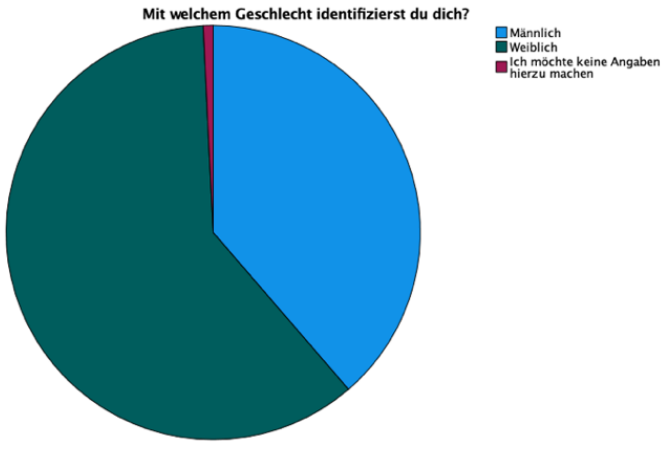
Wenn du an den Ergebnissen der Studie interessiert bist und du an dem Gewinnspiel teilnehmen willst, trage bitte hier deine E-mail Adresse ein:

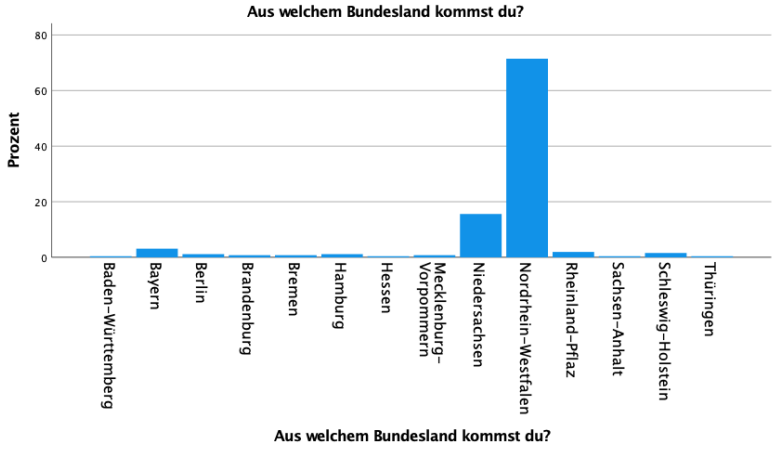
## Part 4: Video Links

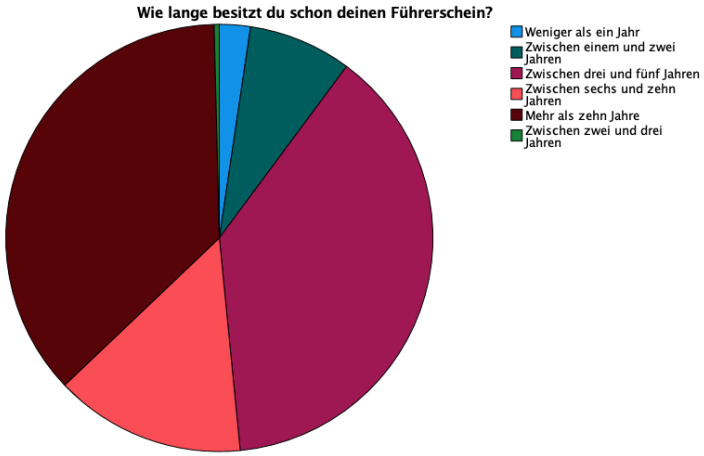
Scenario 1
https://youtu.be/q-Y_5zl_VSY
Scenario 2
https://youtu.be/UP1biFg5FHc
Scenario 3
https://youtu.be/4W0oE9F7IDg
Scenario 4
https://youtu.be/t7i7mhaQHrc
Scenario 5
https://youtu.be/dfCqo3JlhKk
Scenario 6
https://youtu.be/3l_LtqgcsDk
Scenario 7
https://youtu.be/RRFvUFniSmU
Scenario 8
https://youtu.be/WP_dh1WgCa0

