

Hey Google, speak human to me!

A quantitative experiment of human-likeness and personalization in voice assistant interaction

Voice assistants are a new customer touchpoint for businesses, and the market for voice shopping is expected to grow substantially in upcoming years. Despite high societal and business interest, academic marketing research on consumer interaction with voice assistants is scarce. There is a critical need for application developers to understand how to best design the interaction with the voice assistants, and for brands to understand the impact of being advertised through the device. This study contributes to research by studying voice assistant interactions as a service encounter and as an advertising context. Specifically, it focuses on consumer interaction with voice assistant in a low-involvement purchase situation.

Through an audio experiment using a quantitative survey method, the study compares a series of two manipulations: making the voice appear more human-like, and the dialogue more personalized. 203 answers were collected using *Amazon Mechanical Turk*. Results strikingly show the importance of voice human-likeness for a positive perception of the device as well as improved brand attitudes. An explanatory path was found in how users anthropomorphized the device to a greater extent and experienced a sense of social presence during the interaction. However, dialogue personalization showed no significance for either dependent variable. Finally, attitudes towards the voice assistant were found to mediate the attitudes towards the brand. Such results will be useful for voice app developers and marketers alike as they establish the benefits of a strong focus on voice human-likeness. Likewise, this thesis is a first building stone to the nascent field of voice assistant research.

Keywords: Voice Assistant, Social Presence, Anthropomorphism, Service Encounter, Advertising Context

Authors:

Louise Bengtsson (23375)
Violaine Lulin (93087)

Supervisor:

Magnus Söderlund

Presentation date:

May 27, 2019

Examiners:

Sarah Jack
Jesper Blomberg

Word Count:

19 989

Acknowledgements

We would like to thank our supervisor, Magnus Söderlund, for his thoughtful comments and precious advice during all of the writing process.

*

A further warm thank you to Reema Singh for sharing her research experience on the topic of voice assistant and stirring us in an appropriate direction.

*

We are also thankful to ICA X for meeting us and sharing their experience and thoughts around the development of a Google action.

*

Thank you Sarah Wiggins for lending your voice to our experiment and impersonating an amazing voice assistant.

*

Eventually, thank you to all of the Master in Business and Management students for the great class spirit and mutual support.

List of definitions

Voice assistants	Software agents that can interpret human speech and respond via synthesized voices (Hoy, 2018)
Personalization	This study bundles two sub-definitions of personalization into one. Programmed Personalization which focuses on “making each customer feel like an individual” by using small talk, friendly greetings and names etc. Customized Personalization focuses on “offering services that are tailored towards the customers’ needs and taste” (Surprenant & Solomon, 1987)
Human-likeness	Human physical similarity in humanlike interaction and communications (speech, gaze, gestures) (Cheetham et al., 2011; Fink, 2012)
Synthesized voice	Artificial production of human speech (Hande, 2014)
Dialogue	Language used in conversation between two or more persons, or robotic computer agent (Merriam-Webster)
Conversational commerce	Commercial transaction through voice (Kinsella, 2018)
Human to robot interaction	“The science of studying people’s behavior and attitudes towards robots in relationship to the physical, technological and interactive features of the robots, with the goal to develop robots that facilitate the emergence of human-robot interactions” (Dautenhahn, 2007)
Service encounter	"The dyadic interaction between a customer and service provider" (Surprenant & Solomon, 1987)
Anthropomorphism:	“The tendency to attribute human characteristics to inanimate objects, animals and others with a view to helping us rationalize their actions.” (Duffy, 2003)
Social presence	“The feeling of proximity, closeness, warmth and sociability between agents” (Qiu & Benbasat, 2008)
Personal innovativeness	“The willingness of an individual to try out any new information technology” (Agarwal & Prasad, 1998)

Table of Contents

1. Introduction.....	7
1.1 Background	7
1.2 Problem discussion	8
1.3 Purpose and expected contribution.....	9
1.4 Research question	10
1.5 Delimitations	10
1.6 Thesis disposition	11
2. Literature Review	12
2.1 Defining voice assistants.....	12
2.1.1 Privacy and security	13
2.1.2 Personification of the device	13
2.1.3 Adoption and use of the voice assistant	13
2.1.4 Marketing through the device	14
2.2 Knowledge gaps	15
3. Theoretical Framework and Hypothesis Development	16
3.1 Anthropomorphism in robot interaction.....	16
3.2 Social presence in robotic interaction	18
3.3 Voice assistant as a service encounter.....	19
3.3.1 Trust	20
3.3.2 Service quality.....	20
3.3.3 Customer loyalty	21
3.3.4 Customer satisfaction	22
3.3.5 Social presence as a mediator for attitudes towards the voice assistant.....	23
3.4 Personal innovativeness as a moderator for attitude towards the voice assistant.....	23
3.5 Perception of brands and advertisement in voice assistants.....	24
3.5.1 Voice assistants as an advertising context.....	24
3.5.2 Effects of manipulations on brands	24
3.6. Summary of hypothesis	27
4. Methodology.....	29
4.1 Scientific approach to research design	29
4.2 Research design.....	29
4.3 Preparatory work for the main study	30
4.4 Pre-study 1: Understanding consumers and application developers.....	31
4.4.1 Pre-study 1.1: Task-application development at ICA	31
4.4.2 Pre-study 1.2: Consumer insights.....	32
4.4.3 Conclusions from consumer and industry insights	34

4.5 Pre-study 2: Choice of brand	34
4.5.1 Purpose and survey	34
4.5.2 Result and conclusion	35
4.5.3 Criticism of pre-study 2	36
4.6 Pre-study 3: Test of independent variables	36
4.6.1 Purpose and survey	36
4.6.2 Manipulation 1: Levels of voice “human-likeness”	36
4.6.3 Manipulation 2: Levels of dialogue personalization	37
4.6.4 Criticism of pre-study 3: voice and dialogue	38
4.7 Main Study	38
4.7.1 Pilot test of questionnaire	38
4.7.2 Main study questionnaire design	39
4.7.3 Dependent variables	39
4.7.4 Sampling approach	42
4.7.5 Analytical tool	42
4.7.6 Data quality check	43
4.8 Data quality	43
4.8.1 Reliability	43
4.8.2 Validity	44
4.8.3 Replicability	45
5. Analysis and Results	46
5.1 Hypotheses testing	46
5.1.1 Anthropomorphism	46
5.1.2 Social presence	47
5.1.3 Anthropomorphism as a mediator for social presence	48
5.1.4 Attitude towards the voice assistant	49
5.1.5 Social presence mediates the attitude towards the voice assistant	51
5.1.6 Personal innovativeness moderates the perception of the voice assistant	53
5.1.7 Attitudes towards voice assistant mediate the relationship with brand attitudes	54
5.1.8 Attitudes towards the brand	56
5.2 Summary of results	58
6. Discussion	60
6.1 Interacting with a voice assistant or an assistant with a voice?	60
6.2 Effects of voice assistants as a service encounter	61
6.3 Brand related effects of voice assistant interaction	63
7. Conclusion	65
7.1 Conclusion	65
7.2 Contribution to theory	66
7.3 Contribution to practice	67
7.3.1 Implication for task-application development	67
7.3.2 Implication for brand managers	68
7.4 Critique and limitations	68
7.4.1 Limitations to ecological validity	68

7.4.2 Limitations to internal validity	69
7.4.3 Limitations to external validity	69
7.5 Further research	70
8. References	71
9. Appendix.....	81
9.1 Pre-study 1: Understanding consumers and application developers.....	81
9.1.1 Interview guide: Interview with a task-application developer at ICA	81
9.1.2 Interview guide: Focus group with consumers	81
9.2 Pre-study 2: Brand choice.....	82
9.3 Pre-study 3: Test of independent variables	83
9.3.1 Voice manipulation pre-study questions	83
9.3.2 Dialogue manipulation pre-study questions	84
9.4 Voices ranked by levels of human-likeness	84
9.5 Dialogues ranked by levels of personalization	84
9.5.1 Dialogue 1 – D1 – low personalization	85
9.5.2 Dialogue 2 – D2 – medium personalization.....	85
9.5.3 Dialogue 3 – D3 – high personalization.....	86
9.6 Main study questionnaire items	86

Figures and Tables

Figure 1. VA Ecosystem Source.....	12
Figure 2. Conceptual Framework.....	16
Figure 3. Conceptual Framework with hypothesis.....	27
Figure 4. Experiment Design.....	30
Figure 5. Research Design.....	30
Figure 6. Mediation Anthropomorphism.....	49
Figure 7. Mediation Social Presence.....	52
Figure 8. Moderation Personal Innovativeness.....	53
Figure 9. Mediation Attitudes towards VA.....	54
Figure 10. Summary Hypotheses Results.....	58
Table 1. Summary of Hypotheses.....	28
Table 2. Pre-study Choice of Brand - Awareness.....	35
Table 3. Pre-study Choice of Brand – Attitudes.....	35
Table 4. Pre-study Independent Variable Voice.....	37
Table 5. Pre-study Independent Variable Dialogue.....	38
Table 6. Cronbach’s Alpha Dependent Variables.....	40
Table 7. Anthropomorphism Mean Voice.....	46
Table 8. Anthropomorphism Mean Dialogue.....	46
Table 9. Anthropomorphism Scheffe test Voice.....	47
Table 10. Social Presence Mean Voice.....	47
Table 11. Social Presence Mean Dialogue.....	47
Table 12. Social Presence Scheffe test Voice.....	48
Table 13. Mediation Anthropomorphism.....	49
Table 14. Mean Attitudes VA Voices.....	50
Table 15. Mean Attitudes VA Dialouge.....	50
Table 16. Attitudes VA Scheffe test Voices.....	50
Table 17. Dependent Variables Attitudes VA Voices.....	51
Table 18. Mediation Social Presence.....	52
Table 19. Moderation Personal Innovativeness.....	53
Table 20. Mediation Attitudes towards VA Voice.....	55
Table 21. Mediation Attitudes towards VA Dialogue.....	55
Table 22. Attitudes Towards the Brand Means Voice.....	56
Table 23. Brand Attitudes Scheffe test Voice.....	57
Table 24. Brand Attitudes Means Dialogue.....	57
Table 25. Brand Loyalty Scheffe test Dialogue.....	57
Table 26. Summary Hypotheses.....	59

1. Introduction

The introductory chapter provides a foundation for understanding the contemporary relevancy of voice assistant research. The section further elaborates on the purpose of this research paper and the delimitation and disposition of the thesis.

1.1 Background

By 2023, there will be 8bn voice assistants embedded in our daily lives: in our homes, in our cars, in our offices and in our stores. The growth rate of voice assistants is spectacular. In 2018 alone, the number of smart speakers in the US increased by 78%, to reach 118,5mn devices (Levin & Lowitz, 2019).. In the Netherlands, Google Home was adopted by a whole 5% of Dutch households within the first 4,5 months following its introduction (Kinsella, 2019). Amazon was first on the mass market, launching its software Alexa in November 2014. Since then, other smart speakers reached the market, each with their own assistant: Google’s Assistant and Apple’s Siri – embedded in the HomePod. Currently, the two most popular smart speakers on the market are Amazon Echo and Google Home, together sharing 94% of the market in the US (Levin & Lowitz, 2019). It is predicted that 55% of US households will have access to a smart speaker by 2022 (Braiker, 2018) and 3,3% worldwide by 2020 (Gartner, 2016).

Smart speakers are intended to act as personal assistants and can perform a wide range of tasks, from setting alarms, timers and reminders to sending and reading text messages, emails or even making phone call. The device answers basic informational queries, manage media playback from connected devices such as iTunes, Netflix and Spotify, and controls Internet-of-Things-enabled devices such as thermostats, lights, alarms and locks (Hoy, 2018). Third-party companies are developing further features, called “skills” by Amazon and “tasks” by Google. These can be described as “voice applications” and expand the smart speaker’s ability by interfacing with other programs via voice commands (Hoy, 2018; Smith, 2018). Such skills are, for instance, playing Jeopardy through the smart speaker, ordering a Starbucks drink, an Uber lift, or asking it to read the headlines from one’s favorite news provider. Another feature is to facilitate e-commerce purchases. ICA in Sweden has recently released Monica, their own “shopping assistant” accessible through a voice app on Google Home (ICA interview, 2019). Similarly, H&M launched an action on Google Home to let shoppers breeze through H&M Home’s catalogue, and order through the voice-app (Iribarren, 2018).

Smart speakers are therefore a new consumer touchpoint for brands (Smith, 2018) and will become an increasingly important sales channel in the upcoming years. It is predicted that by 2019, 27% of smart speaker users will make a purchase through the device, while 37% will shop through the device by putting items in their digital shopping cart (Hu, 2018). Furthermore, the market for voice shopping is expected to be worth 40 billion USD by 2022, from 2 billion in 2018. Currently, the four most shopped categories through voice are commoditized: grocery (20%), entertainment (19%), electronics (17%) and clothing (8%) (OC&C Strategy Consultants, 2018). Additionally, research shows that 84% of consumers expect voice-activated personal assistants to transform their expectations of companies within five years (Salesforce, 2018).

In a nutshell, the technology is still in its nascent stage and there are logistical difficulties for launching an e-commerce channel in Sweden. Yet the European e-commerce market is growing fiercely, reaching 9.4% between 2017-2018 (PostNord, 2018). As voice assistants are becoming more widely adopted, it is critical for retail providers to seize a first mover advantage. Businesses have great freedom in designing their own task application, both in terms of how much data to collect on the consumers, how to copywrite the interaction and what voice tonality to use (ICA interview, 2019). For these reasons it is fundamental for them to understand how the design of a task application will affect the consumer. To prepare for the future it is crucial for brands to develop a deep understanding of voice commerce, and a strategy to best reach out to potential customers. The channel presents new opportunities to gain customers, expand sales, and promote themselves.

1.2 Problem discussion

The pressing problem today is the gap between a large societal interest and the lack of research on understanding voice assistants as a sales medium. Despite promising voice shopping forecasts and the countless opportunities for businesses, the research field is new to the academic community and has not been extensively explored (Moriuchi, 2019).

From a company perspective, there is an important need for more research and understanding of the phenomena. This need is illustrated by two insights: first, the plethora of business reports publication from both large and smaller consulting firms (Moore, 2018; PwC, 2018; Chaudhuri & Terlep, 2018), as well as articles published by specialized retail industry websites and major international newspapers. These publications are emphasizing threats and opportunities associated to voice shopping, both in terms of sales opportunities but also brand variables. Some are even

making recommendations for adapting to the change, such as how to optimize your product or service offering description for voice-based search (Bentahar, 2018; Vozza, 2018). Second, by private initiatives of large corporations such as partnerships with voice assistant device manufacturers, and the development of companies' own voice app. Walmart has recently announced that it will partner with Google for voice-assisted grocery shopping (Meyersohn, 2019), and ICA has developed its own voice-app on Google Home that allows users to create shopping lists, ask for recipes, and more (ICA interview, 2019). Conclusively, there is an apparent interest from corporations in understanding the phenomena of voice commerce.

From an academic research perspective, the current bulk of research focuses on themes such as data security and privacy, adoption and usage, as well as personification of the device. These topics are related to the machine itself, and not on sales, voice commerce or brand relationship after a commercial interaction with the machine. There is therefore both a need and an opportunity to update and apply theories from a broad spectrum of disciplines, including: retail research, consume research, advertising research, and human-computer interaction research.

Ultimately, the medium of communication is particularly unique: it offers no visual support, and in a retail context it is very difficult to push promotional offers. Since the voice assistant differs fundamentally from other e-commerce supports – being a hybrid between an in-person sales encounter and digital shopping – it should be investigated.

