



Conversing with Automated Cars:

Exploring Voice Assistants for Trustful User Interactions with fully Automated Vehicles

Julian Schwieren, 1877828

Bachelor Thesis in Communication Science (BSc)

Faculty of Behavioral Management and Social Sciences (BMS)

Supervisor: Dr. H.A. Van Vuuren

University of Twente

June 26, 2020

**UNIVERSITY
OF TWENTE.**

Abstract

This paper explores the effects of a voice assistant's (VA) speech quality (synthetic, natural human voice) and gender (female, male) on the user's trust in the assistant (measured by integrity, competence, benevolence) as well as in the automated vehicle (AV) (measured by system transparency, technical competence, situation management), in which the VA operates. First, trust in the VA was expected to positively affect trust in the AV. Second, assistants with a natural human voice were expected to score best on all the given constructs. Third, a female voice was expected to score better on all constructs, except for competence and technical competence, where a male voice was seen superior. The proposed hypotheses have been tested with an experimental 2x2 study design, incorporating four videos with manipulated VAs, followed by a survey (N=100). It was identified that trust in the VA causes trust in the AV, further stimulating the need to explore the design of trustful VAs. Accordingly, organizations can improve the adoption of AVs, and finally decrease human-caused road fatalities. However, no main effects were found for speech quality or gender. Lastly, the study contributes to future work in the same domain, by giving recommendations based on the findings and identified shortcomings.

Keywords: Automated Vehicle, Self-driving Car, Autonomous Vehicle, Voice Assistant, Digital Assistant, Speech Quality, Gender, Human-Computer-Interaction, Trust

Table of Content

2. Theoretical Framework	7
2.1. Trust in Automated Vehicles	7
2.2. Trust in Voice Assistants	8
2.2.1. Speech Quality of Voice Assistants	9
2.2.2. Gender of Voice Assistants	10
2.2.3. Interaction of a Voice Assistant's Speech Quality and Gender	11
2.3. Conceptual Model	14
3. Method	15
3.1. Procedure	15
3.2. Stimulus material:	15
3.3. Pre-test	17
3.4. Design, Instruments, and Measurements	17
3.5. Respondents	18
3.6. Validity	19
3.7. Reliability	20
4. Results	22
4.1. The Effects of Trust in the Voice Assistant on Trust in the Automated Vehicle	22
4.2. Multivariate Analysis of Variance	22
4.3. Main Effects of Gender and Speech Quality	23
4.3.1. Effects on Trust in the Voice Assistant	23
4.3.2. Effects on Trust in the Automated Vehicle	24
5. Discussion and Conclusion	27
5.1. Discussion of Findings	27
5.2. Limitations	30
5.3. Conclusion	31
6. Literature	32
7. APPENDIX	37
Appendix A – Final Questionnaire	37
Appendix B – Manipulated Videos	45
Appendix C – Literature Study Log	45

1. Introduction

Fully automated vehicles (AVs), also referred to as self-driving-, driverless-, and autonomous vehicles, define society's expectations and hopes for the future of traffic. Their promise – overcoming human mental shortcomings to make traffic safer and more efficient. Within the stage of full automation, which is depicted as the 5th and final level in the widely accepted scale of SAE International (2020), cars can drive without manual interferences in any circumstances. Therefore, the human driver's role shifts to that of a passenger, or user, of the technology. As a consequence, safety, efficiency, and mobility are improved (Beiker & Calo, 2010). Theoretically speaking, that stage is the optimal solution for safe traffic – assuming that the technology is “perfect” in sensing the environment, has “perfect” decision making algorithms, and “perfect” actuators (Kyriakidis, Happee, & Winter, 2015).

Based on the promising possibilities, countries and institutions seek chances in the technology. For instance, has the European Union already disclosed the plan to transform traffic and industry accordingly. Moreover, could the invention significantly reduce human-caused road fatalities, which make up for 95% of all deadly traffic accidents (“Self-driving cars in the EU”, 2019). Although the EU has accomplished a decrease in road deaths of 20.7% in the period of 2010 to 2018 (“Road Deaths in the European Union”, n.d.), they are still behind their initial goal of lowering the number by 43%. However, even though some modern cars have reached partial automation, a fully automated vehicle that is safe enough to interact legally within the EU, has not been invented yet.

Not only do fully automated cars redefine traffic, industry, and the role that humans play within both; they also bring up new ethical questions about responsibility, behavior, intelligence, and autonomy. Instead of focusing on these domains only, Coeckelbergh (2016) asserts that society should pay more attention to shifts in subjectivity and perception. Specifically, the potential change of the humans' perception of other traffic entities, may shift from “others” to “machines”, or to “quasi-others”. Hence, people shall not only be required to act responsibly, but rather it shall be asked *how* people can act responsibly. Since opinions on these topics can differ, looked upon from diverse cultural backgrounds, Coeckelbergh (2016) stresses the importance of considering cultural differences in future work.

Moreover, people possess different attitudes towards fully automated cars, determining their acceptance of the technology. Nielsen & and Haustein (2018) segment potential users in three different stages: skeptics, indifferent stressed drivers, and enthusiasts. While enthusiasts are mainly male, young, highly educated, and live in large urban areas, the less trusting skeptics are older, the most car reliant, and come from less densely populated areas more often. In

addition, König and Neumayr (2017) underline the real concerns among car drivers towards AVs. Especially people that own a car that provides little automation features and pleasure drivers seem to dislike the technology.

Consequently, one that seeks to improve the acceptance of highly automated vehicles must improve the users' willingness to engage with the technology through changing their attitude. In that manner, literature stresses the significant effect of the users' trust – “the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability” (Lee & See, 2004) – deciding about their acceptance of AVs (Choi & Ji, 2015; Kaur & Rampersad, 2018; Raue et al., 2019; Zhang et al., 2020).

As cars are progressively being improved in automation, the nature of how people interact with the driving technology changes. Thereupon, several studies have investigated human-like voice assistants (VAs) within automated cars (Chi, Dewi & Huang, 2017; Ji, Liu, & Lee, 2019; Wong, Brumby, Babu, & Kobayashi, 2019). These assistants are often programmed with human-like features, such as gender traits, overt personality, and a role (that of an assistant) (Guzman, 2019). Since anthropomorphizing symbolic information in AVs has been identified as improving trust in the vehicle (Niu, Terken, & Eggen, 2018), one can reason that a trustful and human-like VA might contribute to the users' attitude towards the technology. However, digital personal assistants, such as VAs, are still underexplored within Human-Computer Interaction (Søndergaard & Hansen, 2018), with open questions on how to specifically design trustful agents.

In order to shed light on this area, did Ji, Liu, and Lee (2019) gather data of 30 participants about their preferences on gender (male and female) and voice content (“how” and “why”) for a text-to-speech (tts) VA in an AV. They conducted a 2x2 study, where participants watched a first-person point of view video of a ride in an AV while listening to the modified VA. Their findings suggest no initial preference for gender, but higher ratings for the female voice in trust, acceptance, and pleasure after the study. A significant number of respondents mentioned that female voices are more familiar to them due to other commercial VAs. Regarding the voice content, their results show significant preferences in messages that were framed as “how” messages. However, their findings are limited by a small number of participants and their geographic location in China and Japan, potentially manifesting results that could differ in other cultures. Finally, they suggest future work on the interaction of speech quality (natural human voice, tts) and voice gender (male, female) since these interactions have not been studied yet.

To fill this research gap and to contribute to the open questions in designing trustful

VAs for fully automated driving technology, this paper's focus lies on the interaction of a VA's speech quality and voice gender. Thereby, the findings could give insights in how to design trustful interactions of users and AVs. Moreover, users' increased trust may add to the transformation of the current traffic standard into a fully automated one. Thus, people may adapt to the new technology faster, potentially leading to less fatal traffic incidents caused by human error. Finally, the study's potential contributions are aligned with the EU's vision of traffic and industry transformation, as well as with their specific goal of reducing road fatalities. Based on the presented literature, the following research questions have been formulated:

***RQ1:** How is the user's trust in the fully automated vehicle affected by the user's trust in the voice assistant?*

***RQ2:** How is the user's trust in the fully automated vehicle affected by the voice assistant's gender?*

***RQ3:** How is the user's trust in the fully automated vehicle affected by the voice assistant's speech quality?*

***RQ4:** How is the user's trust in the fully automated vehicle affected by the interaction of the voice assistant's gender and speech quality.*

In order to answer these research questions, the combination of the VA's speech quality (natural human voice, tts) and the VA's voice gender (male, female) will be examined on their influence on trust in VAs and its dimensions 1) integrity, 2) competence, and 3) benevolence, as well as on trust in AVs and its trusting beliefs a) system transparency, b) technical competence, and c) situation management. The variables and constructs will be further substantiated in the theoretical framework and measured by means of a 2x2 experimental study design, which is explained in the method section. Finally, the outcome of the study is presented in the results part and further interpreted in this paper's discussion section.

2. Theoretical Framework

2.1. *Trust in Automated Vehicles*

Trust has been found to be a critical belief that determines if users are willing to use a specific automation technology (Lee & Moray, 1994; Muir & Moray, 1996; Sheridan, 1975; Sheridan & Hennessy, 1984). Moreover, do Lee and Moray (1994) reason that the absence of a user's trust might lead to incorrect usage of automated devices. However, nearly 80% of trust in automation studies was conducted in highly specified fields, such as military, security, and industry (Hoff & Bashir, 2015). Thus, Hoff and Bashir (2019) conclude that it remains unclear if the results from high criticality domains can be transferred to the process of designing more consumer-oriented systems. Despite their conclusion, several studies suggest that trust also determines the acceptance of automated driving technology (Choi & Ji, 2015; Gold, Körber, Hohenberger, Lechner, & Bengler, 2015; Kaur & Rampersad, 2018; Raue et al., 2019; Zhang et al., 2020).