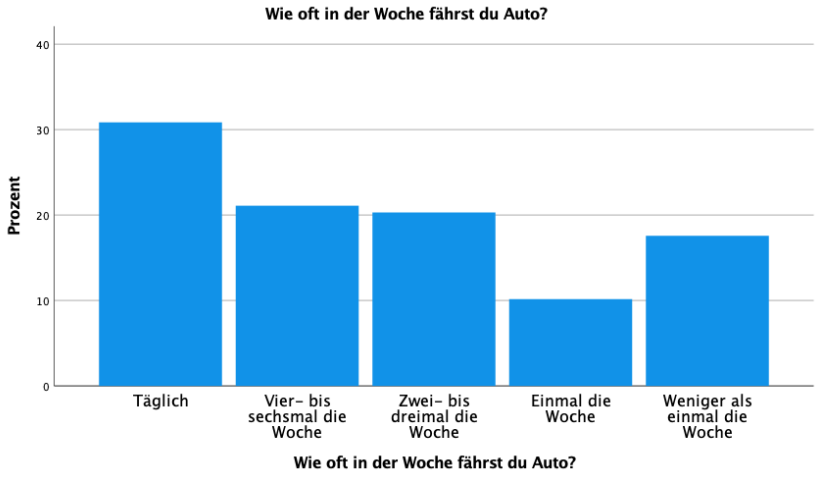
## Part 5: Demographic Data about Respondents




Aus welchem Bundesland kommst du?
Prozent
80
60
40
20
0
Baden-Württemberg
Bayern
Berlin
Brandenburg
Bremen
Hamburg
Hessen
Mecklenburg-Vorpommern
Niedersachsen
Nordrhein-Westfalen
Rheinland-Pflaz
Sachsen-Anhalt
Schleswig-Holstein
Thüringen
Aus welchem Bundesland kommst du?


Wie lange besitzt du schon deinen Führerschein?
Weniger als ein Jahr
Zwischen einem und zwei Jahren
Zwischen drei und fünf Jahren
Zwischen sechs und zehn Jahren
Mehr als zehn Jahre
Zwischen zwei und drei Jahren


Wie oft in der Woche fährst du Auto?
Prozent
40
30
20
10
0
Täglich
Vier- bis sechsmal die Woche
Zwei- bis dreimal die Woche
Einmal die Woche
Weniger als einmal die Woche
Wie oft in der Woche fährst du Auto?

**Häufigkeiten von $Fahrverhalten**

| | | Antworten | | Prozent der Fälle |
|---|---|---|---|---|
| | | N | Prozent | |
| Fahrverhalten[a] | Beschreibe dein Fahrverhalten (mehrfach ankreuzen möglich) Ich bin ein ruhiger Fahrer*in. | 158 | 37,9% | 61,7% |
| | Beschreibe dein Fahrverhalten (mehrfach ankreuzen möglich) Ich fahre sehr umsichtig und bedacht. | 174 | 41,7% | 68,0% |
| | Beschreibe dein Fahrverhalten (mehrfach ankreuzen möglich) Ich lasse mich sehr leicht aus der Ruhe bringen beim Fahren. | 35 | 8,4% | 13,7% |
| | Beschreibe dein Fahrverhalten (mehrfach ankreuzen möglich) Ich bin ein sehr ungeduldiger Fahrer*in. | 32 | 7,7% | 12,5% |
| | Beschreibe dein Fahrverhalten (mehrfach ankreuzen möglich) Ich bin meistens gestresst beim Fahren. | 18 | 4,3% | 7,0% |
| Gesamt | | 417 | 100,0% | 162,9% |

a. Dichotomie-Gruppe tabellarisch dargestellt bei Wert 1.

# Part 6: Cross table analysis

**Wann wurde das Video aufgenommen? * Scenario Kreuztabelle**

Anzahl

| | | Scenario | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 | Scenario 5 | Scenario 6 | Scenario 7 | Scenario 8 | Gesamt |
| Wann wurde das Video aufgenommen? | Tagsüber | 0 | 34 | 0 | 32 | 31 | 31 | 0 | 0 | 128 |
| | Nachts | 30 | 0 | 34 | 0 | 0 | 0 | 31 | 33 | 128 |
| Gesamt | | 30 | 34 | 34 | 32 | 31 | 31 | 31 | 33 | 256 |

## Part 7: Overview of hypotheses

| Hypothesis | |
|---|---|
| H1: Trust in a VA is higher when the VA has a human likely voice than when the VA has a robot-likely voice. | Rejected |
| H2: Trust in a VA is higher when the VA has a female voice than when the VA has a male voice. | Rejected |
| H3: The driver's emotional state is better if the VA has a human-likely voice than when the VA has a robot-likely voice. | Accepted |
| H4a: The driver's emotional state is positively influenced if the VAs gender is female. | Rejected |
| H4b: The driver's emotional state is negatively influenced if the VAs gender is male. | Rejected |
| H5a: Anger is lower when a VA with a human likely voice is used during the drive than when a VA with a robot-likely voice is used. | Accepted |
| H5d: Concentration is higher when a VA when a VA with a human likely voice is used during the drive than when a VA with a robot likely voice is used. | Accepted |
| H6: Purchase Intention is higher when the voice assistant has a female voice than when the VA has a male voice. | Rejected |