1.3 Purpose and expected contribution

The purpose of the thesis is to contribute to the understanding of consumer interaction with voice assistants. More specifically, this thesis aims to understand how design elements in a voice assistant will affect consumers' perception of the device, as well as the brands being advertised through the medium. The study intends to broaden the understanding of how human-likeness and personalization of the device can affect how users perceive and internalize it. By testing ascending levels of human-likeness in the voice, the study strives to understand how this element affects the customer perceptions of the medium as well as the brands that are advertised through marketing messages on the platform. Similarly, the user's response towards the voice assistant personalizing the encounter by using collected data on them, is researched through testing ascending levels of personalization in the dialogue. Furthermore, an explanatory path for consumer responses are investigated through the extent to which the user regards the voice assistants as a human, in the degree to which they anthropomorphize the assistant and feels a

social presence in the interaction. Additionally, the thesis explores whether the user's reported personal innovativeness, in terms of familiarity with and interest in new technology, impacts their perception of the device.

The thesis is expected to contribute theoretically to a field of research that is yet in its nascent stage. As the interest of voice assistants is expected to grow significantly in the research community, this thesis aims to bridge the field of robotic anthropomorphism with voice assistants as a context for service encounters and advertisement. Moreover, the thesis strives to create an empirical contribution to organizations developing voice assistant task applications for commercial purposes. The thesis is one of the first to investigate how design characteristics of the voice applications can affect the marketing messages carried out through the applications.

1.4 Research question

The study will aim to answer the following research question;

Will the level of human-likeness in the voice and the level of personalization in the dialogue of a voice assistant affect users' perception of the device and the brands advertised through it?

The thesis aims to add nuance to the first research question by exploring two further questions;

What explains the different levels in perception of the VA and the brands advertised through it?

Will customers' reported interest and familiarity with technology effect how they perceive the interaction with the device?

1.5 Delimitations

For the sake of feasibility, this thesis is delimited in a number of ways. Firstly, the study is delimited in terms of what is being explored. In order to expand the research topic and enable comparative conclusions, the study researched both context and content of a voice assistant interaction. Context in terms of voice tonality and content in terms of copywriting of the dialogue. However, the content is delimited to the situation of purchasing and ordering a selected product, toothpaste. Similarly, the context is delimited to only studying the voices selected for this experiment. Yet we believe that by testing three levels in each design element, the study captured multiple variations of a personalized dialogue and human-like voice.

Secondly, the delimitation was made to focus the study on product of FMCG industry. For reasons of simplification the research is delimited to only testing the effect on a single brand and has not compared across product categories.

Moreover, the study is delimited in terms of participants. The study was distributed through *Amazon Turk* which allows for a demographically diverse respondent pool (Buhrmester, Kwang & Gosling, 2011). The respondents were mainly male (67%) and from the US (57%) and India (36%), which shows delimitations in the sample group.

Lastly, the delimitation had to be made in terms of what dependent variable to measure. Based on our preliminary interviews, voice assistants as a service encounter and as an advertising context appeared as dependent variables that could bring much practical contributions.

1.6 Thesis disposition

The thesis is structured into six parts; *Introduction*, *Literature Review*, *Theoretical Framework*, *Methodology*, *Results*, *Discussion* and *Conclusion*. The initial *Introduction* section deals with the topic of voice assistants at large, outlining industry trends, research and defining contemporary problems with adapting the technology as an e-commerce channel. The *Literature Review* and *Theoretical Framework* goes to present previous research of chosen concepts and subsequently elaborates on dependent variables for establishes connecting hypotheses. The *Methodology* section presents choices and motivation for the research design. Moreover, it presents purposes, methods and findings of the 3 pre-studies that were used to form the main study. The chapter is concluded with a discussion on the data quality of the main study. The *Results* section presents and analyses data collected in the main study. Finally, the *Discussion* reasons around the finding and the *Conclusion* summaries the study and how it can contribute to research and practice.

2. Literature Review

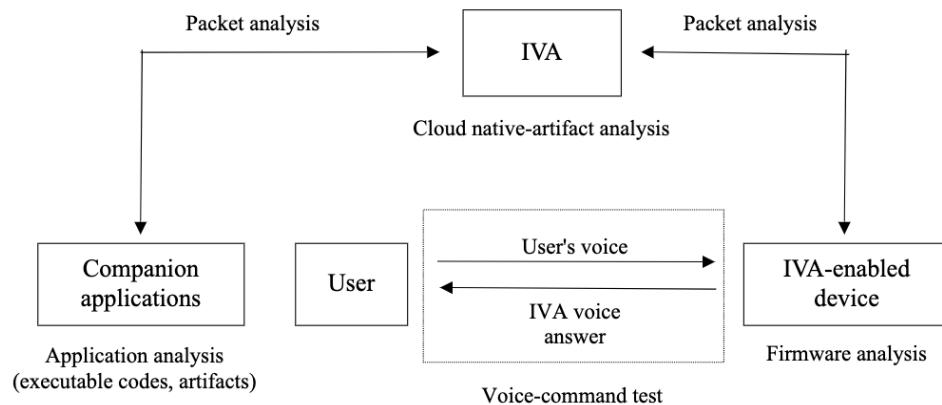
The following chapter explores previous research within the field of intelligent voice assistants. Firstly, a definition of voice assistant from previous research will be presented. Secondly the chapter elaborates on four fields of research on voice assistants. Lastly, a knowledge gap in current research is presented.

2.1 Defining voice assistants

Voice assistants are software agents designed to provide aid to users in conducting daily routine tasks, such as answering phone calls, taking notes, shopping, making appointments, web browsing, and alike (Orehovacki, Etinger & Babic, 2019). They have been referred to as Intelligent Virtual Assistants, smart assistants, digital assistants and Personal Virtual Assistants. For reasons of simplification and clarity, the term Voice Assistants (VA) will be used throughout this thesis.

VAs are derived from chatbots, AI-based natural language processing systems that strive to generate a conversation between a human and machine. Following the reception of a voice command, VAs' "intelligence" performs linguistic analysis and carries out the requested action (see figure 1). The system's "brain" is a cloud-hosted service, and the device in itself does not carry embedded intelligence (Chung et al., 2017).

Figure 1. VA Ecosystem. Source: (Chung et al., 2017)



Previous VA research has primarily been made within the field on computer science. A search on Scopus.com for the term "Voice Assistant*" gives 91 results, out of which 77 are within the subject area of computer science (16/4-2019). Four themes studying consumer perception of voice assistants have been identified.

2.1.1 Privacy and security

Users experience of privacy and security in an interaction context with voice assistants has been studied by Lau et al. (2018). They found that many users feel deeply uncomfortable with the idea of a “microphone-based” device that a company with malicious intent could use to listen in on their homes. Indeed, microphones are perceived as one of the most privacy-invasive sensors next to video cameras (Chung et al., 2017). People are concerned about speaker data being used for targeted advertising, in the same way as the data from their online behavior is used. In addition, the study found that most participants would feel uncomfortable if their smart speaker voice logs were used for targeted advertising (Lau et al. 2018). A study from Ford et al. (2018), during which the researchers plotted Amazon Dot Echo in households for 21 days, shows that the device recorded private home conversations without users using a wake word, and that not all interaction recordings were properly logged to the Alexa application. Wiretapping and device hacking are therefore important worries regarding voice assistants (Chung et al., 2017).

2.1.2 Personification of the device

Previous research has studied the ways in which users assign human characteristics to the voice assistant through the process of anthropomorphism. Purington et al. (2017) have found in a study of Amazon Echo’s reviews that half of reviewers refers to the device with the personified name Alexa, with some even using the pronoun “she”. The same study finds that the device can be considered as a companion and a conversation partner, or even as a friend or a family member. Software developers are furthermore trying to give the program a personality. This is expressed by Alexa being given a gender and a name, but also people perceiving Siri as being “sassy” and “friendly”. The perception of the software’s personality is further influenced by attributes such as having a regional accent (Cowan et al., 2017). Eventually, Sciuto et al. (2018) finds that Amazon’s Alexa adopters usually test Alexa in different ways to see how intelligent the device is by asking personality related questions about Alexa’s personal life.

2.1.3 Adoption and use of the voice assistant

How and in what setting a voice assistant is used has been studied by previous research. Porcheron et al. (2018) studied the interaction of a family with an Alexa-enabled device. They found that people perceived awkwardness and artificiality in the dialogue with the machine, specifically in the time lag between emitting a request and waiting for the reply. They further studied how the smart speakers are embedded into conversational settings like family dinners

where various simultaneous activities are being conducted and how requests to the device are embedded into the conversational setting. Sciuto et al. (2018) brought light on how users probed the smart speaker capabilities by testing queries for which the users are uncertain whether they are supported by the device.

Furthermore, Lau et al. (2018) found that the primary factors driving smart speaker adoption are convenience and identity as an early adopter. Convenience is defined as the utility that they get from using the smart speaker frequently and effectively. Identity as an early adopter is the fact that users recognize themselves as someone who derives pleasure from being among the first users of new gadgets. On the other hand, as their study finds, the primary motivators for not using smart speakers are perceived lack of utility, and privacy and security concerns (Lau et al., 2018).

2.1.4 Marketing through the device

Academic research on voice assistants as a context for marketing messages is clearly limited, and to our knowledge there are few prominent articles on this subject. Moriuchi (2019) researches user's adoption of voice assistants and found that consumers who use the device for transactional purchases make habitual purchase where little thought processing is required. Similarly, those purchasing activities that are non-transactional appear to be made early in the buyer journey, at the stage of awareness and consideration for alternatives.

Kim et al. (2018) tested consumers responses to advertising messages made through Amazon Alexa and found that consumers that were asked interactive questions in relation to the product reported better ad-recall. Similarly, they found that advertising for low involvement products showed a higher advertising effectiveness, as these require less involvement and information. For both high and low involvement products, contextual relevance was essential for recall.

Smith (2018) researched which types of marketing messages users find acceptable on smart speakers. They suggested that cognitive messages that combine practical information with rational arguments would utilize cognitive produces of the user and result in better recall. Similarly, they concluded that the most essential was that message provides value for the listener.

2.2 Knowledge gaps

Despite growing academic and business interest in the field of voice assistant research, a review of published research highlights how scarce investigations of marketing aspects of voice assistants are. If some have published works about reactions to advertising messages (Smith, 2018; Kim et al., 2018), none have studied the particular impact of the voice assistant medium on brand attitude. Similarly, despite extensive research on sales encounters in a retail setting as well as on the effects of interaction personalization, no academic work has applied such theory to the context of voice assistants.

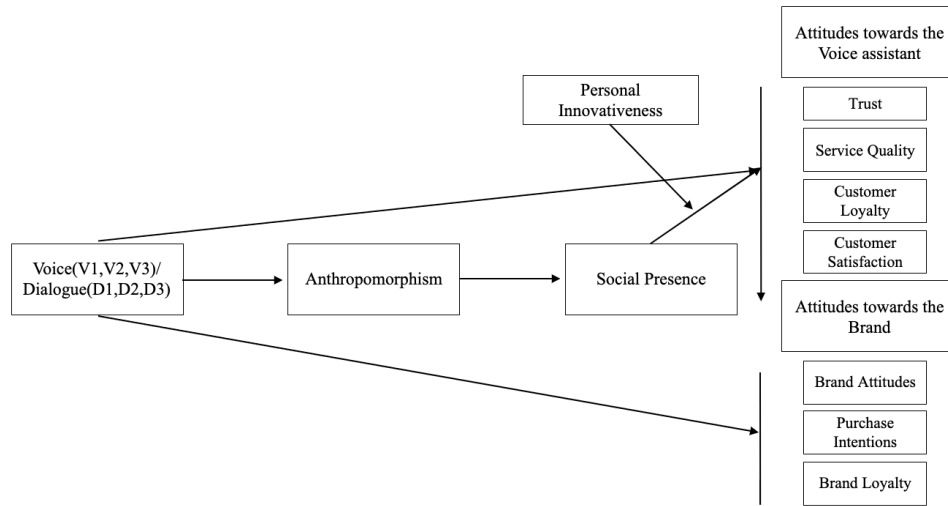
Overall, no studies have, to our knowledge, studied voice assistants as a setting for service encounters and advertising context. This thesis therefore intends to bring together academic fields that have rarely been associated before and fill this gap to contribute to research on voice assistants from a marketing perspective.

3. Theoretical Framework and Hypothesis Development

The following chapter will begin with presenting our conceptual framework. Thereafter it will introduce the hypotheses that form the foundation of this study by outlining the theories that justify these hypotheses. Lastly a summary of the hypotheses will be presented.

The following conceptual model presents an overview of the hypotheses of this study. Theoretical justifications for these hypotheses will be elaborated on further in this chapter. Theory on anthropomorphism, social presence, service encounter, advertising context and customer-based brand equity have been applied to study the under-researched field of consumer perception of voice assistants.

Figure 2. Conceptual Framework



3.1 Anthropomorphism in robot interaction

Anthropomorphism is the process of attributing cognitive and emotional characteristics that are uniquely human to a non-human agent (Duffy, 2003). It is a recurring theme in the field of human-robot interaction and a process that humans do to rationalize the behaviors of non-human agents (Epley et al., 2007). The process is unintentional despite being rationally aware that the object cannot exhibit any human traits (Duffy, 2003).

Previous research on robotic anthropomorphism has mainly focused on behavioral changes resulting from the anthropomorphizing of the agent, few have measured how users internalize it (Zlotowski et al., 2014). Nass et al. (1995, 2001, 2005) have long researched how the voice of a robot effect the personality we assign to it. Walters et al. (2008) researched how three different levels of human-likeness in robotic voices impact their willingness to physically approach the

robot and found that people are more likely to break norms of human-human spatial zones the more they anthropomorphize the robot. Similarly, Sims et al. (2009) tested three different levels of artificially synthesized voices and found that as they anthropomorphized the robot, they extended the rules of human interactions to that with the robot. Such findings generate the hypothesis that the higher degree of perceived human-likeness in the voice, the more the voice assistant will be anthropomorphized.

To understand our ability to anthropomorphize robots, Zlotowski et al. (2014) studied it as a reverse process of dehumanization. They divided traits into *uniquely human* and *human-nature*. Such traits that fall into the *uniquely human* are such as broadminded, humble, conservative, and rude. A human without these traits would make it appear more non-human and animal-like. In contrast, traits related to *human nature* are those that, when absent, would make humans appear more mechanical. Example of these emotional traits are friendly, sociable, trusting and aggressive. Robots appeared more human-like when they were described with *human nature traits*, whereas traits that were *uniquely human* made the robot seem more intelligent and lifelike, but not necessarily more human. Zlotowski et al. (2014) concluded that people already associate and expect intelligence from robots and these trait does therefore not increase their level of anthropomorphism nor make them perceive the robot as more human-like. The dialogue of the voice assistant would become more personalized using social language, by using affective responses such as humor and self-disclosing behavior (Hamman, 2006). Therefore, one can expect that it will exhibit more *human nature* traits such as sociable and trusting. Consequently, the higher degree of personalization of the dialogue, the more the voice assistant should be anthropomorphized.

H1a. The higher degree of perceived human-likeness in the voice the more the voice assistant will be anthropomorphized

H1b. The higher degree of personalization of the dialogue, the more the voice assistant will be anthropomorphized

3.2 Social presence in robotic interaction

The term social presence has been defined as the “sense of being with another” (Biocca et al. 2003). It refers to the feeling of proximity, closeness, warmth and sociability between agents (Qiu & Benbasat, 2008). This theory is most often discussed with regards to a communication medium’s ability to product an illusion of physical proximity between humans (Hess et al., 2009). High levels of social presence are found in physical human to human interaction; however the concept has been used to explain how an interface of artificial intelligence in robots can carry social cues in a human to robot interaction (Biocca et al. 2003).

The voice of an anthropomorphized computer agent has shown to generate a greater sense of social presence than less anthropomorphized, more synthesized computer voice (Lee & Nass, 2005b). Social presence is especially strong if the voice appears similar to the user’s in terms of the expressed personality (Lee & Nass, 2005). The same study also found that extroverted voices are more likely to create social presence than introverted ones. One might expect that as the voices of a voice assistants become less synthesized, voice assistants will also appear more extroverted. A further study by (Chérif & Lemoine, 2019) found that consumers who interacted with a virtual assistant having a human voice experienced a stronger sense of social presence than those interacting with a voice assistant having a synthesized voice. Therefore, one would expect that this study will find similar results and that sense of social presence will increase as voice synthetization decreases.