In order to examine technology acceptance in AVs, literature proposes a variety of different models, such as the theory of reasoned action (TRA), the motivational model (MM), the unified theory of acceptance and use of technology (UTAUT) as well as the technology acceptance model (TAM). The latter has been extended with the factor trust, or modified to examine trust, in a variety of studies and suggests investigating technology acceptance by studying the effect of the factors perceived usefulness and perceived ease of use, on attitude toward using, behavioral intention, and finally on actual system use (Davis, 1989). These models have been constructed for fields such as mobile internet (Alalwan, Baabdullah, Rana, Tamilmani, & Dwivedi, 2018), online service applications (Pavlou, 2003; Wu & Chen, 2005), automation technologies (Ghazizadeh, Lee, & Boyle, 2011) and also for AVs specifically (Zhang et al., 2019). For AVs, Zhang et al. (2019) propose a TAM that has been extended by the factors trust and social influence.

Despite the scientific consensus on trust – being a critical belief in the acceptance of AVs – results differ in rating its importance. Next to other studies, that stress the role of trust as either significant (Kaur & Rampersad, 2018), supplementary (Buckley, Kaye & Pradhan, 2018) or as mediated by other factors (Choi & Ji, 2015), did Zhang et al. (2019) underline trust as a key factor, that influences the intention to use AVs in the first place. Hence, they argue that studies that do not highlight trust as a key factor suffer from the limitations of their research models or survey population.

In order to examine trust in automated cars, Choi and Ji (2015) sum up the suggestions in the literature and establish three dimensions of trust in technology, that all correspond to an

interpersonal trusting belief: a) system transparency: “the degree to which users can predict and understand the operating of autonomous vehicles”, b) technical competence: “the degree of user perception on the performance of the autonomous vehicles” and c) situation management: “the user’s belief that he or she can recover control in a situation whenever desired”. Their findings show that 47.4% of the construct trust’s variance was explained by these three trusting beliefs. Therefore, they conclude that these dimensions are suited for researching trust in AVs. Due to their specific context, these three trusting beliefs will serve to examine the effects on trust in AVs in the present research.

2.2. Trust in Voice Assistants

Voice assistants (VAs), alternatively referred to as digital assistants, virtual assistants, and (artificially) intelligent personal assistants (Golden & Fleischmann, 2018), can help people to manage their lives, be in control, facilitate, assist, inform, and entertain. Moreover, their application is developing into contextual habits; At home, voice assistants help to remain relaxed and, in a car, they are highly functional (Huisman & Huisman, 2018). Furthermore, they assist users in performing tasks while using fewer mental resources. Thus, users can easily multi-task, communicate more effectively, and have enough support in a case of restricted mobility (Boonrod & Ketayan, 2018). Since human factors, such as panic reactions, mental handicaps, and inattention lead to traffic accidents (Bucsuházy et al., 2020), reducing cognitive overload through voice assistants could lead to higher traffic safety levels, as long automation is not covering these threats already.

As aforementioned, digital assistants are often designed with human-like features. According to Golden and Fleischmann (2018), this development has been influenced by the users’ lack of understanding the technology. For instance, users can hear a smile in the voice of a virtual assistant and tend to prefer a “smiling voice” over a neutral one (Torre, Goslin, & White, 2020).

On the contrary, too much humanness could lead to a decrease in trust in artificially intelligent VAs (Garcia & Lopez, 2019). Following Dr. Masahiro Mori’s theory of the uncanny valley (Mori, MacDorman & Kageki, 2012), the likeability of humans towards anthropomorphized robots does increase to a certain point, where they are too human-like, which humans will react negatively upon. Nonetheless, a certain level of anthropomorphism seems to have a positive effect on trust in VAs, as well as on trust in AVs (Niu et al., 2018).

Concerning examining trust in technology, Lankton et al. (2015) propose a distinction between system-like and human-like technology. With this, they state that developing trust in

technology fundamentally depends on its humanness. In particular, they identify three trusting beliefs for human-like technology in the literature, which are 1) integrity: “the belief that a trustee adheres to a set of principles that the trustor finds acceptable” (Mayer, Davis, & Schoorman, 1995), 2) competence: “the belief that the trustee has the ability to do what the trustor needs to have done” (McKnight, Choudhry, & Kacmar, 2002), and 3) benevolence “the belief that the trustee will want to do good to the trustor, aside from an egocentric profit motive” (Mayer et al., 1995). Their suggested trusting beliefs have been further studied by Califf, Brooks, and Longstreet (2020) in the context of online travel booking websites, that further stress that human-like and system-like trusting beliefs should be balanced on the user, technology, and context. Since the three specified trusting beliefs of Choi and Ji (2015) serve to cover the system-like dimensions of the AV, the human-like trusting beliefs are used to measure the trust regarding the human-like voice assistant only.

Finally, based on the potential of anthropomorphizing symbolic information in AVs to increase trust in the vehicle, and due to a human-like VA’s ability to facilitate such potential, the following hypothesis has been formulated:

H1: Trust in the voice assistant positively affects trust in the automated vehicle.

2.2.1. Speech Quality of Voice Assistants

VAs do not only include anthropomorphized cues in their content, but also differ in the humanness of their voice. Researchers have examined how the two different speech qualities, synthetic/tts, and pre-recorded natural human voice differ in their effects on perception in a variety of studies (Nass, Foehr, Brave, & Somoza, 2001; Lewis, Commarford, & Kotan, 2006; Stern, Mullenix, Yaroslavsky, 2006; Gorenflo & Gorenflo, 1997). The natural human voice did generally score better for emotions and was preferred when talking with a human (Nass et al., 2001). Thus, happy content, as well as sad content, were perceived as happier, or sad, when matched with the fitting natural voice, compared to the fitting synthesized one. It was not clarified whether that effect appeared due to the poor detectability of emotions, or due to the fact that computers cannot be emotional. Furthermore, did Stern et al. (2006) find that natural voices are perceived as more persuasive than synthetic speech.

Contrary to other research that states that synthetic voices will be preferred while talking with a computer (Nass & Brave, 2007), Stern et al. (2006) do not support the findings that synthetic speech will be preferred if the source is a computer. Here, their findings suggest that the participants equally prefer both types of speech quality. Further, it has been found that even

though humans react emotionally different to robots than to humans, the initial trust does not differ (Schniter, Shields, & Sznycer, 2020). Additionally, mixing natural and synthesized voices is not recommended. In the context of interactive voice response, it has been found that a consistent tts response has better effects than mixing it with a natural voice, even though natural voice might score higher alone (Lewis et al., 2006).

Considering commercial VAs, such as Mercedes's Mbox, Apple's Siri, Microsoft's Cortana, or Amazon's Alexa, one identifies that all of them make use of synthetic speech. Regarding the fact, that the likeliness towards synthetic voices strongly depends on how easy the voice is to listen to (Gorenflo & Gorenflo, 1997) and that the body of research within the topic does not include recent studies, it may be the case, that new synthetic voices have become better than those researched. Moreover, natural voice use is not always feasible and might lead to high costs, unlike synthetic voices (Lewis et al., 2006).

Nevertheless, synthetic voices do lack in specific dimensions, such as natural pauses and accents, as well as in continuity between phonemes and syllables (Nass & Lee, 2001). In addition, Stern et al. (2006) clarify that, features such as pace, verbal errors, and hesitations, affect trust in the assistant in a computer-mediated situation. Accordingly, it is concluded that natural human voice has better effects on the users' trust than the synthetic one.

2.2.2. Gender of Voice Assistants

As aforementioned, are gender traits part of the anthropomorphized attributes of VAs. Lee, Nass, and Brave (2000) identified that people automatically respond to these gender traits in tts voices, as they would to a real male or female. By comparing commercial VAs, one can identify the shared features of feminization. Except for Siri, which can be chosen to respond with a male voice, other VAs possess a female voice only, already manifested in some of their product names, like Cortana or Alexa.

According to Costa (2018), this is due to gender stereotypes of the "female secretary". Thereby, the VAs show stereotypical behaviors of submission as well as commonalities in empathizing, understanding, and accommodating the user. Additionally, female-voiced computers have been found to be more socially attractive and trustworthy (Lee et al., 2000). Further, they lack in manly attributes like assertiveness, dominance, aggression, or willingness to take stand. Søndergaard and Hansen (2018) see these stereotypes critically and motivate designers to take a feminist approach on the creation of personal digital assistants, by troubling gender stereotypes and focusing more on matters of care, trust, and interdependency. Even though organizations seek to create neutral entities, feminized VAs are reflecting society's

assumptions and gender stereotypes of women (Costa, 2018).

Differences in male and female voices involve format, frequencies, breathiness, and other that can be examined by categorical encoding through a detailed and complex auditory psychophysical process (Costa, 2018). Not only are male and female voices technically different from each other, but they also lead to other findings, depending on the context and specific attributes of the participants. For instance, speech automatically triggered a social identification process with the VA, embodying the same gender as the participants. Accordingly, female participants conform more to the female-voiced technology, than men, that conformed more to male-voiced technology (Lee, Nass, & Brave, 2000).

Regarding male voices, Reeves & Nass (1997) found that male speech outputs were better received by users than those with a female voice. Further, they identify that their participants evaluate male output as more friendly as well as better in providing information than the female output. In addition, Lee et al. (2000) underline that a male-voiced computer had a more significant influence on the users' decision making than the female-voiced version. Notwithstanding, none of these studies framed the voice as an assistant. Thus, one could argue that without a precise framing of the voice's role as an assistant, the stereotypes of the "female secretary" are not transferred by the users.