The use of social language has been shown to generate social presence, even to a greater extent than the use of visual cues (Hamman, 2006). The study defines social language as threefold: interactive responses such as referential posts and use of first name; affective responses, such as the self-disclosure, humor and emoticons; and cohesive responses, such as inclusive terminology like the word “we”. Many websites use strategies to generate an online social presence, such as by using the consumers’ name and making e-mail communication more personalized (Gefen & Straub, 2004a). Consequently, as dialogues with the voice assistant become more personalized, they will use both interactive, affective and cohesive responses. As a result, one would expect the user to experience a greater sense of social presence the more personalized the dialogue appears.

H2a. The higher degree of perceived human-likeness in the voice, the more social presence will be experienced in interaction with the voice assistant

H2b. The higher degree of personalization of the dialogue, the more social presence will be experienced in interaction with the voice assistant

Several previous studies have suggested that improving anthropomorphized elements of a robot will enhance the social presence in the interaction (Nowak & Biocca, 2003). Edlund et al., (2008) suggested the benefits of anthropomorphism in spoken dialogue systems and Stern et al., (2006) compared across synthetic and human voices. Gong (2008) found that the levels of anthropomorphism correlates with the level of social presence. Similarly, they found that anthropomorphism and social presence affected the users' sense of trustworthiness in the agent. Building on these studies, one can hypothesize that anthropomorphism will act as a mediating explanatory variable for social presence.

H3a. Anthropomorphism will function as a mediator for social presence in terms of human-likeness in voice

H3b. Anthropomorphism will function as a mediator for social presence in terms of personalization in dialogue

3.3 Voice assistant as a service encounter

Several studies have shown that people tend to apply the same social behavioral tendencies towards a robot as they would towards a human (Reeves & Nass, 1996; Nass & Moon, 2000). The interaction between a human and a robotic assistant is similar in nature to that of a human-service provider and their customer (Zhuang et al., 2011; Bellini & Convert, 2016; Tunvall, 2018). These interactions are defined as a service encounter, and the dynamics of "the dyadic interaction between a customer and service provider" (Solomon et al., 1985) have been extensively researched (Arnould & Price, 1993; Pugh, 2001; Bitner, Booms & Tetreault, 1990). This relationship is characterized by two conflicting goals: efficiency and personalization (Surprenant & Solomon, 1987). During a service encounter, customers and employees are interacting through a "service script" of inexplicit role expectations from both parties (Grove, Fisk & John, 1983). According to Surprenant and Solomon (1987) personalization is a behavior that neglects these scripted roles and aims to individualize the customer. These behaviors include both verbal and non-verbal communication as well as offering customized products and services tailored towards the individual's needs. Both the use of social language and customer knowledge can be interpreted as indicators of personalization (Surprenant, Solomon, 1987; Pugh, 2001). Coelho and Henseler, 2012 defines how personalization impacts the perception of the service encounter through 4 dimensions: Trust, Service Quality, Customer Loyalty and Customer Satisfaction.

3.3.1 Trust

Personalization of the service encounter is key to creating a trusting customer relationship and the connection has been well established empirically in previous studies (Hawes et al. 1989; Swan, Bowers & Richardson, 1999; Morgan & Hunt, 1994). Ligas (2004) explains how trust in a service encounter is created through a three-stage process. Firstly, the service provider create rapport by attempting to interpret the customer cues and adapt the process towards the customer's expectations. Secondly, to create reassurance by listening to the customer and using humor. Lastly, they relate to the customer by validating their value to the firm. Shamdasani, and Balakrishnan (2000) further validates the idea that the employees' levels of friendliness, knowledge about the customer and ability to disclose information about themselves generated trust in the relationship. This effect has been replicated for an e-commerce setting. It has been indicated that if website vendors showed consumer knowledge in the website experience, customers will form positive attributes and trusting intentions with the e-commerce website (McKnight et al., 2002).

An increased sense of trust in the voice assistant can also be expected as the voices are manipulated to become more human-like. (Nass & Lee, 2001) found that computer-synthesized voices were able to generate liking and trust towards the computer the more similar its voice was to that of the user. Moreover, Elkins & Derrick (2013) indicated how a wider vocal range can generate trustworthiness for a person. Lastly, Chérif (2019) found that trust was higher for human voices than synthesized ones. As a voice assistant's voice becomes more human-like, increases in trustworthiness can be expected.

3.3.2 Service quality

Furthermore, personalization in service encounters have also shown to influence customers' reported level of perceived quality of the service. Personalization is an underlying dimension of several of the key determinants by Parasuraman (1985) framework of key service quality determinants. Courtesy, credibility, reliability, access, communication and understanding the customer are all concepts affected by a level of personalization (Coelho & Henseler, 2012). For businesses delivering services in interactive encounters with customers, personalization was indicated to be the most determinant factor for perceived service quality (Mittal & Lassar, 1996). Attempts to personalize the service encounter is seen as a signal for both higher quality of the service and of the service provider themselves. Accordingly, Yang (2001) found that the level of

personification significantly affected the overall service quality assessment of internet service providers, such as e-commerce websites.

In terms of voice, it has been found that the more users anthropomorphize a synthesized voice, the higher their expectations on the quality of the computer's intelligence and functionality (Nass & Brave, 2005). Pugh (2001) found that displays of positive emotions in voice tonality were related to better perceived quality. For these reasons, one can expect that both a more human-like voice and personalized dialogue will generate a positive impression of the users perceived service quality.

3.3.3 Customer loyalty

Customer loyalty can in general be explained by social exchange theory, the idea that people stay with a service provider as long the attractiveness value of the alternative minus its switching costs does not exceed the value of the current alternative (Thibaut & Kelley, 1959). Service personalization allows for the retention of customers through building up a relationship between the customer and the service provider, by learning about the customer and exhibiting this knowledge by providing an individualized service. It requires the service provider to spend time and effort learning about the customer and for the customer to express their needs, which may decrease economies of scale but can payoff due to increased customer retention and loyalty (Coelho & Henseler, 2012). On the contrary, Ball et al. (2006) found that it was not behavioral loyalty, the intention of repurchasing behavior, but rather emotional loyalty that was enhanced during a personalized service encounter. They defined the concept as “the desire on the part of the customer to continue the relationship even if competitors have lower prices, willingness to recommend to friends, and intention to continue to patronize”. Burnham et al. (2003) further found that personalized service encounters create relationship equity, “consumer’s willingness to stay with the brand beyond objective measures”. As the dialogue becomes more personalized, it can hence be expected that customer loyalty will increase.

The level to which a computer is anthropomorphized has shown to impact user loyalty towards the computer. Sundar (2004) was even able to show that the level of anthropomorphism of a computer was a stronger indicator of behavioral loyalty towards than personal preferences for consistency. Specifically, in terms of voice assistants, it has been shown the more engaged users feel with the assistant, the more behavioral loyalty they will exhibit (Moriuchi, 2019). As the

voice becomes more human-like, it can be expected to result in increased customer loyalty towards the voice assistant.

3.3.4 Customer satisfaction

Customer satisfaction is the last concept that Coelho & Henseler (2012) conceptualize as an output from personalization. It is a multi-dimensional concept, often described as the value in relation to price of a service (Cronin, Brady & Hult, 2000). Service encounters that are individualized towards the consumer should naturally improve the satisfaction of the service. Empirical findings from many previous studies have been able to confirm this positive relationship (Srinivasan et al., 2002; Meuter et al., 2000). Similarly Surprenant (1987) reported that customization decreased customers' satisfaction with the service providers' competence and rather employee sociability in terms of warmth and friendliness had the most significant effect on satisfaction. These are behaviors of sociability that divert from the role-script of service interactions (Grove et al., 1983) and can hence be interpreted as a personalization of the service encounter (Surprenant & Solomon, 1987).

Kiseleva et al., (2016) researched ways of measuring satisfaction in interactions with voice assistants. They found that what users considered satisfactory varied to a great extent depending on the complexity of the task. The most important factors for user satisfaction for complex mission tasks were to provide elaborate answers, clarity and understandability of the voice assistants way of communicating. Nass & Lee (2001) were also able to indicate how a more human-like voice will result in better satisfaction with the device. These findings give rise to the hypothesis that both voice human-likeness and dialogue personalization improves customer satisfaction.

H4a. A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.

H4b. A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.

- a. Trust
- b. Service Quality
- c. Customer Loyalty
- d. Customer Satisfaction

3.3.5 Social presence as a mediator for attitudes towards the voice assistant

The level of social presence conveyed through a medium has shown to relate to the sense of trust experienced towards the medium itself. This has been particularly studied in the context of e-commerce (Gefen, 2002). By building social presence, an e-commerce website can enhance both trust and purchase intentions towards the medium, as well as perceived usefulness and enjoyment (Gefen & Straub, 2004b; Cyr et al., 2007). It has been proved that social presence has a mediating effect on both trust and purchase intentions in an e-commerce setting (Weisberg, Te'eni & Arman, 2011). In a voice assistant context, Chérif (2019) similarly found that social presence mediated the relationship to behavioral intentions. For this reason, it is hypothesized that social presence will have a mediating effect on the attitudes towards the voice assistant.

H5. Social presence will function as a mediator for attitudes towards the voice assistant in terms of:

- a. human-likeness in the voice
- b. personalization in the dialogue

3.4 Personal innovativeness as a moderator for attitude towards the voice assistant

According to individual innovativeness theory, the personal innovativeness level of a consumer affects their likelihood to adopt a new technology (Lu et al., 2005; Rogers, 1983). The idea of personal innovativeness can be defined as “the willingness of an individual to try out any new technology.” (Turan et al., 2015). A person with high personal innovativeness is characterized by, among others, prior ownership of product, interest in novelty and active information search for new information (Vishwanath & Goldhaber, 2003). Especially product involvement, particularity in terms of previous interactions, has shown to moderate the relationship with consumers behavior with new technology (Bloch, 1981).

The level of personal innovativeness does not only affect user’s likelihood of adopting new technology but will also enhance their perception of the interaction with the voice assistant. This relationship has previously been uncovered by research on consumers’ interaction with online-services. McKnight et al., (2002) were able to show a relationship between a user’s level of personal innovativeness with their perceived trustworthiness in an e-commerce website. Several studies found that personal innovativeness affected the both users’ perceived ease of use with IT

products as well as their trust in the service provider (Hwang, 2009). Finally, Agarwal and (1998) have found that the level of personal innovativeness of a user moderates the relationship with their perception of new information technology. For these reasons, it can be expected that this relationship will hold in a context of voice assistants.

H6. Personal innovativeness will function as a moderator for attitudes towards the voice assistant in terms of:

- a. Human-likeness in the voice
- b. Personalization in the dialogue

3.5 Perception of brands and advertisement in voice assistants

3.5.1 Voice assistants as an advertising context

The context of brand recommendations through a voice assistant can be viewed as a form of advertising context. Advertising context is a broad concept composed of both the media context and receiver context (Karlsberg, 2016). Looking at voice assistants as a media context, it has been found that both involvement with, and affect towards the context has positive effects for the brand that is being advertised (Pelsmacker et al., 2002; Moorman et al., 2012). If the receiver feels involved in the medium in terms of enhanced attention, the ad and consequently the brand will be more favorably reviewed (Moorman et al., 2012; Krugman, 1983). If the receiver feels strong affection towards the media context, this will also spill over towards the brand being advertised (Moorman et al., 2012; Krugman, 1983; Goldberg & Gorn, 1987). Advertising context has previously mainly focused on media context such as magazine, TV and web advertising segments, however it is likely that these findings are applicable in the context of voice assistants as well. Hence it is natural to hypothesize that respondent's affect towards the voice assistant will spill over on their attitudes towards the brand being advertised.

H7. Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of:

- a. Human-likeness in the voice
- b. Personalization in the dialogue

3.5.2 Effects of manipulations on brands

When measuring the effect of advertising on brands through the perspective of the customer, customer brand-equity is a common measure. The concept is defined as “consumers overall attitude with brand, and anything linked in memory to the brand” (Mitchell & Olson, 1981a). Rosengren & Dahlén (2015) defined the concept as consisting of three dimensions: Brand Attitudes, Purchase Intentions and Brand Loyalty.

3.5.2.1 Brand attitudes

The concept of brand attitudes have been defined as individual's internal evaluation of the brand." (Mitchell & Olson, 1981) When seeing voice assistants as a context for advertisement, personalization of the advertising message in terms of relevancy for the consumer has been linked with better attitudes toward the advertisement and product promoted (Leung & Cheung, 2004). Favorable ad attitudes is similarly strongly related to improved brand attitudes (Stone et al., 2000).

When viewing voice assistants as a platform of service encounters, it has previously been stated how personalization leads to better satisfaction with the service (Coelho & Henseler, 2012). With this in mind, Choi (2016) studied the setting of a retail department store, where the sales-employee were selling brands for a third party. They found that the customer satisfaction with the salesperson spilled over to the brand the sales representative endorsed. Similarly, Brexendorf et al. (2010) found that salesperson interaction competence and customer satisfaction with the sales encounter were positively associated with brand attitudes. Interaction competence was in this study defined as the salesperson's ability to fulfil interpersonal needs by being a cooperative partner and listen to the customer in order to pick up on details of how to help them in an appropriate manner. Consequently, personalization of the dialogue can be hypothesised to create better brand attitudes towards the brand being advertised in the interaction.

In terms of the effect of human-likeness, Nass and Lee (2011) have found that users are more favourable towards voices that sounds like our own in terms of perceived personality. Enhancing similarities between sales representatives and associates have long been researched as an effective method of influence for sales (Cialdini, 1993). Likewise, Jiang et al., (2010) found that similarities between salesperson and customer can significantly improve attitudes in a sales context. For these reasons, human-likeness in voice can be expected enhance brand attitudes.

3.5.2.2 Brand purchasing intentions

Purchase intentions are personal action tendencies relating to the brand (Bagozzi et al. 1979; Ostrom, 1969). As previously stated, personalization of service encounters has been linked with increased service quality (Parasuraman et al. 1985). Likewise, literature has identified that improved service quality enhances purchase intentions of a brand. This relationship is particularly significant for private label brands, in which brand purchase intention has a strong correlation with service quality (Wu et al., 2011). Pornpitakpan et al. (2017) were able to indicate how retail service quality also impacted the likelihood of impulse buying of a brand. Similarly, by personalizing the encounter through showing responsiveness; empathy and courtesy of the

salesperson have, independently of previous product perception, a direct effect on purchase intentions (Sweeney et al., 1997). Therefore it can be expected that a personalized dialogue will increase purchase intentions.

As voice becomes more human-like, both parties are likely to experience a greater extent of social presence in the interaction (Gong, 2008). Social presence has shown an effect on purchase intentions (Gefen & Straub, 2004) in e-commerce settings, hence this can be expected to apply to the setting of a voice assistant interaction.

3.5.2.3 *Brand loyalty*

Brand Loyalty has been defined as a biased behavioral response that is expressed over time by some decision-making unit, with respect to one brand out of several alternatives (Jacoby & Kyner, 1973). As previous presented, personalization leads to better satisfaction with the service (Coelho & Henseler, 2012). Similarly, clarity and understandability has shown to improve the satisfaction with a voice assistant (Kiseleva et al., 2016)

With regards to customer satisfaction, Lemmink & Bloemer (1992) found that satisfaction with a service has a mediating effect on satisfaction with brand loyalty towards the brand that they decide to purchase during the sales encounter. Furthermore, Bloemer & Kasper (1995) stated that customer satisfaction was would moderate the relationship with brand loyalty.

In line with these arguments, loyalty and trust in a person spills over to the things the he endorses (Björkman & Kock, 1995). Feelings of trust towards the sales representative are moderating factors for brand loyalty during the brand decision process (Chow & Holden, 1997). For all these reasons, it can be expected that brand loyalty will increase as voices becomes more human and dialogue more personalized.

H8a. A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.

H8b. A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.

e. Brand Attitudes

f. Purchase Intentions

g. Brand Loyalty

3.6. Summary of hypothesis

Based on this theoretical foundation, eight hypotheses have been formed which explores five areas of research. These hypotheses will compose the base for our main study, analysis and conclusions about the research questions. Following is a summary of all hypotheses.

Figure 3. Conceptual Framework with hypothesis

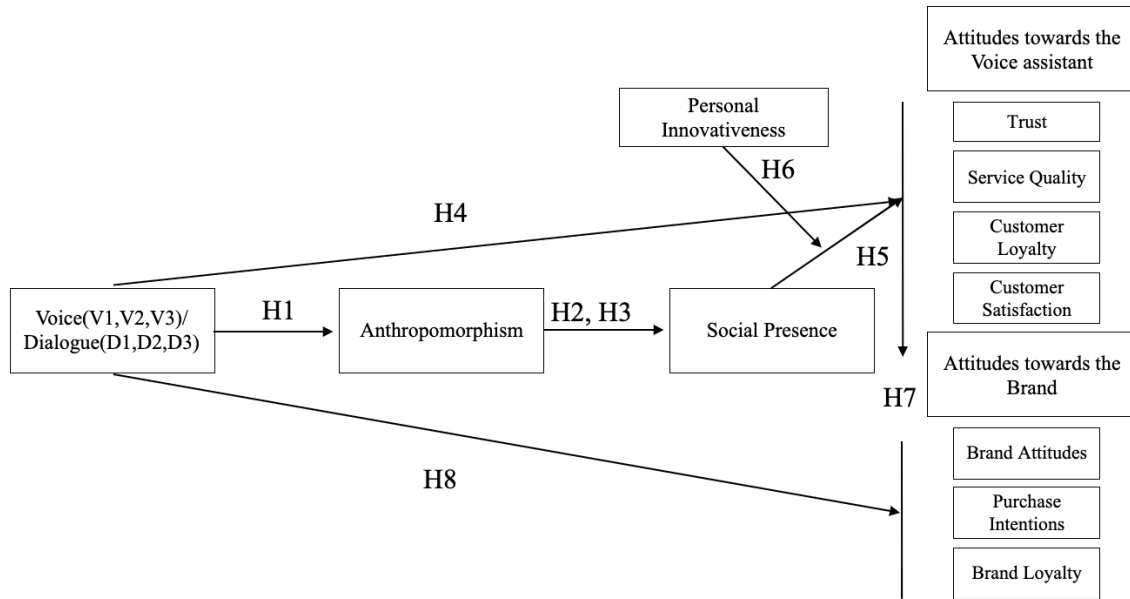


Table 1. Summary of Hypotheses

<i>Anthropomorphism</i>	H1	H1a	The higher the degree of perceived human-likeness in the voice, the more the voice assistant will be anthropomorphized
		H1b	The higher the degree of personalization of the dialogue, the more the voice assistant will be anthropomorphized
	H2	H2a	The higher the degree of perceived human-likeness in the voice, the more social presence will be experienced in interaction with the voice assistant
		H2b	The higher the degree of personalization of the dialogue, the more social presence will be experienced in interaction with the voice assistant
<i>Social Presence</i>	H3	H3a	Anthropomorphism will function as a mediator for social presences in terms of human-likeness in voice
		H3b	Anthropomorphism will function as a mediator for social presences in terms of personalization in dialogue
<i>Attitudes towards Voice Assistant</i>	H4	H4a	A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.
		H4b	A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.
	H5	H5a	Social Presence will function as a mediator for Attitudes towards the voice assistant in terms of human-likeness in the voice
		H5b	Social Presence will function as a mediator for Attitudes towards the voice assistant in terms of personalization in the dialogue
<i>Personal Innovativeness</i>	H6	H6a	Personal innovativeness will function as a moderator for Attitudes towards the voice assistant in terms of human-likeness in the voice
		H6b	Personal innovativeness will function as a moderator for Attitudes towards the voice assistant in terms of personalization in the dialogue
<i>Attitudes towards the brand</i>	H7	H7a	Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of human-likeness in the voice
		H7b	Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of personalization in the dialogue
	H8	H8a	A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.
		H8b	A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.

4. Methodology

The following chapter elaborates on the scientific approach to the research design. Conclusions and insights from the qualitative pre-studies with industry and consumers are presented. Thereafter, the chapter will illustrate the preparatory work for the main study through three pre-studies. Subsequently the choices for the design of the main-study will be explained. Lastly the chapter will elaborate on the data quality of the study.

4.1 Scientific approach to research design

This study has been built from a deductive approach to research. As mentioned, there is a lack of research in the field of voice assistants. As a result, literature applicable to retail service encounters and advertising context has been combined with theory of human-robot interaction to generate hypotheses. These have subsequently been empirically tested through analysis (Bryman & Bell, 2015). The study strives to test the causal effect between several manipulations in the design of the voice assistant with chosen independent variables. Hence an experimental study through a quantitative survey was deemed most suitable, as this is a common research approach in the field of human-computer interaction (Dautenhahn, 2007). Due to the lack of research on this topic area, the choices of independent and dependent variables were supported by insights found through qualitative pre-studies. As for our justification of using an imaginative scenario experiment as a priming manipulation for the main study survey, this is a widely accepted methodology able to generate interesting yet useful and rigorous research findings (Ramirez et al., 2015). A deductive research approach is the most common method to be used for marketing research (Hunt, 2010) and as a result we deemed it most suitable for our main study. However, we believe an inductive and adductive approach could potentially have brought new contributions to the field of voice assistant research.

4.2 Research design

The main study was a quantitative survey using an experimental research design priming the respondent with audio-manipulations. This design allowed for deductively testing existing theory with collected data (Bryman & Bell, 2015). The main study tested changes for two independent variables; Manipulation 1: Voice, and Manipulation 2: Dialogue. The experiments had a 2x3 group design in order to test three different levels of each manipulation. Respondents were randomly assigned to one experiment and were only exposed to the manipulation once. Three different manipulations of Voice were recorded (V1, V2, V3), with the relationship $V1 < V2 < V3$

in terms of perceived human-likeness. Manipulation 2 had three different levels of Dialogue (D1, D2, D3), with the relationship $D1 < D2 < D3$ in terms of perceived personalization. These were then combined as illustrated in Figure 4, resulting in 5 sample groups since combination D2, V2 was used for both experiments. In order to eliminate any possible compounding factors and to isolate cause and effect relationships, the main study questionnaires were the same for all 5 studies, only differing in terms of the manipulation (Lynn & Lynn, 2003). These results can subsequently be generalized towards the sample population.

Figure 4. Experiment Design

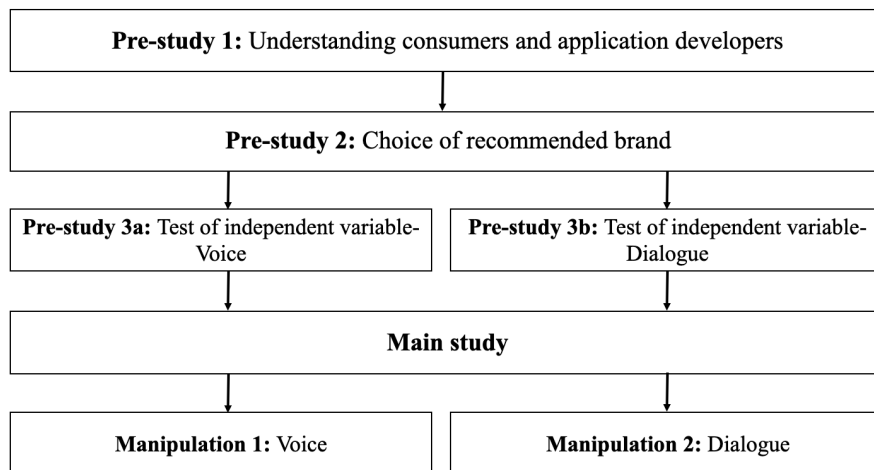
	Low	Medium	High
Manipulation 1: Voice	D2, V1	D2, V2	D2, V3
Manipulation 2: Dialogue	D1; V2	D2, V2	D3, V2

4.3 Preparatory work for the main study

Three pre-studies were conducted in order to prepare for the development of the main study. The pre-studies built on each other and helped pre-test the choice of independent and dependent variables for the main study.

The first pre-study consisted of one exploratory interview and one focus group. The purpose of our first pre-study was to grasp opinion and attitudes towards voice assistants of both consumers and industry task-developers. The second pre-study aimed to determine which brand to use in the main study, while the third assessed the internal validity of our chosen independent variables.

Figure 5. Research Design



4.4 Pre-study 1: Understanding consumers and application developers

As voice assistant is a field with limited academic research, it was necessary to understand what direction the study should take to benefit both the academic community and the industry in practice. Consequently, we made one interview with the innovation department of Sweden's largest grocery retailer (Nielsen et al., 2017) and one focus group to understand the consumer perspective of voice assistants. These qualitative interviews helped form the research questions for the main-study.

4.4.1 Pre-study 1.1: Task-application development at ICA

The first pre-study interview was conducted with a Product Manager at the Swedish food retailer ICA, working with the company's innovation team currently developing the retailer's application for Google Home (ICA interview, 2019).

ICA was one of the first companies in Sweden to release a native app within Google Home, meaning that the user is transferred to the ICA application through asking the device to speak with ICA. Their application was launched in August 2018, aligned with the launch of the Swedish speaking Google Home. Within the ICA native app, the company itself is responsible for all programming of the dialogues, how customer data is being used and collected as well as which voice the assistant will speak in.

The user can ask the assistant to create a purchase list which will then be transferred to their phone application and physical scanners used in-store, or make an e-commerce purchase out of the shopping list. By translating generic words such as "cheese" and "milk", the application can recommend specific products for the consumer based on previous purchases data or generally popular products.

ICA purposely designed the persona speaking in their task-application by giving her the name Monica and thoughtfully copywriting her conversational style. They aimed to making the dialogue personal and quirky by telling jokes and stories about Monica's background. *"We want people to create a relationship to Monica, it should be like having a friend in the kitchen"*. The dialogue is something they are continuously working on in order to make it more advanced and add new features. They have recently updated Monica's voice to the standard voice used by Google Home, as the previous voice was much more synthetic.

In the future they are considering enabling e-commerce purchases directly through the voice application. The Product Manager imagines that this could be done by having consumers pre-program standard purchases in an app, such as which brand of milk they usually buy. The main challenge would be the visual restriction of the application. It would also require them to build a new infrastructure for connecting purchase decisions with one's ICA card. The feasibility for releasing an e-commerce channel has recently expanded as Google has released the permission for applications to set up their own e-commerce shop, and not forcing them to funnel consumers to the Google shop. However, currently ICA is simply focusing on making consumers comfortable with using the voice assistant application.

During the discussion the insight emerged that Monica, in the future, might take on the role of a salesperson as well as that of a kitchen assistant. For example, by reminding customers about discounts on products they usually buy at their local ICA store. *"In the end our purpose is to push and sell products"* says the interviewee. To finish up our discussion, he emphasizes that most discussions about the ICA voice assistant are still in the idea stage. There has been little testing made around how consumers act in response to various design elements in their application. If voice assistants will be just as well received in Sweden as in the US is yet to see.

4.4.2 Pre-study 1.2: Consumer insights

4.4.2.1 Purpose and focus group

The purpose of this pre-study was to understand how consumers perceive and understand voice assistants, what uses they make of them and if using a voice assistant has an impact on their perception and more broadly relation to brands. The focus group, conducted on February 14th, 2019, gathered six participants, among which five of them had tried using a voice assistant, with three being regular users. The method of a focus group was chosen over individual in-depth interviews as the pre-studies had an explorative purpose, trying to uncover trends and general consumer opinions. Indeed, focus groups create a group dynamic that allows new topics to emerge as the discussion develops. (Bryman & Bell, 2015). The session ran for one hour and all participants were present voluntarily, without any economic or material benefit in exchange of their participation. It was audio-recorded and notes were taken. Both notes and recording were subsequently brought together to identify recurring themes in the conversation.

4.4.2.2 Result and conclusion

A number of topics stirred the conversation. First, a recurrent theme was data, with two sub-themes: data collection and data privacy. Data collection refers to how the voice assistant would

collect data points about you – habits, preferences, behaviors. Opinions here were divided: some accepted to have their personal data collected as a price to pay for the convenience of personalized services, while others were anxious about companies spying on them. “*You speak about something and then you get an ad about it*”, said one participant. This relates to the concept of surveillance, i.e the feeling that the device would be spying on you at a time when you don’t need nor desire its services. Participants referred to such behavior by mentioning that “[*the voice assistant*] is picking up a lot of stuff that’s not supposed to be recorded”, or “[*the voice assistant*] would light up out of the blue”. This means that data collection is acceptable, but only when there is a somewhat explicit agreement to do so, which takes place when the user is using a wake-up word like “Alexa” or “Hey Google”. This links to data privacy, the second sub-theme. Some participants talked about the desire to have their data stay within a device, and not have it owned by a company that could use the data for other services, such as online searches on other devices and more. Others were concerned about data security, like someone hacking the device or others in the home accessing one's private data.

Participants potential usage of the voice assistants was subsequently discussed. Answers revolved primary around one theme: convenience. Several participants would use a voice assistant while cooking, by asking for recipes and measurements, others as a smart home hub for controlling media devices, and some would use the voice assistant to perform tasks hands-free, such as texting or checking one’s schedule.

When it comes to the voice assistant in a shopping setting and advertisement context, participants were overall agreeing on appropriate uses and messages. Participants stated they would use the voice assistant for low-involvement purchase, such as groceries. Some talked about food and beverage products, while others mentioned personal hygiene products (toothbrush, soap). Regarding advertisement, participants agreed that it should not be pushed aggressively towards you – i.e the device suddenly lighting up to talk about a product or service deal – but rather be integrated to another conversation and be relevant in the conversational context.

4.4.2.3 Criticism

As the purpose of the focus group was purely exploratory, the focus group was only conducted once, which could question its external validity. Similarly, the participants were made up of the authors social network, resulting in a skewed convenient sample (Jacobsen, 2002). Lastly focus groups are inherently problematic as true opinions might not emerge due to the group dynamics (Reed, Roskel & Payton, 1997)

4.4.3 Conclusions from consumer and industry insights

The qualitative pre-studies generated insights about consumer usage of voice assistants and the process of developing the application software. Firstly, applications developers have a great deal of control in developing their own “persona” to which the user speaks with when using the company’s own voice application. This includes both the voice and the copywriting of the dialogues. Little internal A/B testing have been made and developers are unsure to what extent they should use data on the customer during the conversation. Data privacy is also a dilemma from the consumer perspective, as many are concerned about the degree to which the voice assistant appears to be collecting data. From the focus group, it appeared that the participants actively using a voice assistant were less concerned about data privacy compared to those who had not tried the product. Finally, FMCG products such as groceries and hygiene products appeared to be the most relevant product category to purchase through a voice assistant.

4.5 Pre-study 2: Choice of brand

4.5.1 Purpose and survey

The purpose of pre-study 2 was to determine which brand to advertise in the main study. As the manipulation is based on a hypothetical scenario, the aim was to ground the experiment in a realistic setting, hence an existing brand was used in the main study (Geuens & De Pelsmacker, 2017). Since the main study would be distributed to a worldwide sample, it was critical to choose a globally-recognized brand. It was also important to choose a brand which the sample group did not have greatly diverting brand attitudes towards. However, individual differences in brand attitudes could be expected to cancel out due to sample size (Söderlund, 2005).

Pre-study 1 and previous theory had found that consumers considered FMCG products as one of the product categories most relevant for purchase through a voice assistant (Smith, 2018). Based on the most popular FMCG brands worldwide (Kantar Worldpanel, 2017), food and beverage were excluded as they are more likely to be influenced by personal taste and the top five hygiene brands were tested instead: Pepsodent, Colgate, Lifebuoy, Dove, and Sunsilk.