This is further stressed by Ji et al.'s (2019) findings, that stress higher scores for the female voice, that has been framed explicitly as an assistant. Specifically, it scores higher in trust, acceptance, and pleasure in the context of AVs, than the male voice. Nevertheless, they also show that fewer participants prefer a female voice after the experiment, pointing towards a difference of thought preferences and real experiences. Furthermore, they find differences between and similarities among gender for the explanation of their preferences. On the one hand, 8 out of 12 males that chose a female voice argued, that it is like the voices they already know from other VA's, such as Siri. On the other hand, 7 out of 10 females explained their choice for a male voice with its familiarity in the context of driving, because they made their first driving experiences with male relatives.

2.2.3. Interaction of a Voice Assistant's Speech Quality and Gender

Based on the presented literature, the effects of a voice assistant's speech quality and gender on both trust in the VA and trust in the AV can be predicted individually and combined. As aforementioned, does synthetic speech lack in emotion (Nass et al., 2001) as well as persuasion (Stern et al., 2006), and further comes with shortcomings in specific speech dimensions as natural pauses, accents, phonemes, and syllables (Nass & Lee, 2001), opposed to a natural

human voice. Thereupon, one can assume that a VA with a natural human voice will score higher on all three human-like trust dimensions (integrity, competence, and benevolence).

Moreover, grounded on the assumption that trust in the VA positively affects trust in the AV (*H1*), one can argue that the same condition that is causing high trust in the VA, on the one hand, causes high trust in the AV on the other hand. Hence, the VA with a natural human voice is expected to score higher on all three AV related trusting beliefs (system transparency, technical competence, situation management). As a consequence, the following hypotheses are proposed:

H2: A voice assistant's speech quality can improve trust in the voice assistant.

H2a: A voice assistant with a natural human voice will score higher on integrity, competence, and benevolence in comparison to a voice assistant with a synthetic voice.

H3: A voice assistant's speech quality can improve trust in the automated vehicle.

H3a: A voice assistant with a natural human voice will score higher on system transparency, technical competence, and situation management in comparison to a voice assistant with a synthetic voice.

Furthermore, the reviewed literature about a VA's gender stresses societal stereotypes as a major influence on the users' preferences. Consequently, the stereotype of the "female secretary", that has manifested in most commercial VAs (Costa, 2018), can be transferred onto the three human-like trusting beliefs 1) integrity, 2) competence, and 3) benevolence. A secretary is usually subordinated to the principles of his/her supervisor. Thus, users may feel that the VA's actions are easier to predict with a female voice since they take over the supervisor's role in the relationship with the assistant. In addition, typical feminine attributes of empathy and understanding could contribute to the user's perception of benevolence. Following this logic, the factor 2) competence implies a male voice preference since the setting of a car is traditionally connected to masculine attributes (Balkmar & Mellström, 2018).

Once more based on the assumed positive relationship between trust in the VA and trust in the AV (*H1*), the preceded argument about the stereotype of the "female secretary", is transferred to the AV related trusting beliefs. In other words, a female-voiced VA will score highest on a) system transparency and c) situation management. This can be justified by the fact that the provision of information is a substantial part of a secretary's profession, potentially causing users to perceive a female VA as more predictable and understandable than a male VA.

Further, connected feminine attributes of submission may lead to the perception that it

is more comfortable with a female VA to regain control in the desired situation. Lastly, similar to 2) competence, b) technical competence implies a preference of a male voice since now the context of the “masculine” car (Balkmar & Bellström, 2018) is even more determining. Subsequently, the following hypotheses have been formulated:

H4: *A voice assistant’s gender can improve trust in the voice assistant.*

H4a: *A female voice assistant will score higher on integrity and benevolence, in comparison to a male voice assistant.*

H4b: *A male voice assistant will score higher on competence, in comparison to a female voice assistant.*

H5: *A voice assistant’s gender can improve trust in the automated vehicle.*

H5a: *A female voice assistant will score higher on system transparency and situational management than a male voice assistant.*

H5b: *A male voice assistant will score higher on technical competence, in comparison to a female voice assistant.*

In addition, literature does not provide information about specific interactions between speech quality and gender. As a consequence, it is assumed that the effects on VA and AV trust add to each other. Thus, the following hypotheses have been formulated:

H6: *The interaction of a voice assistant’s speech quality and gender can improve trust in the voice assistant.*

H6a: *A female voice assistant with a natural human voice will score highest on integrity and benevolence, in comparison to all other combinations.*

H6b: *A male voice assistant with a natural human voice will score highest on competence, in comparison to all other combinations.*

H7: *The interaction of a voice assistant’s speech quality and gender can improve trust in the automated vehicle.*

H7a: *A female voice assistant with a natural human voice will score highest on system transparency and situational management, in comparison to all other combinations.*

H7b: *A male voice assistant with a natural human voice will score highest on technical competence, in comparison to all other combinations.*

2.3. Conceptual Model

Based on the presented literature, a research model has been designed, which is depicted in figure 1. The model explores the effects of a VA's speech quality and a VA's gender on trust in the VA and trust in the AV, both individually and in combination.

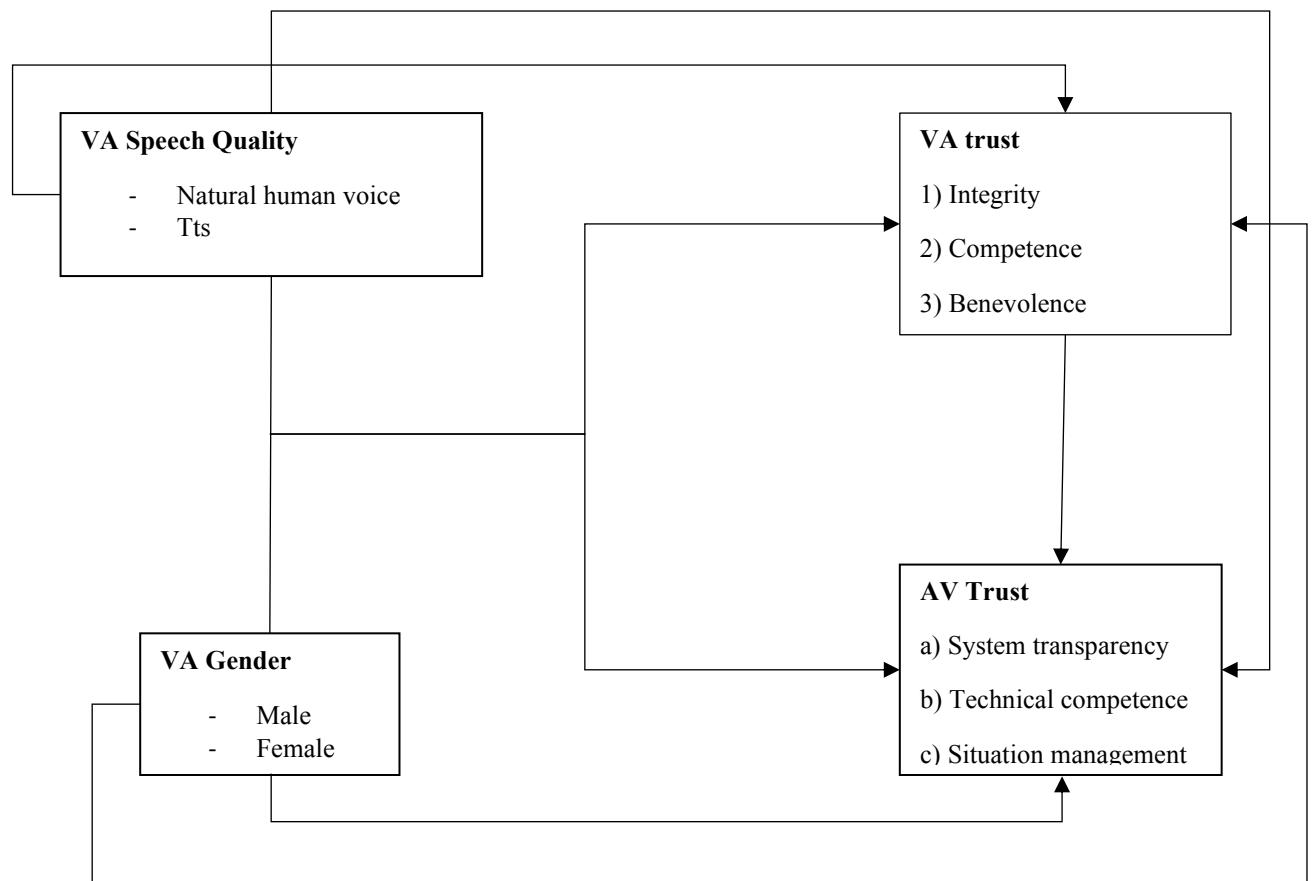


Figure 1. 2x2 Research Model

3. Method

3.1. Procedure

This study aims to identify how a VA's speech quality and gender influence the users' trust in the VA and subsequently in the AV. Accordingly, a survey in combination with a modified video has been chosen as the best-suited research method. The survey has been assessed and granted permission by the Ethics Committee of the University of Twente. To gather participants, the survey has been shared on the researcher's personal social media profiles (Instagram, Facebook), per private messenger (WhatsApp), on Surveyswap, and on the University of Twente SONA platform.