The survey presented the brand name and included three statements related to brand familiarity; *“I am very familiar with the brand”*, *“I have strong preconceptions about this brand”* and *“I have previously interacted with this brand”*. Following they were asked to assess their general attitudes towards the brand on; *“Negative/Positive”*, *“Dislike/Like”* and *“Unfavorable/Favorable”* (Rosengren & Dahlén, 2015). Respondents answered on a seven-point Likert scale (ranging from *strongly disagree* to *strongly agree*). Such scale has been found to be

more suited to electronically-distributed surveys (Finstad, 2010). To avoid any appearance order bias, the order in which brands appeared was randomized.

The survey was designed in *Qualtrics* and distributed through *Amazon Turk* (commonly abbreviated *MTurk*). It was central to distribute the pre-study through *MTurk* for the pre-study results to be directly applicable for the sample used for the main-study. An attention test was included to ensure quality responses.

4.5.2 Result and conclusion

The data was analyzed in *SPSS 25*. The survey collected answers from 31 respondents dispatched around the world through *MTurk*. Prior to analysis, the questions were merged into an index called *Awareness* (Cronbach's Alpha, 0.91) and *Attitudes* (Cronbach's Alpha, 0.98). A one-sample t-test was performed, with the assumed test-value '4. From the results we found that Colgate had the highest mean in terms of awareness (6,236), Colgate also had the lowest std.deviation (0,9437) and highest t-value (13,425), indicating that most respondents had similar levels of awareness and that the results do not differ significantly from the general population. Similarly, Colgate had the highest mean for brand attitudes (6.387) and lowest standard deviation (0.898). For all these reasons Colgate was selected.

Table 2. Pre-study Choice of Brand - Awareness

Brand	N	μ	σ	T	p
Pepsodent	31	4.188	2.034	0.522	0.606
Colgate	31	6.236	0.944	13.425	0.000***
Lifebuoy	31	3.237	2.440	-1.743	0.092
Dove	31	6.054	1.152	9.928	0.000***
Sunsilk	31	3.086	2.480	-2.052	0.049*

Significance levels: *** $\leq .001$

Table 3. Pre-study Choice of Brand - Attitudes

Brand	N	μ	σ	T	p
Pepsodent	31	5.065	1.470	4.034	0.000***
Colgate	31	6.387	0.898	14.785	0.000***
Lifebuoy	31	4.644	1.541	2.289	0.030*
Dove	31	6.067	1.116	10.116	0.000***
Sunsilk	31	4.533	1.542	1.890	0.069

Significance levels: *** $\leq .001$, ** $\leq .01$, * $\leq .05$

4.5.3 Criticism of pre-study 2

One could criticize the use of *Amazon Turk* as a distribution method. Since respondents are paid for their responses, they might want to skim through our study as fast as possible without reading thoroughly our questions. To mitigate the risk of lack of attention during the survey (Jacobsen, 2002), we have included a control question around the middle of our survey and excluded all answers that did not respond correctly to it.

4.6 Pre-study 3: Test of independent variables

4.6.1 Purpose and survey

The third pre-study was conducted in order to assess the internal validity of the two independent variables in the main study. This pre-study is therefore divided into two main manipulations: levels of voice human-likeness, and levels of dialogue personalization. Each manipulation had three levels, ranging from *low* to *high*.

4.6.2 Manipulation 1: Levels of voice “human-likeness”

4.6.2.1 Presentation of manipulation 1

For testing the impact of different levels of human-likeness in the voice tonality of a voice assistant, three different voice levels were recorded. V1 was a low robotic voice recorded with the voice Victoria from the website onlinetonegenerator.com. V2 was recorded with the advance robotic voice in a Google Home assistant (Vincent, 2018). V3 was a conversation with a human impersonating a voice assistant, mimicking its speech style in terms of speed and neutral voice tone. As the main study was distributed through *MTurk*, there most users are American native speakers (Ross et al., 2010), all voices spoke in an American accent as to eliminate regional accent biases. Lastly, all voices were feminine and used the same dialogue to eliminate any further compounding factors. The audio-recording was a 1-minute-long interaction during which one person tried to purchase toothpaste through the voice assistant (appendix 9.5). One of the authors assumed the role of interacting with the voice assistant. The dialogue used had a medium level of personalization (see Figure 4) and was predominantly focused on speech from the voice assistant. For testing three different levels of human-likeness in the dialogues, the pre-study aimed to test the relationship $V1 < V2 < V3$ in terms of perceived human-likeness in the voice.

The recording was followed by three questions. One question was directly assessing the perceived level of human-likeness, while another measured if people forgot they were listening to a robot during the recording, and the last aimed to measure social presence in a sales setting, i.e if the conversation sounded like two humans talking during a sales encounter (appendix 9.3.1). A seven-point Likert scale was used ranging from *Strongly disagree* to *Strongly agree*. The survey was distributed in-person over a day in the hallway of the Stockholm School of Economics.

4.6.2.2 Result and conclusion

The data was analyzed in *SPSS 25*. 33 answers were collected in total, divided equally between the 3 manipulation groups, to which the participants had been randomly assigned. The three survey questions were made into an index variable with an acceptable Cronbach's Alpha of 0.796 (Söderlund, 2005). A one-way ANOVA test with a post-hoc Scheffe test was performed to assess the mean difference relationship between the results. Due to the low N, no all results were significant. However, the results showed mean differences that could confirm the relationship of $V1 < V2 < V3$ in terms of perceived human-likeness of the dialogue.

Table 4. Pre-study Independent Variable Voice

Manipulations	Mean difference	Std. error	Sig
V1-V2	-0.367	0.443	0.071
V2-V3	-0.515	0.450	0.047*
V3-V1	0.881	0.404	0.107

Significance levels: * $\leq .05$

4.6.3 Manipulation 2: Levels of dialogue personalization

4.6.3.1 Presentation of manipulation 2

The second independent variable that was pre-tested was the level of personalization in the dialogue. Three contextually similar dialogues were developed with levels of personalization deemed different. The pre-test aimed to assess whether a sample group perceived their level of personalization according to the relationship $D1 < D2 < D3$. The dialogues were made more personalized using social language (Hamman, 2006), by including vocabulary such as personal pronouns and friendly greeting phrases. They also exhibited knowledge about its user through referring to data inexplicitly collected about her (appendix 9.5). D1 had a low level of personalized dialogue and used very standardized replies. It did not use any stored data whatsoever on the user. The user had to repeatedly state her preferences throughout the order, and the voice assistant simply executed her will. D2 used some user data in the dialogue such as saved grocery list and location, and employed words such as “*Okay*”, “*Sure*” and the personal pronoun “*I*” to appear as more friendly to the user. D3 exhibited further human-like dialogues and emitted judgments such as “*good decision*” and “*great deal*” as well as used further user data like repeated behaviors and drew conclusions from it.

Recordings were approximately one minute-long, and respondents were subsequently asked to answer three questions. One question asked the respondent directly how personalized s/he - thought the dialogue was, another asked about how well the respondent felt the voice assistant

knew about him/her, and the third asked about perceived amount of data points about purchase behavior stored by the device (appendix 9.3.2). All questions could be answered to on a seven-point Likert scale ranging from *Strongly disagree* to *Strongly agree*. The survey was distributed in-person over a day in the hallway of the Stockholm School of Economics, through the online survey tool Qualtrics.

4.6.3.2 Result and conclusion

The data was analyzed in *SPSS 25*. 34 answers were collected in total, divided equally between the 3 manipulation groups, to which the participants had been randomly assigned. The three survey questions were made into an index variable with the acceptable Cronbach's Alpha of 0.787. A one-way ANOVA test with a post-hoc Scheffe test was performed to assess the mean difference relationship between the results. The results showed significant mean differences that could confirm the relationship of $D1 < D2 < D3$ in terms of perceived personalization of the dialogue.

Table 5. Pre-study Independent Variable Dialogue

Manipulations	Mean difference	Std. error	Sig
D1-D2	-1.622	0.581	0.032*
D2-D3	-0.700	0.607	0.052
D3-D1	2.322	0.581	0.002*

Significance levels: * $\leq .05$

4.6.4 Criticism of pre-study 3: voice and dialogue

One could criticize both pre-studies for the relatively low number of respondents to each sub-survey. However, the analysis has shown significant results and the number of respondents is therefore not affecting the quality of our conclusions. Another critique could be about the potential background homogeneity of the sample – due to material and time constraints, a convenience sample was used (Jacobsen, 2002). Finally, the internal validity of the study could be questioned. It is the subjective opinion of the authors that the questions asked are measuring level of personalization and dialog. However, the strong Cronbach's Alpha supports the coherency of the questions.

4.7 Main Study

4.7.1 Pilot test of questionnaire

The main study survey was piloted 8 times with an audience in order to correct any misunderstandings. The aim was to understand if any questions were misinterpreted, if the order

of the questions was appropriate and other misunderstandings that could skew the data. In alignment with Malhotra (2010), respondents were asked to verbalize their experience of taking the pilot survey. In this way we were able to understand how they interpreted each question. The design of the survey followed an iterative approach in which changes to the design were made based on the feedback after each pilot study. After 7 tests, saturation has been reached as the number of feedback comments were fewer and no amendments appeared necessary.

4.7.2 Main study questionnaire design

The questionnaire was designed with the survey tool *Qualtrics*. The questionnaire consisted out of 8 blocks of questions. Initially the respondents were introduced to a consent agreement in which the *MTurk* users agreed to comply with several conditions (appendix 9.6). Next, they were introduced to a scenario in which they are buying toothpaste through a voice assistant. They were then randomly assigned to listen to one of the five audio recordings and asked two control questions in order to check that they understood it.

The next section asked them to answer from the point of view of the person they had heard interacting with the voice assistant. The survey began with questions regarding anthropomorphism and social presence of the voice assistant. By putting these questions first, we hoped that they would enhance the priming effect of the recording by making the user reflect on the human-likeness of the voice assistant (Strack, 1992). Later questions followed regarding their perception of the voice assistant and questions regarding their perception of the brand.

Lastly, survey takers were asked to answer questions regarding their personal innovativeness with technology along with basic demographic variables.

Questions were primarily organized in a matrix structure in which the survey taker assessed a statement on a 7-point Likert scale. With this scale the respondent could be both natural and extreme in their answers, which facilitates when comparing the results for the analysis (Bryman, Bell, 2015).

4.7.3 Dependent variables

The dependent variables were assessed through several multi-item measures that were later index into each dependent variable. Internal consistency was found for each measure as the Cronbach's Alpha >0.7 (Söderlund, 2005). See appendix 9.6 for full questionnaire.

Table 6. Cronbach's Alpha Dependent Variables

DV	Cronbach's Alpha
Anthropomorphism	0.966
Social Presence	0.973
Trust	0.847
Service Quality	0.833
Service Satisfaction	0.899
Loyalty	0.833
Brand Attitudes	0.914
Purchase Intentions	0.967
Brand Loyalty	0.950
Personal Innovativeness	0.767

Anthropomorphism

A common way of measuring anthropomorphism of robotic agents is with the Godspeed questionnaire developed by Bartneck (2009). The questionnaire assesses the level of anthropomorphism directly, while also assessing the perceived intelligence, likeability and animacy of the robot. The questionnaire was introduced in the first block of our survey and consisted of 19 questions on a bi-polar 7-point Likert scale.

Social Presence

Social presence was measured based on literature defining the concept as a sense of human contact, personalness, warmth and sociability (Qiu & Benbasat, 2008). The respondent was presented with five statements regarding their experienced level of human contact, personal connection, human warmth, sociability and human sensitive. These statements were then assessed on a 7-point Likert scale ranging from *Strongly Disagree* to *Strongly agree*.

Trust

The measures of trust were based on Coelho's (2012), questionnaire on trust in retail service encounters. The questions were measured on a 7-point Likert scale in which the respondent assessed the strength to which they agreed with the following statements; "I trust the voice assistant", "I expect the voice assistant to be sincere and genuine", "I believe the voice assistant performs its role very well", "I believe the voice assistant acts in my best interest".

Service quality

Similarly, service quality took inspiration from measures by Coelho (2012), however the measures were adapted to be suitable for the context of a voice assistant interaction. The following measures were used: "The quality of the service provided by this voice assistant was

high”, “This voice assistant gave me valuable recommendations”, “This voice assistant was a reliable source of information”, “This voice assistant give me clear and transparent information”. The measures were made on a 7-point Likert scale in which the respondents assessed the strength to which they agreed.

Loyalty

As previously, the measures were inspired by Coelho (2012), The statements were formulated as: “I would use this voice assistant the next time I am searching for products” and “I would recommend using this voice assistant to friends and family if they were asking for advice”. The measures were made on a 7-point Likert scale in which the respondents assessed the strength to which they agreed.

Service Satisfaction

Our measurements took inspiration from Coelho (2012), measures of satisfaction of service encounters by asking “Overall I am satisfied with this voice assistant”, “This voice assistant fulfills my expectations” and “The interaction with this voice assistant is ideal”. The measures were made on a 7-point Likert scale in which the respondents assessed the extent to which they agreed.

Brand attitudes

To measure how brand attitudes were effected by each manipulation, we found inspiration in a three item developed by Spears et al. (2004) and Bruner (2016) on how to study the concept. On a seven-point Likert scale, the respondents answered to: “bad”/good”, “negative/positive”, and “dislike/like”.

Purchase Intentions

Purchase intentions were measured with two statements. The first one “How likely are you to buy Colgate Sensation White in the near future?” measured on a bi-polar 7-point scale, ranging from *Not likely* to *Likely* and *Not probable* to *Probable*. This measure was inspired by Rosengren (2015) interpretation of Grohmann, (2009) brand equity measures. Secondly the survey asked “I would buy this brand the next time I buy toothpaste” (Dodds, Monroe & Grewal, 1991) in which the respondents assessed their level of agreement with the response on a 7-point Likert scale.

Brand Loyalty

Inspiration was found from related studies in the area of brand equity and loyalty. The variable is measured through four questions. “I will buy this brand the next time I buy toothpaste” and “I would be willing to pay a higher price for this brand compared with other brands of toothpaste” were adapted from Rosengren (2015). The following questions “I would recommend this brand to

others” and “I would choose this brand over other brands selling toothpaste” were adapted from Aaker (1997) and Jacoby (1973). The items were measured on a 7-point Likert scale, ranging from *Strongly disagree* to *Strongly agree*.

Personal Innovativeness

Personal Innovativeness were measured using five statements inspired from Bloch (1981). The questions asked: “I am very positive towards using voice assistants in my everyday life”, “I am curious to try new technologies”, “I am curious about new technologies “, “I am usually the first to try out new technologies in my social circle”, “I have no concern for companies collecting my personal data”.

Demographics and Final Control Measures

To ensure a diverse and representative sample, demographic information was collected about the respondents by asking them to provide their age and gender.

4.7.4 Sampling approach

The survey was distributed online through *Amazon Mechanical Turk* as several studies have shown that the platform generates high quality answers and a more demographically diverse pool of respondents (Buhrmester et al. 2011; Paolacci et al. 2010). Five surveys were distributed between the 27-03-2019 and 08-04-2019. Out of the 523 answers collected, 40 answers from each group were included in the data sample due to strict data quality checks (see 4.8.7). All respondents were paid \$0.7 for each answer. The final sample consisted of 33% female and 67% male population, and the mean age was 38 years old. 57% lived in the US, 36% from India and 7% from other nations.

4.7.5 Analytical tool

The data was directly exported from the survey collection tool *Qualtrics* into *IMB SPSS 25* in which the statistical data analysis was performed, which avoided faulty manipulations by human error. Multi-scale measures were checked for internal reliability by having a Cronbach’s alpha > 0.7. Means comparison were made with one-way ANOVA test to compare means between three groups. Later a Scheffe post-hoc test was made as it is most suitable when number of participants are approximately similar in each group (Marascuilo & Levin, 1970). Mediating and moderating effects were tested through the add-on program Haye’s PROCESS tool for *SPSS*. Hypotheses were supported at a 95% confidence interval with significance of 5%. No results were deemed

partially supported as to avoid ambiguous contributions (Szucs & Ioannidis, 2017). All numbers were rounded up to three decimals.

4.7.6 Data quality check

As the survey was made through *MTurk*, rigorous quality checks were made in order to ensure appropriate answers. Only answers that completed the integrality of the survey were included. To ensure that all answers had been made in a thoughtful manner, all answers were reviewed individually according to five pre-communicated conditions. Firstly, all participants had to correctly answer control questions to make sure that they had listened to the audio-recording and understood our manipulation. Secondly, any rushed response was excluded, defined as answering below three minutes. Thirdly two attention checks were planted in the survey, in order to ensure that participants were not picking random answers. Fourthly we made sure that no one completed the survey twice. Lastly every answer was reviewed to ensure that the respondents had not responded in any obvious pattern. However, any significant outliers were not removed from the data set, as this could be argued to manipulate the data (Orr, Sackett & Dubois, 1991). Out of the 523 answers collected, 203 were used for the survey which resulted in an acceptance rate of 39% (Bryman & Bell, 2015). According to Söderlund (2010) more than 30 answers is needed in each group to motivate statistical tests, hence each experiment manipulation had around 40 answers.

4.8 Data quality

It is necessary to critically assess the quality of the data collected through looking at the reliability, validity and replicability of the study (Bryman & Bell, 2015).

4.8.1 Reliability

A reliable study should be able to generate the same results if repeated. If the study has been correctly conducted without any apparent measurement errors, it should generate reliable and accurate results. The aim has been to describe the methodological research process in an accurate and transparent way so that the reader of this thesis is able to assess the reliability of the study themselves (Saunders et al. 2012). Moreover Bryman and Bell (2015) claims that reliability can be assessed through measures of stability and internal reliability.

Internal reliability

For this study, the dependent variables were measured through multi-item measures that were merged into indexes. The internal reliability of a study assesses if the multi-item measures of the study are consistent and coherent measures of the intended variable (Bryman & Bell, 2015). To

ensure internal consistency, a multi-item variable was checked for Cronbach's Alpha ($>0,7$) (Söderlund, 2005). Moreover, all measures were inspired by previous research to ensure that sub-measures had previously been connected to the intended variable. To the further possible extent, literature was also checked for its citation impact. As a last step to ensure internal reliability the main study was pilot tested eight times in order to not create misinterpretations of the questions. Measures of human-likeness and personalization in the voice and dialogue were also pre-tested in order to validate the relationship between the manipulation levels.

Stability

Stability of the study measures that the results are reliable over fluctuating time frames and varying contexts. The fact that the experiment was only conducted once could negatively affect its stability (Bryman & Bell, 2015). For further research it could be useful to replicate the study in another time frame, on other survey respondents and by using other varieties of voices and dialogues. To still ensure the stability of responses, two pre-tests were made to ensure statistically significant reliability of the manipulations voice and dialogue. The relationship between the independent and dependent measures have been indicated by previous research, further ensuring the stability of this study.

4.8.2 Validity

Validity concerns the extent to which the collected data is measuring what the research is intended to study. In order to draw accurate conclusions about causal relationships, high validity is necessary. Validity consists of four main dimensions; internal validity, external validity, ecological validity and reliability (Bryman & Bell, 2015).

Internal validity

Internal validity refers to the accuracy of causal relationships that is concluded from the study. To minimize possible confounding effects the exact same study was used for all five experiments, with only the audio recording being different. All surveys were distributed through *Amazon Turk* in order to not risk that different contexts were affecting the reader. The survey was pilot-tested several times to make sure that no questions could be subjectively misinterpreted and would distort the data. Attention tests were used to check that all responses were mindful. Moreover, all hypotheses were backed up with literature that previously had indicated a causal relationship between the variables. Lastly, internal validity was improved by only using variables with a high Cronbach's alpha ($>0,7$) (Söderlund, 2005).

External validity

To ensure strong internal validity, some sacrifices to external validity had to be made. External validity refers to the generalizability of the result, the extent to which the conclusions can be applied to other contexts or sample groups (Saunders et al. 2012).

Firstly, the results are limited to a low-involvement purchase through a voice assistant. It is likely that users would have other expectation of, e.g service quality in another interaction setting. Secondly, only one brand and product category were researched, limiting the conclusion to these. However, subjective differences in initial brand attitudes can be expected to be cancelled out due to the sample size. Likewise, the choice of product category was backed up both by pre-test 2 and previous research (Kim et al., 2018), hence we believe this is a springboard for future research.

Moreover, the sample of survey respondents only consists of *Amazon Mechanical Turk* users, mainly from the US and India. However, the main benefits with *MTurk* is the demographic diversity of respondents (Paolacci et al., 2010). The sample consisted of 33%/66% in gender diversity, 18-63 age range and five different nationalities which would enhance generalizability of the results.

Ecological validity

By exposing participants to an experimental arrangement and using a survey questionnaire to collect data, one is putting the respondents into an unnatural setting and risk distorting the data (Bryman & Bell, 2015). For the sake of feasibility and due to the newness of the research area, some sacrifices on ecological validity had to be made in order to provide a foundation for future research. The most severe criticism of this study is the fact that the study is testing respondent's perception about the voice assistant in a hypothetical scenario, by listening to an audio recording of the interaction. To partly compensate for this, the study was based on a real brand and reinforced strong internal validity.

4.8.3 Replicability

A research study of high quality should be possible to replicate in order to confirm the finding through further research (Bryman & Bell, 2015). To ensure the replicability of this study the research process is thoroughly described both in terms of theory used and methodological approach. Moreover, all measures have been previous replicated in other studied, enhancing the ability to reproduce this study.

5. Analysis and Results

The following section introduces and analyzes the results of the main study. The results will be described in the order of the hypotheses and voice and dialogue manipulations will be presented in parallel.

5.1 Hypotheses testing

5.1.1 Anthropomorphism

According to the first hypothesis, survey takers should anthropomorphize the voice assistant to a greater extent as the voice becomes more human-like and the dialogue more personalized. A one-way ANOVA test was performed to compare the means between different levels of manipulations. The results showed significant ($p=0.000$) ascending means for the voice manipulation ($M_{V1} = 3.848 < M_{V2} = 4.792 < M_{V3} = 6.21$). However, no significance was found in the dialogue manipulation ($p=0.752$). Hence hypothesis 1b was rejected.

Table 7. Anthropomorphism Mean Voice

	N	μ	σ	P
V1	40	3.848	1.085	0.000***
V2	40	4.792	1.052	
V3	41	6.211	0.599	

Significance levels: *** $\leq .001$

Table 8. Anthropomorphism Mean Dialogue

	N	μ	σ	P
D1	40	4.851	1.052	0.752
D2	40	4.792	1.052	
D3	42	4.974	1.236	

Voice was further analyzed by performing a post-hoc Scheffe test in order to understand which of the three manipulations significantly differed from the mean. Between all manipulations showed a significant ($p=0.000$) mean difference, indicating that the greatest difference is between V2 and V3 ($M_{V2-V3} = -1.419$). Hence hypothesis 1a is supported.

Table 9. Anthropomorphism Scheffe test Voice

Manipulations	Mean difference	Std. error	P
V1-V2	-0.943	0.209	0.000***
V2-V3	-1.419	0.208	0.000***
V3-V1	2.363	0.208	0.000***

Significance levels: *** $\leq .001$

H1a. The higher degree of perceived human-likeness in the voice, the more the voice assistant will be anthropomorphized – SUPPORTED

H1b. The higher degree of personalization of the dialogue, the more the voice assistant will be anthropomorphized – NOT SUPPORTED

5.1.2 Social presence

The second hypothesis tested whether survey takers experienced an increasing degree of social presence as the voice became more human-like and the dialogue more personalized. Similarly, a one-way ANOVA test was performed to compare group means. The results showed significant ($p=0.000$) ascending means for Voice manipulation ($M_{V1} = 3.350 < M_{V2} = 4.870 < M_{V3} = 5.630$). Even though results from Dialogue showed an increasing mean difference between manipulation D1 and D3 ($D1=4.125 < D3=4.481$), these results showed no significance ($p= 0.22$) and hypothesis 2b is therefore rejected.

Table 10. Social Presence Mean Voice

	N	μ	σ	P
V1	40	3.350	1.785	
V2	40	3.870	1.716	0.000***
V3	41	5.630	1.265	

Significance levels: *** $\leq .001$

Table 11. Social Presence Mean Voice

	N	μ	σ	P
D1	40	4.125	1.562	
D2	40	3.870	1.716	0.220
D3	42	4.481	1.485	

Looking further at the mean differences between Voice through a post-hoc Scheffe test, a significant ($p=0.000$) mean difference was found between manipulation V2-V3 and V3-V1. Even though no significance was found between V1-V2 ($p=0.319$), the mean difference was still aligned with the hypothesis ($M_{V1-V2} = -0.52$). Most importantly, significance was found between extreme values V3-V1 ($M_{V3-V1} = 2.279$, $p=0.00$), hence hypothesis 2a is supported.

Table 12. Social Presence Scheffe test Voice

Manipulations	Mean difference	Std. error	P
V1-V2	-1.520	0.358	0.319
V2-V3	-0.759	0.356	0.000***
V3-V1	2.279	0.356	0.000***

Significance levels: *** $\leq .001$

H2a. The higher the degree of perceived human-likeness in the voice, the more social presence will be experienced in interaction with the voice assistant - SUPPORTED

H2b. The higher the degree of personalization of the dialogue, the more social presence will be experienced in interaction with the voice assistant – NOT SUPPORTED

5.1.3 Anthropomorphism as a mediator for social presence

Hypothesis 3 tests whether the variable *Anthropomorphism* can function as a mediator for the variable *Social Presence*. To test the relationship, a mediation analysis was performed using Haye's macro Process for SPSS. The test used mediation model 4, a 95% confidence interval with a bootstrap sample of 5000 to receive better data representation and avoid non-normality (Preacher & Hayes, 2008). Results are presented according with Zhao et al. (2010). No significance was found for direct effects, path C. However, the results showed significance for indirect effects, path AB, for the Voice manipulation because no confidence interval (BootLLCI-BOOTULCI) crossed zero. No significance was found for Dialogue manipulation. Hence hypothesis 3a is supported and hypothesis 3b rejected.

Figure 6. Mediation Anthropomorphism

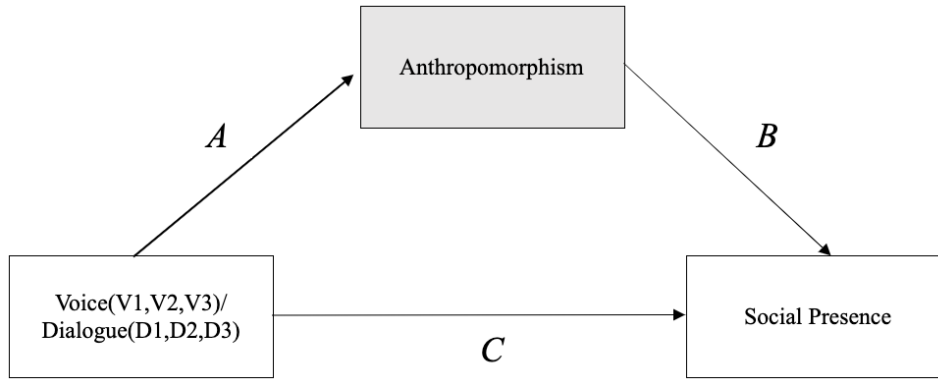


Table 13. Mediation Anthropomorphism

IV.	DV.	Mediator	Direct Effects		Indirect effects	
Voice	Social Presence	Anthropomorphism	C	-0.271	AB	1.413
			SE	0.189	BootSE	0.150
			T	-1.434	BootLLCI	1.130
			P	0.154	BootULCI	1.718
Dialogue	Social Presence	Anthropomorphism	C	-0.035	AB	0.097
			SE	0.079	BootSE	0.097
			T	-0.435	BootLLCI	-0.081
			p	0.664	BootULCI	0.274

Bootstrap sample = 5000, 95% confidence interval

H3. Anthropomorphism will function as a mediator for social presences in terms of:

- Human-likeness in the voice – SUPPORTED
- Personalization in the dialogue – NOT SUPPORTED

5.1.4 Attitude towards the voice assistant

Hypothesis 4 tests the effect of manipulation on the users' attitudes towards the voice assistant, measured through the dependent variables *Trust*, *Customer Loyalty*, *Service Quality* and *Customer Satisfaction*. These were indexed into one variable called Attitude towards Voice Assistant. The variable had a Cronbach's Alpha of 0.865 indicating internal consistency. Similarly, significant mean differences were found for Voice ($p=0.00$) with increasing means ($M_{V1} = 5.206 < M_{V2} = 5.499 < M_{V3} = 5.993$). No significance was found for Dialogue ($p=0.402$). Hypothesis 4b is hence rejected.

Table 14. Mean Attitudes VA Voices

	N	μ	σ	p
V1	40	5.206	0.832	0.000***
V2	40	5.499	0.938	0.000***
V3	41	5.993	0.826	0.000***

Significance levels: *** $\leq .001$

Table 15. Mean Attitudes VA Dialogue

	N	μ	σ	p
D1	40	5.293	0.806	0.806
D2	40	5.531	0.944	0.944
D3	41	5.256	1,170	0.870

Similarly, when looking at the mean differences through a post-hoc Scheffe test, although no significant results are found between V1 and V2 ($p=0.323$) all differences involving manipulation V3 show significant mean differences ($p=0.000$).

Table 16. Attitudes VA Scheffe test Voices

Manipulations	Mean difference	Std. error	Sig
V1-V2	-0292	0.193	0.323
V2-V3	-0.494	0.192	0.000***
V3-V1	0.787	0.192	0.000***

Significance levels: *** $\leq .001$

Looking more closely at each of the individual dependent variables, significance is found for Voice but not for Dialogue. All dependent variables for Voice indicate significant results and a relationship of means ($M_{V1} < M_{V2} < M_{V3}$) aligned with hypothesis 4a. Hence hypothesis 4a is supported for all dependent variables.

Table 17. Dependent Variables Attitudes VA Voice

DV.	IV.	N	μ	σ	p	IV.	N	μ	σ	p
Trust	V1	40	5.201	0.843	0.016*	D1	40	5.131	1.030	0.146
	V2	40	5.423	0.877		D2	40	5.544	0.837	
	V3	41	5.805	1.077		D3	42	5.147	1.253	
Service Quality	V1	40	5.463	0.850	0.010**	D1	40	5.457	0.916	0.222
	V2	40	5.732	0.910		D2	40	5.759	0.881	
	V3	41	6.136	0.833		D3	42	5.414	1.233	
Customer Loyalty	V1	40	5.136	1.148	0.010**	D1	40	5.390	0.990	0.765
	V2	40	5.417	1.382		D2	40	5.415	1.382	
	V3	41	5.944	1.074		D3	42	5.237	1.349	
Customer Satisfaction	V1	40	5.044	1.283	0.000***	D1	40	5.233	1.074	0.746
	V2	40	5.427	1.243		D2	40	5.431	1.243	
	V3	41	6.108	0.844		D3	42	5.319	1.188	

Significance levels: *** $\leq .001$, ** $\leq .01$, * $\leq .05$

H4a. A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d. - SUPPORTED

H4b. A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d. - NOT SUPPORTED

- a. Trust
- b. Service Quality
- c. Customer Loyalty
- d. Customer Satisfaction

5.1.5 Social presence mediates the attitude towards the voice assistant

According to hypothesis 5, *Social Presence* should act as an explanatory variable – mediator – for the *Attitude Towards the Voice Assistant*. As no relationship between dialogues and attitudes towards the voice assistant could be found in hypothesis 4b, only Voice manipulations are tested for mediating effects. Haye's bootstrapping macro Process for *SPSS*, model 4, was used for the mediation analysis. No significance was found for direct effects, path C. However, the results showed significance for indirect effects as the confidence interval between BootLLCI to BootULCI did not cross zero for each of the independent variables (Zhao et al., 2010). This indicates that *Social Presence* acts as an indirect explanatory variable for *Attitude Towards the Voice Assistant*. Hence hypothesis 5a is supported.