The first part of the survey includes the informed consent as well as information about associated risks and privacy implications. Then, the participants receive a brief introduction to the research topic, without going into detail about the specific aim and the independent variables researched. Afterward, the participants are equally distributed and randomly assigned to one of four videos. Here, participants watch a first-person point of view ride within an AV, while listening to the manipulated VA, modified according to one of the four conditions displayed in table 2. Subsequently, the participants are confronted with the central part of the survey, aiming to examine effects on the three VA as well as on the three AV specific trusting beliefs. Finally, the respondents answer questions regarding their demographics.

Table 1

2x2 experimental design with four conditions

		Gender	
		Female	Male
Speech quality	Natural human voice	Natural human voice & female	Natural human voice & male
	Text-to-speech	Tts & female	Tts & male

3.2. Stimulus material:

The presented video is an animated clip of 50 seconds, taken from "[Waymo 360° Experience: A Fully Self-Driving Journey](#)". To be in line with the study's goal, the 360° function of the original video has not been transferred to the clip, since the interaction might lead to a lack of focus on the VA. However, the video has explicitly been chosen because of its realistic feeling of sitting within the self-driving car from Waymo.



Figure 2. Screenshot made of “Waymo 360° Experience: A Fully Self-Driving Journey”

While seeing the car driving around, the participant listens to the VA commenting on the car’s moves as well as informing the user about a scheduled stop at the gas station, a meeting with a friend, and the estimated time that is left to arrive at the destination. The latter were included to set a clear frame on the voice’s role being an assistant. The natural human voices have been recorded by personal contacts from the researcher. It has been paid attention to select people with voices that embody stereotypical attributes of their gender as well as having profound skills in the English language. In addition, the tts Software “Polly” from Amazon served as the basis for the synthetic voices. The voice persona “Salli” has been chosen as the female voice, whereas the persona “Matthew” was chosen as the most fitting male voice. Next to the voice assistant, the clip’s soundscape incorporates the sounds of driving, the cars’ indicator, and a notification sound. The videos have been uploaded on YouTube, with the restriction that one can only find it with the specific link ([female & synthetic](#), [male & synthetic](#), [female & natural human voice](#), [male & natural human voice](#)). Finally, the VAs say the following phrases:

1. “Heading right onto Picton lane.”
2. “The car is low in gas; I have scheduled a stop at the gas station on your way to work tomorrow.”
3. “Heading right onto Houston street.”

4. “You have a calendar notification: Meeting with Emily at 7 pm.”
5. “Stop sign ahead.”
6. “Heading left onto Texas street.”
7. “We will arrive at the destination in approximately 11 minutes.”

3.3. Pre-test

To reduce misunderstandings during the data collection, a pre-test has been conducted. Therefore, three participants have been chosen to watch all the provided videos and to review the survey questions, with the task to report parts they do not understand or have questions about. Respondent one reported a spelling mistake, and respondent two did not report anything. Respondent three, on the other hand, mentioned that the phrasing of the nine questions regarding the VA is formulated too specific. Respondent three was confused since he did feel that he cannot answer specifically. He recommended to frame the questions in the same way as the AV questions, which begin with “I believe that”, clarifying that the answer shall be based on perception rather than on specific knowledge. This recommendation has subsequently been discussed with respondent one and two, whom both preferred the suggested alternative over the original version. Thus, the framing of all VA related questions has been adapted to the ones targeted at the AV trusting beliefs.

3.4. Design, Instruments, and Measurements

First, the respondents are asked to read the informed consent and to agree on it if they want to pursue with the survey. It starts with the information about who conducts the research and is followed by information about the content of the study, stating that the relationship of a user’s trust in a VA of an AV and the vehicle itself is investigated, without mentioning the specific researched variables. Then, the participant is presented with a list of preconditions that he/she can either agree or disagree on. These were derived from the Ethics Committee of the Faculty of Behavioral, Management and Social Sciences at the University of Twente.

Afterward, the respondents are asked to watch one of the four manipulated videos, before continuing with the main part of the survey. First, the trust in the presented VA is assessed by answering relevant statements on a 7-point Likert scale from (1) strongly disagree to (7) strongly agree (McKnight et al., 2002). The respondents answer on the following statements, proposed by Lankton et al. (2015), and are further adjusted to the given context and

the results of the pre-test: 1) Integrity: “I believe that the voice assistant is truthful in its dealing with me.”, “The voice assistant is honest”, “The voice assistant keeps its commitments.”; 2) Competence: “I believe that the voice assistant is competent and effective in assisting me during the ride in the automated vehicle.”, “I believe that the voice assistant is a capable and proficient assistant”; “I believe that the voice assistant performs its role of assisting me very well.” 3) Benevolence: “I believe that the voice assistant acts in my best interest.”, “I believe that the voice assistant does its best to help me if I need help.”, “I believe that the voice assistant is interested in my well-being, not just its own.”

Subsequently, the participants are asked to answer relevant questions to assess their trusting beliefs in the AV. Here, the following questions have been proposed by Choi and Lee (2015): a) System transparency “I believe that the autonomous vehicle acts consistently and its behavior can be forecast.”, “I believe that I can form a mental model and predict the future behavior of the autonomous vehicle.”, “I believe that I can predict that the autonomous vehicle will act in a particular way.”; b) Technical competence: “I believe that the autonomous vehicle is free of error.”, “I believe that I can depend and rely on the autonomous vehicle.”, “I believe that the autonomous vehicle will consistently perform under a variety of circumstances.”; c) Situation management: “I believe that the autonomous vehicle provides alternative solutions.”, “I believe that I can control the behavior of the autonomous vehicle.”, “I believe that the autonomous vehicle will provide adequate, effective, and responsive help.”

Lastly, the participants are asked to answer questions regarding their demographics, being age, gender, experience with VAs, experience with AVs, and country of origin. The participants are particularly asked about their experiences with VAs and AVs (partial and fully automated), because the answers may clarify whether specific conditions in past usage lead to different outcomes regarding trust.

3.5. Respondents

The researched sample consisted of 100 participants (N=100), equally distributed to one of the four cases. The respondents are between 18 and 60 years old, whereby 90% of them belong to the group of 18 to 25 years. Regarding the single conditions, the percentage of participants in that age group does not differ by more than 2%. Moreover, 51% of the participants are female, 46% are male, 1% is diverse, and 2% preferred not to answer. Within the single conditions, there are no deviations in gender that are greater than 6%. Since the study’s goal includes supporting the EU in facilitating the transformation of traffic, all respondents live in the EU, whereby 59% are German, 29% are Dutch, and 12% come from other European countries. In

terms of the conditions, deviations reach up to 11%.

Further, 75% of the sample have experiences with VAs, such as Apple Siri, Google Assistant, Amazon Alexa, Microsoft Cortana, and Mercedes Mbux; 25% have stated to have not have used a VA. Additionally, 21% have had experiences with automated cars, whereby 11% drove within a partially automated car, 15% sat next to someone driving a partially automated car, 1% declared to have had experiences with a fully automated car, and 3% have had other experiences. Moreover, accepting the informed consent included declaring that one is physically able to listen to the VA and to watch the video. In addition, sufficient English skills were a requirement to take the survey to understand the VA and the asked questions.

3.6. Validity

To test the validity of the study, a factor analysis has been conducted. In total, the analysis included 18 items and six factors. Based on the analysis, one can tell if the items measure the factors, and subsequently, the constructs appropriately. Since the KMO value is .79 and therefore scores above the recommended minimum value of .60, the chosen data is suitable for a factor analysis.

Moreover, the table “Total Variance Explained” displays the Eigenvalues of the identified factors. On the one hand, it shows Eigenvalues above 1 for the first five factors, which are therefore selected for further analysis. On the other hand, factor 6 has an Eigenvalue of 0.97 and therefore disqualifies for further analysis.

The “Rotated Component Matrix” displays the exact items that relate to the identified factors with values above the minimum of 0.65. Following the expectations, the retrieved triplets of questions do load together on one factor, with one deviation for factor 6. They are further displayed in table 3, ordered by the amount of variance that is explained by the factor. Thus, the questions related to competence (Q4, Q5, and Q6) load on factor 1; the questions related to system transparency (Q10, Q11, and Q12) load on factor 2; the questions related to benevolence (Q7, Q8, and Q9) load on factor 3; the questions related to integrity (Q1, Q2, and Q3) load on factor 4; the questions related to technical competence (Q13, Q14, Q15) load on factor 5; and lastly, the questions Q16 and Q18, both explaining situation management, load on factor 6, whereby Q17 does not sufficiently load on any factor. Subsequently, the factor loadings support the rejection of factor 6.

3.7. Reliability

To find out more about the data's internal consistency, a reliability analysis has been added. In particular, the items have been tested on their relationship as single constructs. As displayed in table 3, does each construct have a Cronbach's Alpha of above 0.65, which suggests sufficient internal consistency of the items and thus sufficient reliability of the data.