Figure 7. Mediation Social Presence

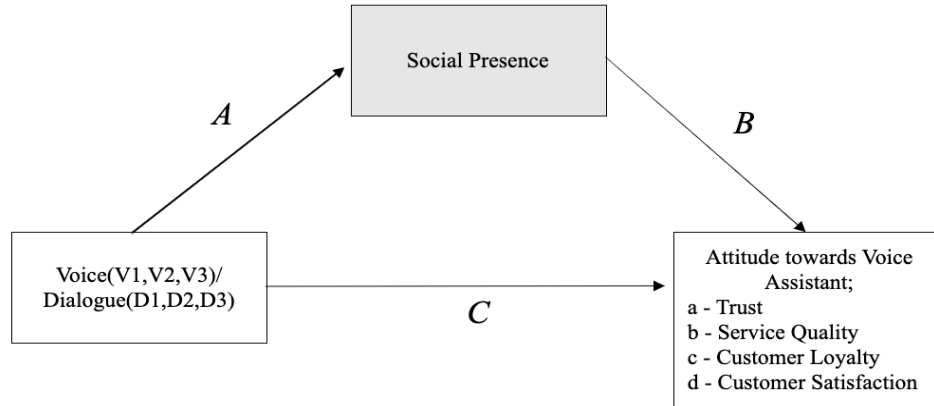


Table 18. Mediation Social Presence

IV.	DV.	Mediator	Direct Effects		Indirect effects	
Voice	Attitude Towards VA	Social Presence	C	0.057	AB	0.337
			SE	0.093	BootSE	0.070
			T	0.613	BootLLCI	0.212
			p	0.541	BootULCI	0.488
Voice	Trust	Social Presence	C	0.049	AB	0.352
			SE	0.102	BootSE	0.073
			T	-0.483	BootLLCI	0.221
			p	0.629	BootULCI	0.510
Voice	Service Quality	Social Presence	C	0.191	AB	0.145
			SE	0.108	BootSE	0.052
			T	1.768	BootLLCI	0.051
			p	0.079	BootULCI	0.254
Voice	Customer Loyalty	Social Presence	C	0.033	AB	0.441
			SE	0.133	BootSE	0.106
			T	-0.246	BootLLCI	0.250
			p	0.803	BootULCI	0.664
Voice	Customer Satisfaction	Social Presence	C	-0.033	AB	0.440
			SE	0.1323	BootSE	0.107
			T	-0.249	BootLLCI	0.253
			p	0.803	BootULCI	0.677

Bootstrap sample = 5000, 95% confidence interval

H5. Social Presence will function as a mediator for Attitudes towards the voice assistant in terms of:

- a. Human-likeness in the voice – SUPPORTED
- b. Personalization in the dialogue – NOT SUPPORTED

5.1.6 Personal innovativeness moderates the perception of the voice assistant

Hypothesis 6 tests if the variable *Personal Innovativeness* will act as a moderator for the indexed variable *Attitudes Towards the Voice Assistant* and following sub-variables, illustrated in H5. A moderation analysis using Haye's bootstrapping macro Process for *SPSS*, with model 1, was made on all dependent variables of *Attitudes towards the Voice Assistant*. A bootstrap sample of 5000 and a 95% confidence interval were used. However, hypothesis 5 must be rejected as no significance could be found, meaning that personal innovativeness cannot be deemed to influence the respondent's attitude towards the voice assistant.

Figure 8. Moderation Personal Innovativeness

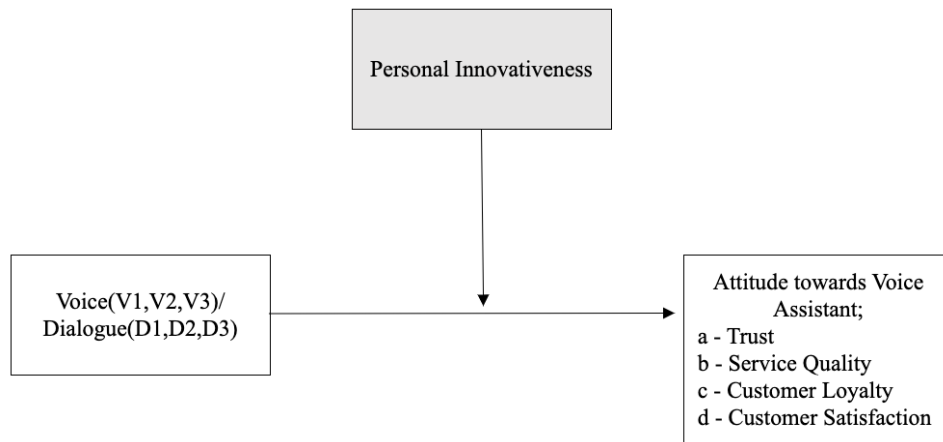


Table 19. Moderation Personal Innovativeness

Dependent variable	Interaction coefficient	p
Attitude towards VA (Indexed)	0.099	0.203
Trust	0.109	0,182
Service Quality	0.126	0,120
Customer Loyalty	0.071	0,522
Customer Satisfaction	0.916	0.392

H6. Personal innovativeness will function as a moderator for Attitudes towards the voice assistant in terms of:

- a. Human-likeness in the voice – NOT SUPPORTED
- b. Personalization in the dialogue – NOT SUPPORTED

5.1.7 Attitudes towards voice assistant mediate the relationship with brand attitudes

Lastly, hypothesis 7 tested whether the variable *Attitudes Towards the Voice Assistant* could act as a mediating explanatory variable for survey takers' *Attitudes Towards the Brand*. Haye's macro Process was used with a bootstrapping sample of 5000 and a confidence interval of 95%. No significance was found for the Dialogue manipulation. Yet all dependent variables showed significance for the Voice manipulation. As the confidence interval (BootLLCI-BootULCI) never crossed zero, it can be concluded that these variables are mediated by *Attitudes Towards the Voice Assistant*. Hence, *Brand Attitudes*, *Brand Purchase intentions* and *Brand Loyalty* can all be partially explained by attitudes towards the voice assistant when the voice of the machine is being manipulated. Hypothesis H8a is therefore supported and H8b is rejected.

Figure 9. Mediation Attitudes towards VA

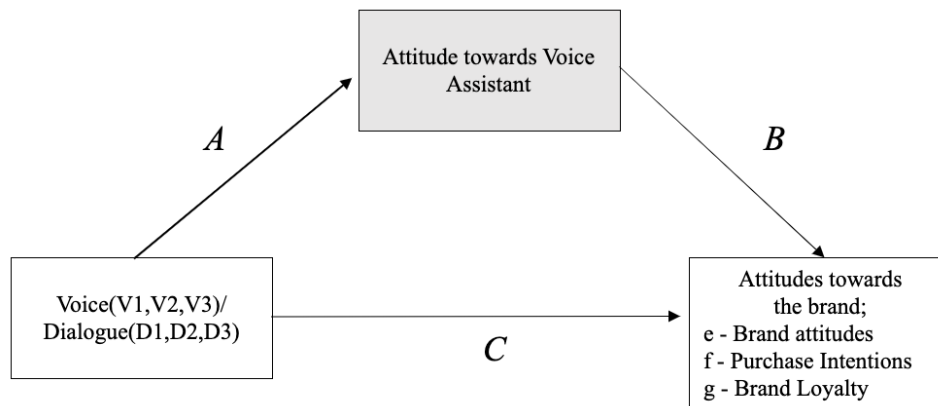


Table 20. Mediation Attitudes towards VA Voice

IV.	DV.	Mediator		Direct effects	Indirect effects	
Voice	Brand Attitudes	Attitudes towards VA	C	-0.160	AB	0.224
			SE	0.119	BootSE	0.067
			T	-1.339	BootLLCI	0.103
			p	0.183	BootULCI	0.367
Voice	Brand Purchase Intentions	Attitudes towards VA	C	-0.147	AB	0.341
			SE	0.192	BootSE	0.118
			T	-0.762	BootLLCI	0.150
			p	0.447	BootULCI	0.604
Voice	Brand Loyalty	Attitudes towards VA	C	-0.015	AB	0.248
			SE	0.179	BootSE	0.094
			T	-0.083	BootLLCI	0.116
			p	0.934	BootULCI	0.484

Bootstrap sample = 5000, 95% confidence interval

Table 21. Mediation Attitudes towards VA Dialogue

IV.	DV.	Mediator		Direct effects	Indirect effects	
Dialogue	Brand Attitudes	Attitudes towards VA	C	-0.117	AB	-0,044
			SE	0.130	BootSE	0.068
			T	-0.903	BootLLCI	-0.144
			p	0.369	BootULCI	0.120
Dialogue	Brand Purchase Intentions	Attitudes towards VA	C	-0.125	AB	-0.005
			SE	0.158	BootSE	0.082
			T	-0.789	BootLLCI	-0.178
			p	0.432	BootULCI	0.148
Dialogue	Brand Loyalty	Attitudes towards VA	C	0.037	AB	-0.005
			SE	0.1490	BootSE	0.068
			T	0.251	BootLLCI	-0.148
			p	0.802	BootULCI	0.123

Bootstrap sample = 5000, 95% confidence interval

H7. Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of:

- a. Human-likeness in the voice – SUPPORTED
- b. Personalization in the dialogue – NOT SUPPORTED

5.1.8 Attitudes towards the brand

Hypothesis 8 aimed to understand if the Voice and Dialogue manipulation influenced respondents' attitude towards the brand advertised in the interaction, measured through dependent variable e-g. A one-way ANOVA to measure differences in means between the manipulation levels was performed. For the Voice manipulation, significance was only found for the variable *Brand Attitudes* ($p=0.037$).

Table 22. Attitudes Towards the Brand Means Voice

IV.	DV.	N	μ	σ	p
e. Brand Attitudes	V1	40	5.608	1.102	0.037*
	V2	40	5.133	1.210	
	V3	41	5.732	1,039	
f. Brand Purchase Intentions	V1	40	4.625	1.771	0.470
	V2	40	4.563	1.784	
	V3	41	5.013	1.791	
g. Brand Loyalty	V1	40	4.019	1.466	0.211
	V2	40	3.963	1.571	
	V3	41	4.543	1.810	

Significance levels: * $\leq .05$

The variable was further analyzed through a post-hoc Scheffe test to understand significance for mean differences. Interestingly, there was a strong non-significance in the results between V3 and V2, however between V1 and V3 the results were found significant ($p=0,037$). As the null-hypothesis can be rejected between these two levels, the data indicates that there is a pattern which cannot be explained by coincidence. Hence hypothesis 8a will be considered supported for dependent variable e.

Table 23. Brand Attitudes Scheffe test Voice

Manipulations	Mean difference	Std. error	p
V1-V2	0.475	0.243	0.155
V2-V3	-0.598	0.243	0.879
V3-V1	0.123	0.243	0.048*

Significance levels: * $\leq .05$

As for testing Dialogue manipulations, significance was only found for Brand Loyalty ($p=0,030$). When further analyzed through a post-hoc Scheffe test (Table 23), significance was similarly only be found between D1 and D3, yet this was deemed sufficient to regard the hypothesis as supported.

Table 24. Brand Attitudes Means Dialogue

IV.	DV.	N	μ	σ	p
e. Brand Attitudes	D1	40	5.593	1.121	0.222
	D2	40	5.139	0.881	
	D3	42	5.361	1.233	
f. Brand Purchase Intentions	D1	40	5.332	1.279	0.092
	D2	40	4.563	1.784	
	D3	42	5.027	1.569	
g. Brand Loyalty	D1	40	4.562	1.200	0.030*
	D2	40	3.962	1.572	
	D3	42	4.720	1.437	

Significance levels: * $\leq .05$

Table 25. Brand Loyalty Scheffe test Dialogue

Manipulations	Mean difference	Std. error	P
D1-D2	0.700	0.315	0.072
D2-D3	-0.758	0.311	0.098
D3-D1	0.157	0.311	0.044*

Significance levels: * $\leq .05$

H8a. A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.

H8b. A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.

e. Brand Attitudes – SUPPORTED for H8a

f. Purchase Intentions – NOT SUPPORTED

g. Brand Loyalty – SUPPORTED FOR H8b

5.2 Summary of results

Figure 10. Summary Hypotheses Results

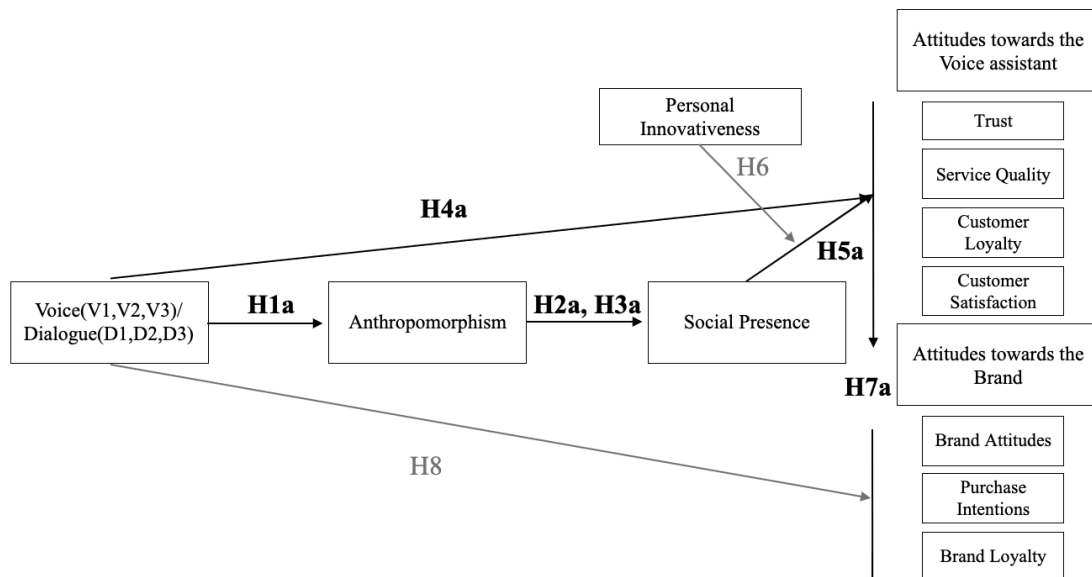


Table 26. Summary Hypotheses Results

H1	H1a	The higher the degree of perceived human-likeness in the voice, the more the voice assistant will be anthropomorphized	Supported
	H1b	The higher the degree of personalization of the dialogue, the more the voice assistant will be anthropomorphized	Rejected
H2	H2a	The higher the degree of perceived human-likeness in the voice, the more social presence will be experienced in interaction with the voice assistant	Supported
	H2b	The higher the degree of personalization of the dialogue, the more social presence will be experienced in interaction with the voice assistant	Rejected
H3	H3a	Anthropomorphism will function as a mediator for social presences in terms of human-likeness in voice	Supported
	H3b	Anthropomorphism will function as a mediator for social presences in terms of personalization in dialogue	Rejected
H4	H4a	A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.	Supported
	H4b	A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the voice assistant, in terms of dependent variable a-d.	Rejected
H5	H5a	Social Presence will function as a mediator for Attitudes towards the voice assistant in terms of human-likeness in the voice	Supported
	H5b	Social Presence will function as a mediator for Attitudes towards the voice assistant in terms of personalization in the dialogue	Rejected
H6	H6a	Personal innovativeness will function as a moderator for Attitudes towards the voice assistant in terms of human-likeness in the voice	Rejected
	H6b	Personal innovativeness will function as a moderator for Attitudes towards the voice assistant in terms of personalization in the dialogue	Rejected
H7	H7a	Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of human-likeness in the voice	Supported
	H7b	Attitudes towards the voice assistant will function as a mediator for attitudes towards the brand in terms of personalization in the dialogue	Rejected
H8	H8a	A higher level of perceived human-likeness in the voice will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.	e. Supported
			f. Rejected
			g. Rejected
	H8b	A higher level of perceived personalization in the dialogue will lead to more positive attitudes towards the brand, in terms of dependent variable e-g.	e. Rejected
			f. Rejected
			g. Supported

6. Discussion

The following chapter discusses the results generated by the study. The section begins with discussing the results regarding interaction with the voice assistants in terms of anthropomorphism and social presence. Secondly results relating to attitudes towards the voice assistant will be discussed. Lastly results dealing with attitudes towards the brand are examined.