Table 3

Validity Factor Analysis

Item	Factor					
	1	2	3	4	5	6
Factor 1: Integrity ($\alpha = .78$)						
Q1 I believe that the voice assistant is truthful in its dealing with me.				.71		
Q2 The voice assistant is honest.				.77		
Q3 The voice assistant keeps its commitments.				.76		
Factor 2: Competence ($\alpha = .86$)						
Q4 I believe that the voice assistant is competent and effective in assisting me during the ride in the automated vehicle.	.83					
Q5 I believe that the voice assistant is a capable and proficient assistant.	.81					
Q6 I believe that the voice assistant performs its role of assisting me very well.	.74					
Factor 3: Benevolence ($\alpha = .82$)						
Q7 I believe that the voice assistant acts in my best interest			.72			
Q8 I believe that the voice assistant does its best to help me if I need help.			.80			
Q9 I believe that the voice assistant is interested in my well-being, not just its own.			.77			
Factor 4: System transparency ($\alpha = .81$)						
Q10 I believe that the autonomous vehicle acts consistently, and its behavior can be forecast.		.80				
Q11 I believe that I can form a mental model and predict the future behavior of the autonomous vehicle.		.72				
Q12 I believe that I can predict that the autonomous vehicle will act in a particular way.		.77				
Factor 5: Technical competence ($\alpha = .69$)						
Q13 I believe that the autonomous vehicle is free of error.				.77		
Q14 I believe that I can depend and rely on the autonomous vehicle.				.76		
Q15 I believe that the autonomous vehicle will consistently perform under a variety of circumstances.				.70		
Factor 6: Situation management ($\alpha = .65$)						
Q16 I believe that the autonomous vehicle provides alternative solutions.					.85	
Q17 I believe that I can control the behavior of the autonomous vehicle.						
Q18 I believe that the autonomous vehicle will provide adequate, effective, and responsive help.					.66	

4. Results

In this study, the effects of a VA's speech quality and gender have been tested on trust in the VA (integrity, benevolence, and competence) and trust in the AV (system transparency, technical competence, situation management). First, by matter of correlation testing and univariate analysis of variance, it has been tested if trust in the VA does indeed positively affects trust in the AV. Afterward, a multivariate analysis of variance has been performed, that was preceded by a test of MANOVA relevant assumptions. Furthermore, the results of the main effects are displayed and further related to the proposed hypotheses.

4.1. The Effects of Trust in the Voice Assistant on Trust in the Automated Vehicle

In order to investigate the relationship between the independent variables, a Pearson Correlation was conducted. Therefore, the constructs have been calculated into two new variables, VA trust, and AV trust. However, the factor situation management was not included in the analysis due to its insufficient validity. According to the Pearson Correlation, are VA trust and AV trust moderately positively correlated, with 0.53 ($p = 0.00$).

Based on the significant correlation, the variables have been tested on causation. Hence, VA trust as an independent variable and AV trust as a dependent variable. The univariate analysis of variance indicates a significant causation with $p = 0.00$ ($F = 2.46$). Finally, one can argue that the significant positive correlation, as well as the significant causation, support H1, which incorporates a positive effect of VA trust on AV trust.

4.2. Multivariate Analysis of Variance

To investigate the formulated hypotheses, a multivariate analysis of variance has been conducted. A MANOVA has been particularly chosen for this study because it supports having more than one dependent variable.

Additionally, the data has been tested on outliers, multivariate normality, and multicollinearity. First, one outlier has been identified by the matter of the Mahalanobis test statistic, which has therefore been removed and leaves 24 participants for the condition tts and female. Moreover, all the constructs measuring trust in the VA (integrity, competence, benevolence) and one construct measuring trust in the AV (system transparency) fail the criteria of a normal distribution, with $p = 0.00$ for the Shapiro Wilk test. However, technical competence meets the criteria of a normal distribution with $p = 0.07$. Lastly, it was proved that no multicollinearity exists between the items since no score violated the maximum of 0.8 for the Pearson Correlation.

In addition, a Pillai's Trace has been performed, particularly chosen based on the

violated assumptions. Table 5 displays the Pillai's Trace, p values, F values, and scores for the squared partial eta. Looking at the outcome, there was no main effect found for speech quality, with Pillai's Trace = 0.09, $F(5, 90) = 1.50$, and $p = 0.19$. In addition, no main effect could be identified for gender, with Pillai's Trace = 0.09, $F(5, 90) = 1.46$, and $p = 0.20$. Lastly, for the combination of both independent variables, no main effect could be identified either, with Pillai's Trace = 0.02, $F(5, 91) = 0.31$, and $p = 0.93$. Finally, these results indicate that there are no effects of speech quality, gender, nor of their combination on the dependent variables.

Table 4

Multivariate Analysis; Descriptive statistics of the independent variables

Effect		Value	F	Sig.	Partial Eta Squared	Observed Power
Speech Quality	Pillai's Trace	.09	1.82	.12	.09	.60
Gender	Pillai's Trace	.03	.60	.70	.32	.21
Speech Quality * Gender	Pillai's Trace	.02	.33	.90	.02	.13

4.3. Main Effects of Gender and Speech Quality

Within the following section, the main effects of the independent variables (speech quality, gender) on the dependent variables trust in the VA (integrity, competence, benevolence) and trust in the AV (system transparency, technical competence) are being displayed and elaborated on in detail.

4.3.1. Effects on Trust in the Voice Assistant

Firstly, the effects of speech quality and gender on trust in the VA are shown, which are further depicted in Table 6. The results show that there is no significant effect of speech quality on integrity, with $p = 0.49$ and $F = 0.49$, no effect on competence, with $p = .83$ and $F(1, 95) = 0.05$, and lastly no effect on benevolence, with $p = 0.26$ and $F(1, 95) = 1.31$. Subsequently, the hypotheses H2 and H2a are not being supported.

In addition, the effects of gender on trust in the VA have been analyzed. The results indicate that there is no effect on integrity, with $p = 0.24$ and $F(1, 95) = 1.42$, no effect on competence, with $p = .84$ and $F(1, 95) = 0.04$, and finally no effect on benevolence, with $p = .68$ and $F(1, 95) = 0.17$. Thus, the hypotheses H4, H4a, and H4b are not supported.

Furthermore, speech quality and gender, in combination, do not show significance

either, meaning that no main effect was found. In detail, did the combination show no effect for integrity, with $p = 0.99$ and $F(1, 95) = 0.00$, no effect on competence, with $p = 0.94$ and $F(1, 95) = 0.01$, and lastly no effect on benevolence, with $p = 1.31$ and $F(1, 95) = 1.31$. Hence, the hypotheses H6 and H6a are not supported.

4.3.2. Effects on Trust in the Automated Vehicle

Secondly, the effects of speech quality and gender on trust in the AV are shown and as well depicted in table 6. To begin with, there was no effect found for speech quality on system transparency, with $p = 0.07$ and $F(1, 95) = 3.26$, and no effect on technical competence, with $p = .74$ and $F(1, 95) = 0.11$. Accordingly, the hypotheses H3 and H3a are not supported.

Moreover, the results point towards an absence of an effect of gender on trust in the AV. Hence, there is no effect on system transparency, with $p = 0.88$ and $F(1, 95) = 0.02$, nor on technical competence, with $p = 0.70$ and $F(1, 95) = 0.16$. Subsequently, the hypotheses H5, H5a, and H5b are not being supported.

Additionally, the findings show no effect for the combination of speech quality and gender on trust in the AV. Meaning, no effect was found for system transparency, with $p = 0.83$ and $F(1, 95) = 0.05$, and neither for technical competence, with $p = 0.86$ and $F(1, 95) = 0.03$. Therefore, the hypotheses H7, H7a, and H7b are not supported.

Table 5

Multivariate Test; Descriptive statistics of the independent variables

Source	Dependent Variable	df	F	Sig.
Speech Quality	Integrity	1	.49	.49
	Competence	1	.05	.83
	Benevolence	1	1.31	.26
	System transparency	1	3.26	.07
	Technical competence	1	.11	.74
Gender	Integrity	1	1.42	.24
	Competence	1	.04	.84
	Benevolence	1	.17	.68
	System transparency	1	.02	.88
	Technical competence	1	.16	.70
Speech Quality *	Integrity	1	.00	.99
Gender	Competence	1	.01	.94
	Benevolence	1	1.31	.26
	System transparency	1	.05	.83
	Technical competence	1	.03	.86

Table 6
Overview of hypotheses

Hypotheses	
H1: Trust in the voice assistant positively affects trust in the automated vehicle.	Supported
H2: A voice assistant's speech quality can improve trust in the voice assistant.	Not supported
H2a: A voice assistant with a natural human voice will score higher on integrity, competence, and benevolence in comparison to a voice assistant with a synthetic voice.	Not supported
H3: A voice assistant's speech quality can improve trust in the automated vehicle.	Not supported
H3a: A voice assistant with a natural human voice will score higher on system transparency, technical competence, and situation management in comparison to a voice assistant with a synthetic voice.	Not supported
H4: A voice assistant's gender can improve trust in the voice assistant.	Not supported
H4a: A female voice assistant will score higher on integrity and benevolence, in comparison to a male voice assistant.	Not supported
H4b: A male voice assistant will score higher on competence, in comparison to a female voice assistant.	Not supported
H5: A voice assistant's gender can improve trust in the automated vehicle.	Not supported
H5a: A female voice assistant will score higher on system transparency and situational management than a male voice assistant.	Not supported
H5b: A male voice assistant will score higher on technical competence, in comparison to a female voice assistant.	Not supported
H6: The interaction of a voice assistant's speech quality and gender can improve trust in the voice assistant.	Not supported
H6a: A female voice assistant with a natural human voice will score highest on integrity and benevolence, in comparison to all other combinations.	Not supported
H6b: A male voice assistant with a natural human voice will score highest on competence, in comparison to all other combinations.	Not supported
H7: The interaction of a voice assistant's speech quality and gender can improve trust in the automated vehicle.	Not supported
H7a: A female voice assistant with a natural human voice will score highest on system transparency and situational management, in comparison to all other combinations.	Not supported
H7b: A male voice assistant with a natural human voice will score highest on technical competence, in comparison to all other combinations.	Not supported

5. Discussion and Conclusion

This study aimed to explore the design of trustful VAs within AVs to create trustful user interactions between users and the driving technology. The gained trust in the vehicle could potentially improve its' public adoption. Based on the presented literature, speech quality and gender have been identified as relevant variables influencing the user's trust perception of VAs. Moreover, according to the potential of trust in VAs to positively affect trust in AVs, it was assumed that speech quality and gender add to it as well. Based on these objectives, four research questions have been formulated (1): *RQ1: How is the user's trust in the fully automated vehicle affected by the user's trust in the voice assistant? RQ2: How is the user's trust in the fully automated vehicle affected by the voice assistant's gender? RQ3: How is the user's trust in the fully automated vehicle affected by the voice assistant's speech quality? RQ4: How is the user's trust in the fully automated vehicle affected by the interaction of the voice assistant's gender and speech quality.* To answer these questions, four main and ten sub hypotheses have been proposed. Thereupon, speech quality and gender have been further explored by an experimental 2x2 study design, on their effects on the constructs measuring trust in the VA (integrity, competence, benevolence) and trust in the AV (system transparency, technical competence, situation management).

5.1. Discussion of Findings

According to the results does trust in the VA indeed positively affect trust in the AV. The positive relationship goes with Niu et al.'s (2018) findings, that emphasized the benefits of anthropomorphizing symbolic information within AVs. More importantly, it builds the foundation for studies that try to expand that knowledge to maximize the user's trust perception by researching trustful VAs. On the contrary, there were no identifiable main effects for speech quality and gender on the dependent factors. The absence of significant findings does challenge their theoretical foundations and further points towards possible improvements in the research design.

The confirmation of an existing positive relationship of trust in the VA and trust in the AV leaves open questions and opportunities on the one hand, and possible concerns about its use on the other hand. As aforementioned, the research conducted was conducted with the motivation to support the EU with their transformation of traffic and industry, as well as with their specific goal of reducing road fatalities. By showing the potential power of improving trust in AVs with the creation of trustful VAs, it is recommended that the EU enforces more

research into that specific domain. Moreover, car manufacturers would benefit financially from the expanded research since trustful designed vehicles will facilitate an easier and faster adoption period to their products. Consequently, it is strengthening the EU's economy and as well creating new business opportunities for organizations.

However, the finding does not give any more information on how much VAs can be anthropomorphized before they are too human-like, which could result in the uncanny valley (Mori et al., 2012). Since modern VAs incorporate its voices that already reach a high point of humanness, it may be the case that voice itself is not enough to trigger such strong reactions. Moreover, the fusion of VAs and AVs leaves room to anthropomorphize the assistant apart from its voice. For instance, could the ambient light of modern cars be used to underline certain statements. In the perspective of the possibility to issue warnings through the VA, colors with alarming traits could be combined accordingly. Thus, future work may expand the perspective on research into anthropomorphized VAs, by including the technical possibilities of its host technology. Thereupon, it can be tested whether trust in the VA can be strengthened by more than just its voice, and further they may identify a critical point of humanness that results in negative reactions from its users.

In addition, the potential of trustful VAs could be used to tackle today's traffic problems by being implemented within cars in the first three stages of automation of the SAE International (2020) scale. Even though humans are still required to interfere, VAs may add to the trust in specific automation features, such as adaptive cruise control and lane centering. Accordingly, literature must show that results from study's with fully automated cars can be transferred to trust in specific automation features of lower automated cars.

Furthermore, a trustful VA may as well be used to enforce positive behavior from the users. As the human-like trusting belief benevolence (Lankton et al., 2015) implies that the trustor expects the trustee to do good to him, suggestions from a trustful VA may be accepted. Thereupon, if future research follows Coeckelbergh's (2016) recommendation of clarifying *how* to act responsibly within automated traffic, a trustful VA may help to facilitate such. Moreover, referring to Coeckelbergh's (2016) standpoint, that users of AVs may not see other traffic entities as real "others", a trustful VA could potentially be used as a communication mean to create awareness of their "otherness". As a result, users may again act more responsible opposed to a situation where others are seen as "machines" only.

However, reviewed from a critical standpoint, one can argue against the research into trustful VAs in the given context. Since VAs are a possibility to create trust in technology other than the assistant, they may be used for technology that does not provide a good reason to trust

it on its own. For instance, a car manufacturer could incorporate a very trustworthy VA in a faulty AV. As a result, users could be endangered through trust in unsafe technology. Following the findings of Schniter et al. (2020), that have shown that people react emotionally different to robots than to humans, the manipulative power could even be strengthened through human-like attributes. To stop organizations from such misuse, legislators must assess the application of VAs legally and ethically. Accordingly, they should publish concrete guidelines to support manufacturers in their decisions.

Furthermore, the absence of significant findings for the main effects of speech quality and gender challenges their theoretical foundation and the concluded hypotheses. To begin with, the results for speech quality could indicate that the aforementioned studies that use older technologies, are indeed not valid as a foundation for modern VAs. Hence, on the one hand, it could imply that modern software for synthetic voices does not lack in emotion (Nass et al., 2001), persuasion (Stern et al., 2006) and in natural pauses, accents, phonemes, and syllables (Nass & Lee, 2001). On the other hand, it could mean that some of the findings are simply not relevant in the given context. In particular, it displays a challenge between the pace of research and technology development. As technology improves rapidly, older scientific findings are always only valid for the technology state at a specific moment in time. Future work may specifically tackle research assumptions from the past, as those presented, and test them frequently as soon technology for synthetic voices enters a new stage of development.

Additionally, the not significant results for gender could mean that the stereotype of the “female secretary” (Costa, 2018) was not applicable in the given context. Potentially it could be the case that the context of an AV was perceived as more masculine as other devices with inbuilt VAs, such as smartphones, whereby stereotypes for assistants and cars canceled each other. Moreover, this could mean that one cannot simply view the VA and the AV as separate technological entities, but rather as a merged device, which makes it difficult to apply stereotypes that have been originated for a single technology. Alternatively, it could imply that only specific attributes are seen from a merged-technology perspective, and in turn, others are seen separately. Future work should, therefore, explore whether users apply gender stereotypes of VAs and AVs together or separately.

Lastly, no effect was found for the interaction of speech quality and gender. As there have been no identifiable effects for them individually, one may argue that the limiting factors for such, as well caused non-significant findings for their interaction. As it must be assumed that the individual effects, if present, add to each other, future research must focus on these separately, before combining them again.

5.2. Limitations

The present study's limitations will be discussed and used to give further recommendations. First, the observed power indicated that the present study did not cover the data of enough participants to produce significant results. Therefore, future studies should consider a larger number of participants, preceded by a priori power analysis, giving information on how large that amount shall be.

Furthermore, the body of literature that builds the foundation of the proposed hypotheses can be seen as a limiting factor. Specifically, the literature on speech quality lacks in recent work. It is mainly consisting of studies that use out-of-date technologies as stimulus, carrying the risk that results cannot be transferred to the present stage of technological development. Even though the present results may not be statistically significant, they can be seen as an indicator pointing towards this assumption. Therefore, future research should tackle this problem by investigating differences in speech quality with current VA technologies, to build a better foundation for studies like the present one.

Moreover, gender research does not provide a clear direction either, despite that experiences and stereotypes play a crucial role in the users' trust and are already manifested in several VAs. Nonetheless, the study did not include extra verification of the existence of these stereotypes among the researched population. Thus, future researchers might want to install such a verification instance within their study to double-check their assumptions. To shed more light on how and if stereotypes are influencing the user's trust perception in VAs and AVs, more context relevant research is needed, that explicitly investigates gender in combination with the technologies.

Next to the limitations of the study's foundation, the actual design incorporates further limitations that can be improved in future research. Regarding the pre-test, future studies may optimize it by consulting more participants. Additionally, a pre-test may have a greater impact if conducted with people that are socially unrelated to the researcher. As friends and family of the researcher have been consulted to do be part of the pre-test, they might not have communicated flaws deliberately.

Furthermore, the exploitation of the users to the modified VA was relatively short. Explicitly, this has been the case, because the underlying video clip was chosen based on a realistic feeling of sitting within an AV, and not based on the actual length in time. This trade-off may have contributed to the insignificant results and can be further explored in the future. Meaning, either future studies make use of longer video material or alternatively let the participants listen to the VA only, while the context is communicated by other means but not

shown. Hence, respondents are exploited to the manipulated VA much longer.

In addition, sharing the survey on the researcher's personal social media, means that a fair part of the researched population stems from the researcher's social environment. Next to the risk of establishing results derived from a homogeneous group that potentially differs from other parts of the population, some may have recognized the natural human voices, whose speakers are as well part of the social environment of the researcher. Thus, this circumstance may have influenced the results.

Furthermore, there are alternative ways to create a stimulus, such as virtual reality and driving simulators, that could approximately create more realistic settings for the study of AVs. Thus, combined with a more qualitative approach, they may lead to better conditions. Especially at the point in time, when AVs become easily accessible, they can be used in research more often. Hence, studies will not have to suffer from the quality of simulations anymore, and subsequently improve the quality of results.

Moreover, the demographic section of this study lacks a question about the educational background of the participants. Since the segmentation of users by Nielsen & and Haustein (2018) explicitly mentioned a higher educational background as a commonality between enthusiasts, the missing question might suppress useful information in segmenting the participants. Drawn from this, future research should include questions about the educational background to allow better user segmentation.

5.3. Conclusion

The present study was designed to contribute to the challenge of AV adoption by exploring the design of trustful VAs in AVs. The design variables speech quality and gender have been researched individually and in interaction. In that manner, their effects on the constructs measuring trust in the VA (integrity, competence, benevolence) and trust in the AV (system transparency, technical competence, situation management) have been tested. The most noticeable finding was that that the users' trust in the VA of an AV is indeed causing higher trust in the AV. Consequently, researchers and organizations should follow this paper's recommendations by further researching the potential of trustful VAs, to improve the adoption of AVs, and to finally reduce human-caused road fatalities. However, no effects for speech quality and gender were identified. To make progress, future work should test findings from older studies with new technology, to update the scientific foundation of speech quality and gender of VAs. Furthermore, it is recommended to explore whether VAs and AVs are perceived as separate technological entities or as a merged device.

6. Literature

- Alalwan, A. A., Baabdullah, A. M., Rana, N. P., Tamilmani, K., & Dwivedi, Y. K. (2018). Examining adoption of mobile internet in Saudi Arabia: Extending TAM with perceived enjoyment, innovativeness and trust. *Technology in Society*, 55, 100–110. doi: 10.1016/j.techsoc.2018.06.007
- Balkmar, D., & Mellström, U. (2018). Masculinity and Autonomous Vehicles. *Transfers*, 8(1), 44-63. doi:10.3167/trans.2018.080105
- Beiker, S., & Calo, R. (2010). Legal Aspects of Autonomous Driving. *SSRN Electronic Journal*. doi:10.2139/ssrn.2767899
- Boonrod, A., & Ketavan, P. (2018). The Future Is Voice. *Research World*, 2018(71), 40–45. doi: 10.1002/rwm3.20688
- Buckley, L., Kaye, S.-A., & Pradhan, A. K. (2018). Psychosocial factors associated with intended use of automated vehicles: A simulated driving study. *Accident Analysis & Prevention*, 115, 202–208. doi: 10.1016/j.aap.2018.03.021
- Bucsuházy, K., Matuchová, E., Zůvala, R., Moravcová, P., Kostíková, M., & Mikulec, R. (2020). Human factors contributing to the road traffic accident occurrence. *Transportation Research Procedia*, 45, 555–561. doi: 10.1016/j.trpro.2020.03.057
- Califf, C. B., Brooks, S., & Longstreet, P. (2020). Human-like and system-like trust in the sharing economy: The role of context and humanness. *Technological Forecasting and Social Change*, 154, 119968. doi:10.1016/j.techfore.2020.119968
- Choi, J. K., & Ji, Y. G. (2015). Investigating the Importance of Trust on Adopting an Autonomous Vehicle. *International Journal of Human-Computer Interaction*, 31(10), 692–702. doi: 10.1080/10447318.2015.1070549
- Chi, C.-F., Dewi, R. S., & Huang, M.-H. (2017). Psychophysical evaluation of auditory signals in passenger vehicles. *Applied Ergonomics*, 59, 153–164. doi: 10.1016/j.apergo.2016.08.019
- Coeckelbergh, M. (2016). Responsibility and the Moral Phenomenology of Using Self-Driving Cars. *Applied Artificial Intelligence*, 30(8), 748-757. doi:10.1080/08839514.2016.1229759
- Costa, P. C. (2018). Conversing with Personal Digital Assistants: On Gender and Artificial Intelligence. *Journal of Science and Technology of the Arts*, 10(3), 2. doi:10.7559/citarj.v10i3.563
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319. doi: 10.2307/249008

- Garcia, M. P., & Lopez, S. S. (2019). Exploring the Uncanny Valley Theory in the Constructs of a Virtual Assistant Personality. *Advances in Intelligent Systems and Computing Intelligent Systems and Applications*, 1017–1033. doi: 10.1007/978-3-030-29516-5_76
- Ghazizadeh, M., Lee, J. D., & Boyle, L. N. (2011). Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work*, 14(1), 39–49. doi: 10.1007/s10111-011-0194-3
- Gold, C., Körber, M., Hohenberger, C., Lechner, D., & Bengler, K. (2015). Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle. *Procedia Manufacturing*, 3, 3025–3032. doi: 10.1016/j.promfg.2015.07.847
- Golden, K. E., & Fleischmann, K. R. (2018). Digital assistants and human values. *Proceedings of the Association for Information Science and Technology*, 55(1), 813–814. doi: 10.1002/prai.2018.14505501128
- Gorenflo, D., & Gorenflo, C. (1997). Effects of synthetic speech, gender, and perceived similarity on attitudes toward the augmented communicator. *Augmentative and Alternative Communication*, 13(2), 87–91. doi:10.1080/07434619712331277878
- Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior*, 90, 343–350. doi: 10.1016/j.chb.2018.08.009
- Hoff, K. A., & Bashir, M. (2014). Trust in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 57(3), 407–434. doi: 10.1177/0018720814547570
- Huisman, D., & Huisman, J. (2018). Voice and the digital assistant. *Research World*, 2018(73), 38–41. doi: 10.1002/rwm3.20738
- Ji, W., Liu, R., & Lee, S. (2019). Do Drivers Prefer Female Voice for Guidance? An Interaction Design About Information Type and Speaker Gender for Autonomous Driving Car. *HCI in Mobility, Transport, and Automotive Systems Lecture Notes in Computer Science*, 208–224. doi: 10.1007/978-3-030-22666-4_15
- Kaur, K., & Rampersad, G. (2018). Trust in driverless cars: Investigating key factors influencing the adoption of driverless cars. *Journal of Engineering and Technology Management*, 48, 87–96. doi: 10.1016/j.jengtecman.2018.04.006
- König, M., & Neumayr, L. (2017). Users' resistance towards radical innovations: The case of the self-driving car. *Transportation Research Part F: Traffic Psychology and Behaviour*, 44, 42–52. doi: 10.1016/j.trf.2016.10.013
- Kyriakidis, M., Happee, R., & Winter, J. D. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. *Transportation Research Part F: Traffic Psychology and Behaviour*, 32, 127–140. doi: 10.1016/j.trf.2015.04.014

- Lankton, N., 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
- Lee, J. D., & Moray, N. (1994). Trust, self-confidence, and operators adaptation to automation. *International Journal of Human-Computer Studies*, 40(1), 153–184. doi: 10.1006/ijhc.1994.1007
- Lee, E. J., Nass, C., & Brave, S. (2000). Can computer-generated speech have gender? *CHI '00 Extended Abstracts on Human Factors in Computing Systems - CHI '00*. doi:10.1145/633292.633461
- Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 50–80. doi: 10.1518/hfes.46.1.50.30392
- Lewis, J. R., Commarford, P. M., & Kotan, C. (2006). Web-based comparison of two styles of auditory presentation: All TTS versus rapidly mixed TTS and recordings. *PsycEXTRA Dataset*. doi:10.1037/e577592012-023
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, 20(3), 709. doi:10.2307/258792
- McKnight, D. H., 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
- Mori, M., Macdorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2), 98–100. doi: 10.1109/mra.2012.2192811
- Muir, B. M., & Moray, N. (1996). Trust in automation. Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics*, 39(3), 429–460. doi: 10.1080/00140139608964474
- Nass, C., & Brave, S. (2007). *Wired for Speech: How voice activates and advances the human-computer relationship*. Cambridge: MIT Press.
- Nass, C., Foehr, U., Brave, S., & Somoza, M. (2001). *The Effects of Emotion of Voice in Synthesized and Recorded Speech*. Retrieved March 23, 2020, from <https://www.aaai.org/Papers/Symposia/Fall/2001/FS-01-02/FS01-02-019.pdf>.
- Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, 7(3), 171-181. doi:10.1037/1076-898x.7.3.171

- Nielsen, T. A. S., & Haustein, S. (2018). On sceptics and enthusiasts: What are the expectations towards self-driving cars? *Transport Policy*, 66, 49–55. doi: 10.1016/j.tranpol.2018.03.004
- Niu, D., Terken, J., & Eggen, B. (2018). Anthropomorphizing information to enhance trust in autonomous vehicles. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 28(6), 352–359. doi: 10.1002/hfm.20745
- Pavlou, P. A. (2003). Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. *International Journal of Electronic Commerce*, 7(3), 101–134. doi: 10.1080/10864415.2003.11044275
- Raue, M., Dambrosio, L. A., Ward, C., Lee, C., Jacquillat, C., & Coughlin, J. F. (2019). The Influence of Feelings While Driving Regular Cars on the Perception and Acceptance of Self-Driving Cars. *Risk Analysis*, 39(2), 358–374. doi: 10.1111/risa.13267
- Reeves, B., & Nass, C. (1997). The media equation: How people treat computers, television, & new media like real people & places. *Computers & Mathematics with Applications*, 33(5), 128. doi:10.1016/s0898-1221(97)82929-x
- Road deaths in the European Union – latest data. (n.d.). Retrieved March 24, 2020, from <https://etsc.eu/euroadsafetydata/>
- SAE J3016 automated-driving graphic. (2020, May 15). Retrieved June 24, 2020, from <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic>
- Schniter, E., Shields, T., & Sznycer, D. (2020). Trust in humans and robots: Economically similar but emotionally different. *Journal of Economic Psychology*, 78, 102253. doi:10.1016/j.joep.2020.102253
- Self-driving cars in the EU: From science fiction to reality: News: European Parliament. (2019, January 14). Retrieved June 24, 2020, from <https://www.europarl.europa.eu/news/en/headlines/economy/20190110STO23102/self-driving-cars-in-the-eu-from-science-fiction-to-reality>
- Sheridan, T. B. (1975). Considerations in modeling the human supervisory controller. *International Federation of Automatic Control, Triennial World Congress, 6th*, 40.
- Sheridan, T. B., & Hennessy, R. T. (1984). *Research and modeling of supervisory control behavior. Report of a workshop*. Washington DC: National Research Council, Committee on Human Factors.
- Søndergaard, M. L., & Hansen, L. K. (2018). Intimate Futures. *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*. doi:10.1145/3196709.3196766

- Stern, S. E., Mullennix, J. W., & Yaroslavsky, I. (2006). Persuasion and social perception of human vs. synthetic voice across person as source and computer as source conditions. *International Journal of Human-Computer Studies*, 64(1), 43-52. doi:10.1016/j.ijhcs.2005.07.002
- Torre, I., Goslin, J., & White, L. (2020). If your device could smile: People trust happy-sounding artificial agents more. *Computers in Human Behavior*, 105, 106215. doi:10.1016/j.chb.2019.106215
- Wong, P. N. Y., Brumby, D. P., Babu, H. V. R., & Kobayashi, K. (2019). Voices in Self-Driving Cars Should be Assertive to More Quickly Grab a Distracted Drivers Attention. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI 19*. doi: 10.1145/3342197.3344535
- Wu, I.-L., & Chen, J.-L. (2005). An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study. *International Journal of Human-Computer Studies*, 62(6), 784–808. doi: 10.1016/j.ijhcs.2005.03.003
- Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies*, 112, 220–233. doi: 10.1016/j.trc.2020.01.027

7. APPENDIX

Appendix A – Final Questionnaire

Introduction and Informed Consent

Dear participant, thank you for taking part in this survey! The study is conducted by Julian Schwierien from the Faculty of Behavioral, Management, and Social Sciences at the University of Twente. Please read the following information carefully.

The study is investigating the relationship of a user's trust in a voice assistant of a fully automated vehicle and the vehicle itself. Since you will be watching a first-person point of view video of sitting within a fully automated vehicle, do not proceed with the survey if you suffer from anxiety towards driving or specifically automated driving. Completing the survey will take about 8 minutes.

In the case of questions, suggestions, or ambiguities, feel free to contact the researcher by j.schwierien@student.utwente.nl

I understand and consent that:

- I am at least 18 years old
- I can withdraw from the study at any given time.
- My participation is voluntary.
- I understand the risks in the case of anxiety towards driving and automated driving.
- I have been informed that the research has been reviewed and approved by the BMS Ethics Committee.
- My data are handled confidentially and anonymously.
- I have been informed about my right to request access, rectification, and erasure of my data.

If you have questions about your rights as a research participant or wish to obtain information, ask questions, or discuss any concerns about this study with someone other than the researcher, please contact the Secretary of the Ethics Committee of the Faculty of Behavioral, Management and Social Sciences at the University of Twente by ethicscommittee-bms@utwente.nl

- ☐ Yes, I do consent.
- ☐ No, I do not consent.

Video

Now you will watch a short video of 50 seconds. Make sure to put up your volume, to be able to hear the voice assistant. Watch and listen carefully!

Equally distributed to one of the four conditions (Displayed in Appendix B)

Questions measuring trust in the VA

With the following statements we aim to examine your perception of the presented **voice assistant**. Please answer as precise as possible.

1. "I believe that the voice assistant is truthful in its dealing with me."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

2. "The voice assistant is honest"

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

3. "The voice assistant keeps its commitments."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

4. "I believe that the voice assistant is competent and effective in assisting me during the ride in the automated vehicle."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree

- Disagree
- Strongly disagree

5. “I believe that the voice assistant is a capable and proficient assistant.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

6. “I believe that the voice assistant performs its role of assisting me very well.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

7. “I believe that the voice assistant acts in my best interest.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

8. “I believe that the voice assistant does its best to help me if I need help.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

9. "I believe that the voice assistant is interested in my well-being, not just its own."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

Questions measuring trust in the AV

With the following statements we aim to examine your perception of the presented **automated vehicle**. Please answer as precise as possible.

10. "I believe that the autonomous vehicle acts consistently and its behavior can be forecast."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

11. "I believe that I can form a mental model and predict the future behavior of the autonomous vehicle"

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree
- ☐ Disagree
- ☐ Strongly disagree

12. "I believe that I can predict that the autonomous vehicle will act in a particular way."

- ☐ Strongly agree
- ☐ Agree
- ☐ Somewhat agree
- ☐ Neither agree nor disagree

- Disagree
- Strongly disagree

13. “I believe that the autonomous vehicle is free of error.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

14. “I believe that I can depend and rely on the autonomous vehicle.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

15. “I believe that the autonomous vehicle will consistently perform under a variety of circumstances.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

16. “I believe that the autonomous vehicle provides alternative solutions.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree

- Strongly disagree

17. “I believe that I can control the behavior of the autonomous vehicle.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

18. “I believe that the autonomous vehicle will provide adequate, effective, and responsive help.”

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Demographics

Q19 What is your age

What is your gender

- Male (1)
- Female (2)
- Diverse (3)
- Prefer not to say (4)

What is your country of origin

- ☐ The Netherlands (1)
- ☐ Germany (2)
- ☐ Other (3)

Have you used one of the following voice assistants in the past? (multiple answers possible)

- ☐ Apple Siri (1)
- ☐ Google Assistant (2)
- ☐ Amazon Alexa (3)
- ☐ Microsoft Cortana (4)
- ☐ Mercedes Mbox (5)
- ☐ other (6)
- ☐ none (7)

How do you stand towards voice assistants within cars?

- ☐ Extremely positive
- ☐ Moderately positive
- ☐ Slightly positive
- ☐ Neither positive nor negative
- ☐ Slightly negative
- ☐ Moderately negative
- ☐ Extremely negative

Some modern cars provide partial automation (e.g. Tesla). Here does the driver need to pay attention to the car's actions to possibly intervene with the driving. High and fully automated cars do not require the attention of their user (as the car in the previous video). Do you have experiences with (partial, high, and fully) automated cars?

- ☐ Yes
- ☐ No

Skip To: Q26 If Some modern cars provide partial automation (e.g. Tesla). Here does the driver need to pay attent... = No

If so, did you (multiple answers possible)

- ☐ drive a partial automated car. (1)
- ☐ sit in a partial automated car while someone else was driving. (2)
- ☐ use a high or fully automated car (similar to the one shown in the video). (3)
- ☐ other (4)

How do you stand towards highly automated cars?

- ☐ Extremely positive
- ☐ Moderately positive
- ☐ Slightly positive
- ☐ Neither positive nor negative
- ☐ Slightly negative
- ☐ Moderately negative
- ☐ Extremely negative

You are almost done with the survey! Thank you for participating. However, you have not been fully informed about the real aim of the study yet. The purpose of the study is to investigate the interaction of a voice assistant's speech gender (male, female) and speech quality (natural human voice, synthesized voice) and its effects on the user's trust in voice assistants of automated vehicles and subsequently on trust in the vehicle itself. You have previously watched one of four possible videos, with a manipulated voice assistant that embodied one of the four conditions. Now that you have been fully informed about the real aim of the study, do you consent that your data are used?

- ☐ Yes, I do consent.
- ☐ No, I do not consent. My data shall be deleted.

Appendix B – Manipulated Videos

Video 1 (synthetic voice, female) https://www.youtube.com/watch?v=o_92hmRk2zg

Video 2 (synthetic voice, male) <https://www.youtube.com/watch?v=XtVY6pMp8po>

Video 3 (natural human voice, female) <https://www.youtube.com/watch?v=A9soxMpn19Y>

Video 4 (natural human voice, male) <https://www.youtube.com/watch?v=4GMANQINVss>

Appendix C – Literature Study Log

1. Research Questions and Main Concepts for the Literature Study;

Research Questions:

***RQ1:** How is the user's trust in the fully automated vehicle affected by the user's trust in the voice assistant?*

***RQ2:** How is the user's trust in the fully automated vehicle affected by the voice assistant's gender?*

***RQ3:** How is the user's trust in the fully automated vehicle affected by the voice assistant's speech quality?*

***RQ4:** How is the user's trust in the fully automated vehicle affected by the interaction of the voice assistant's gender and speech quality.*

Variables:

Automated Vehicles, Voice Assistants, Trust, Gender, Speech Quality

2. Materials and Databases

Scopus, Google Scholar and ScienceDirect have been used to collect the literature for this study. Scopus provides a wide range of sources from a variety of disciplines. Moreover, it is possible to narrow down the search on e.g. Social Sciences, which helps searching for relevant work. However, whenever search on Scopus does not result in much outcome, Google Scholar could be used. Finally, ScienceDirect helps whenever accessing sources with Scopus and Google Scholar does not work, since it provides easily accessible, and high-quality sources.

3. Search Actions and Results

Nr.	Date	Database	Action & Terms	Results
1.	5.03.20	Scopus	“Automated Vehicles” AND Trust	145
2.	5.03.20	Scopus	“Self-Driving Cars” AND Trust	61
3.	5.03.20	Scopus	Trust AND Technology	18,772
4.	6.03.20	Scopus	“Voice Assistants” AND Trust	7
5.	6.03.20	ScienceDirect	“Voice Assistants” AND Trust	39
6.	6.03.20	Scholar	“Voice Assistants” AND Gender	995
7.	12.03.20	Scholar	“Voice Assistants” AND “Speech Quality”	54
8.	12.03.20	ScienceDirect	“Voice Assistants” AND “Speech Quality	1
9.	12.03.20	Scopus	Synthetic AND Natural AND Voice	384
10.	12.03.20	ScienceDirect	Synthetic AND Natural AND Voice	8.091

11.	20.03.20	Scholar	Gender AND Automated Vehicles AND Voice Assistants	21.700
12.	20.03.20	Scopus	“Uncanny Valley”	448

4. Reflection

During my research I have mainly used Scopus, which was helpful to gather literature around trust in automated vehicles. However, specifically for the design variables of the voice assistant, the quality and accessibility of sources lacked. Thus, Google Scholar has been used mainly on that topic. Since Google Scholar provides a huge number of sources, it would have helped to narrow down the search by certain parameters, as it has been done with Scopus. Further, even though ScienceDirect provides easily accessible sources, the findings are limited in comparison to Scopus and Google Scholar, which might have led to a narrow view of a particular field.