6.1 Interacting with a voice assistant or an assistant with a voice?

As illustrated in H1 and H2, there is a clear link between human-likeness of the voice and experienced anthropomorphism and social presence. These findings were expected as they align with previous research on both concepts (Walters et al., 2008; Sims et al., 2009; Nowak & Biocca, 2003). Interestingly, our comparison of the means showed a somewhat greater difference in means between the most human-like robotic voice, V2 and the human voice, V3. Similar results have been found in recent voice assistant studies (Chérif & Lemoine, 2019), suggesting that human voices are far superior when in generating anthropomorphic perceptions and trust.

However, the human voice (V3) was not vastly superior in creating social presence compared to synthesized voices (V1 and V2). Such results could be explained with the fact that our human voice recording tried to mimic a robotic voice. The voice thus eliminated any signs of specifically human speech features such as varying voice pitches, taking breaks in one's sentence, hummed sounds and using various tones of voice during the same interaction. These features have been researched for decades in order to be integrated into synthesized speech (Bänziger & Scherer, 2005; Schröder, 2001; Schuller et al., 2003) and have been suggested create a sense of social presence (Lubold et al., 2016).

For different levels of dialogue personalization, no effect could be found on neither anthropomorphizing the device nor the effect on perceived social presence. This goes against the idea of Zlotowski, et al. (2014) that robotic agents displaying human-nature traits such as friendliness and sociability did not actually make it more likely to be anthropomorphized. Similarly, research states that e-commerce websites will benefit from enhancing social presence through using social language (Hamman, 2006; Gefen & Straub, 2004a), however our findings question the effectiveness of these communications methods.

No significance could be found for personalizing the dialogue, potentially explained by the experimental context. Our experiment simulated a dialogue taking place between a user ordering toothpaste, and a voice assistant. In such situations, some consumers may be very task-oriented,

and a more or less personalized dialogue will not influence their perception of the encounter, while others will be greatly influenced by different levels of personalization. In contrast to measuring the effects of human-likeness in voice, which is a very one-dimensional measure ranging from synthetic to human-like, measuring the effects of dialogue personalization is more difficult as it is multi-dimensional and more easily judged on different scales by different individuals (Abidi, 2002). In other terms, what might be very personalized for one person might not feel like it is for another. The very nature of the dimension personalization can help explain why no significance was found for the measure. A further reason for data insignificance could come from the data sample. According to Söderlund (2005), the number of respondents was enough to get significant results. However, another sample of respondents might have given other results leading to significance.

Eventually, a few words about anthropomorphism as a mediator for social presence. In line with the existing academic literature (Gong, 2008; Chérif & Lemoine, 2019), in the voice data set, the fact that people assigned human traits to the voice assistant explained in part why they perceived a social connection with the device. Since the dialogue personalization showed no significant results for anthropomorphism variable nor for social presence, no significant relationship could be found either for testing anthropomorphism as a mediator for social presence.

6.2 Effects of voice assistants as a service encounter

One of the most significant finding of this study was how H4 established that the level of human-likeness in the voice appears to affect users' attitudes, perceptions and feelings towards the device. The study was able to confirm that by making the voice assistant sound more like a human it can significantly improve the users' feelings of *trust*, *loyalty* and perception of *service quality* and *satisfaction* with the device. In particular, *service quality* showed the highest effect on means differences compared with the other dependent variables. This indicated that improving the voice quality will have the greatest effect on customer's perceived quality of the interaction with the VA.

The fact that using more advanced synthesized voices leads to better attitudes aligns with the findings of Sims et al. (2009), stating that less synthesized voices made users extend rules of human interaction towards the robot. Similar findings have been made by Nass et al. (2001) and indicated that the more a robot sounds like the user, the better it will be rated in terms of trustworthiness. Improved attitudes can be found at every stage of improvement between the three levels. However, that is not to say that the more human-likeness the better. According to Mori

(1970) theory of the Uncanny Valley, increased likability will hold only until a certain threshold, after that more human resemblance will generate feelings of eeriness and repulsion.

A possible explanation for the improved attitudes if found through H5, that social presence act as a mediating variable for attitudes of the VA. This explanatory relationship relates to previous research where human-likeness in voices have been linked to social presence (Chérif & Lemoine, 2019), a concept related to improve attitudes in retail settings (Gefen & Straub, 2004b; Cyr et al., 2007). Likewise, this confirms how Nass et al. (2001) established that the voice affects the perceived personality of the voice assistant, and how Purington et al. (2017) found that users are more satisfied with their voice assistant the more they personify it. Consequently, creating a sense of social connectedness between the service provider and the customer is essential, no matter if the service provider is a robot or a human being.

In contrast to these findings, no significant results were found when manipulating the words used in the dialogue. It appears that the extent to which the voice assistant exhibits knowledge about its user during the conversation has no effect on neither *trust*, *loyalty* nor *perception of quality* and *satisfaction*. Even tough generic measures of personalization were being used, such as friendliness and use of social language, the study was not able to create a sense of personalization on an individual level. These insights goes against much research on service encounters, which states that a more personalized interaction should enhance perceptions of the service provider (Surprenant & Solomon, 1987; Coelho & Henseler, 2012). Instead this study established how service encounter theory applicable to brick-and-mortar retail is not directly transferable to the context of sales through a voice assistant.

Most likely the lack of significant results is due to the idiosyncrasies of the study. The results for voice are significant as it is a more one-dimensional measure and that most people can be expected to prefer a clear, understandable and consequently human-like voice (Yang, 2001; Nass & Brave, 2005). However, personalization is a more complex measure. This experiment simulated a dialogue taking place between a user trying to order toothpaste and a voice assistant. In this situation, participants' perception of an ideal dialogue might diverge. Some might prefer a simple, fast and transaction-focused dialogue whereas others could appreciate a social language with personalized remarks. Recent research suggests in fact that voice assistant users conceptualize a VA-based conversation as nearly purely transactional, meaning that they do not desire to become friends with, or even converse with a VA the same way they would with another human, therefore not seeking to develop a social bond nor getting personalized content (Clark et al., 2019). Moreover, as found in pre-study 2, as well as by Lau et al. (2018) and Chung et al.

(2017), many find target advertising and microphone-based technology intrusive on their personal data. As preferences can vary greatly between respondents, it is reasonable that no significance was found in the results. Hence, improving the voice of the assistant is a more effective way of enhancing attitudes, since opinions are more coherent in that most prefer a clearer and less synthesized voice.

Likewise, in H6 personal innovativeness did not appear to moderate the relationship for attitudes towards the voice assistant, for any dependent variable in neither voice nor dialogue. This is a surprising insight that goes against much theory on technology adoption (Lu et al., 2005; Bloch, 1981). Lau et al. (2018) especially found that self-perception of being an early adopter was a primary factor for smart speaker adoption. However, once users have adopted the technology, being an early adopter does not appear to moderate their opinions about voice assistants as a service encounter. The conclusion would be that all consumers, regardless of their self-reported level of personal innovativeness, will benefit from a more human-like voices.

6.3 Brand related effects of voice assistant interaction

The most impactful brand related findings of this study are related to H7. The analysis establishes how attitudes towards the voice assistants have an explanatory, mediating, effect on the brands being advertised through the platform, at least with regards to the voice manipulations. This is in-line with theory on advertising context and suggests that affection towards the VA will spill over to the brands being advertised (Moorman et al., 2012). Research on advertising context has mainly focused on measuring attitudes towards the ad itself (Pelsmacker et al., 2002), whereas in this study the affection towards the medium context appears to spill over to the brand promoted through the ad. These findings are accurate for all three dependent variables, *brand attitudes*, *brand loyalty* and *purchase intentions*. However, with regards to the diversity of attitudes in the dialogue manipulation, no significance could be found here.

When looking at the direct impact of the two manipulations on the attitudes towards the brand, the effect is weak. As significance can only be found for one out of three dependent variables for both manipulations, it is difficult to make overall conclusions. The direct impact on brands from both personalization in service encounter and especially human-likeness of voice has been scarcely researched. Research justifying these relationships was based on a mediating relationship with variables such as customer satisfaction and loyalty (Lemmink & Bloemer, 1992; Bloemer & Kasper, 1995; Björkman & Kock, 1995). This further explains the mediating relationship found in

H5. Moreover, this study hence provides a first insight into the direct relationship between these two variables and their direct effect on brands.

When looking at voice manipulation, an effect was found for *brand attitudes*. This aligns with theory stating how satisfaction with the sales encounter spills over to brand attitudes (Brexendorf et al., 2010a; Cronin & Taylor, 1992). A possible explanation for why dialogue was not affected could be how social presence appears to mediate the relationship with VA attitudes. Even if the dialogue becomes more personalized, it is not the personalization in itself, but rather the social presence created through the personalization that spills over towards positive brand attitudes. A synthetic robotic voice will struggle more to create social presence compared to a human voice, no matter the dialogue personalisation level, as the words used does not make the voice assistant more anthropomorphised. However, there was an effect on *brand loyalty* in dialogue, relating to the idea of how personalization of service encounter will result in behavioural loyalty (Ball et al., 2006).

Purchase intentions were not affected by any of the variables, possibly due to the hypothetical scenario of the experiment. Previous research (Gefen & Straub, 2004a; Li, Daugherty & Biocca, 2002) has indicated that social presence influences purchase intentions in a retail setting and since manipulation of the dialogue showed no difference in social presence, this could potentially explain why purchase intentions were not affected. As Moriuchi (2019) found, users mostly make habitual purchases through voice assistants, the importance of forming purchase intentions before interacting with the brand on the voice assistant becomes apparent when no significance was found for this dependent variable.

In conclusion, when relating these findings to the pre-study with ICA and their effort to copywrite the dialogue of their task assistant in a unique and entertaining way, one can question the effectiveness of these efforts. To use a more human-like voice would clearly show much greater effects.

7. Conclusion

The following section makes conclusions with regards to the initial research questions. Next the chapter elaborates on potential contributions to theory and practice. Following this, criticism regarding the study will be raised and lastly suggestions for further research is suggested.

7.1 Conclusion

Will the level of human-likeness in the voice and the level of personalization in the dialogue of a voice assistant affect users' perception of the device and the brands advertised through it?

For different level of human-likeness the conclusion is a strong yes, human-likeness has a significant and positive impact on user's attitudes towards the voice assistant. In particular, the perceived service quality is affected by a more human-like voice. For effects on the brand being advertised, there is ambiguity. It appears that general brand attitudes are positively affected, but no other dependent variables of this study were impacted. Hence this is a topic of interest for further research.

For different levels of personalization in the dialogue, the answer is a strong no. No significant effect whatsoever was found on attitudes towards the voice assistant. Similarly, there is ambiguity with regards to brands. *Brand Loyalty* was positively affected, yet no effect was found on any other brand related variables. Likewise, this is a relevant field for further research.

What explains the different levels in perception of the VA and the brands advertised through it?

By manipulating the voice, this study has found an evident explanatory path. Throughout the different voice levels, the user tends to anthropomorphize the device more, which in turn affects the social presence experienced in the duality of the relationship with the device. This experience of presence has shown to effect user's experience of *service quality*, *loyalty*, *trust* and *satisfaction* with the device. Finally, attitudes towards the voice assistants appear to mediate the relationship towards the brand that the device is recommending. Conclusively, the extent to which the user assigns human characteristics to the device will indirectly affect their opinions about the brand the device is recommending.

No explanatory factors or relationships were found for dialogue, likely due to the ambiguity of dialogue as a concept.

Will customers reported interest and familiarity with technology effect how they perceive the interaction with the device?

The answer here is coherent for both studies, no. No relationship could be found for reported levels of personal innovativeness and the effects of the manipulations appear to have the same effect on all groups no matter their reported interest and familiarity with new technology.

In summary it appears that manipulating the voice has a much greater effect on how consumers experience the interaction with the voice assistant. For task-developer to spend a great effort of copywriting the dialogs and interpreting past data on consumers might not be as effective as just developing a more advanced human-like voice. This goes for all types of consumers, no matter their familiar and interest in new technology. Although more advanced voices do not have little direct effect on brands, it appears as if the general opinion about the voice assistant as a context for advertising influences brand opinions.

7.2 Contribution to theory

This thesis generates important contributions both to the broader, nascent field of voice assistant research, and to the narrower field of conversational commerce studies. Its originality and value lies in the fact that the theoretical framework builds on a wide variety of theories coming from fields as diverse as human to computer interaction, retail sales interaction, and advertising.

First, the impact on the larger field of voice assistant research. This research has shown that voice human-likeness impacts the perception of the voice assistant and has identified *social presence* and *anthropomorphism* as mediators explaining this perception. If previous research has shown that social presence impacts overall attitude towards a web-based virtual assistant (Chérif & Lemoine, 2019), none has used anthropomorphism as a mediator to explain this social presence.

Second, this thesis contributes to the field of advertising context by showing that attitude towards the voice assistant mediates the attitudes towards the brand. The advertising context of the voice assistant has been little researched – most likely due to the newness of the device – and our work is among the first building stones for studying voice assistants as a marketing touchpoint (Smith, 2018).

Third, this work shows that widely accepted theories on personalization in service encounters (Reeves & Nass, 1996; Coelho & Henseler, 2012) do not automatically apply to a sales interaction through a voice assistant, despite similarities with human to human interaction.

Further research will be needed to understand to what extent theories from the retail research can be applied to the field of conversational commerce through voice assistants.

Eventually, the findings act as a criticism for the theory that personalization in sales encounter has a positive effect on attitudes towards the brand and more specifically brand purchase intention, which was never supported by the data analysis. However, this finding needs further research as it could be influenced by other variables not considered in our study, such as the nature of goods being purchased through the voice assistant.

7.3 Contribution to practice

7.3.1 Implication for task-application development

This study underlines the importance of using human-like voices for the task-applications in a voice assistant. By enhancing the human-likeness in the voice, task-developers can expect direct improvements on trust and loyalty towards the voice assistant as well as improved perceptions of service quality and satisfaction. These effects appear to hold no matter the users reported level of familiarity and interest in technology. These insights become extremely relevant since the pre-study with ICA showed that, on the Google Home Android platform, application developers have the freedom of choosing which voice to use in their application.

Moreover, it was found that *trust*, *loyalty*, *service quality* and *satisfaction* will improve the greater sense of social presence is experience in the encounter. The study also found that this effect can be enhanced in practice by making the voice assistant more likely to be anthropomorphized. As a result, application developers should focus on making the interaction more human-like. This could potentially be done in other ways than voice, by for example assigning a name and persona to the voice assistant.

Lastly, making the dialogue more personalized did not affect neither *trust*, *loyalty*, *service quality* nor *satisfaction*. The questions is therefore raised weather it is beneficial to assign considerable time and resources into copywriting personal dialogues and extensively using customer data in the voice assistant interaction. The main-take away from this thesis is that task-developers will reap the greatest benefits from voice assistants by making their application appear more human-like.

7.3.2 Implication for brand managers

This study also provides valuable insights for brand managers of brands intended to be mentioned and advertised on a voice assistant. The most prominent finding from this study is how user's attitudes about the voice assistant appears to mediate their relationship to the brand. The extent to which the voice assistant application can deliver a positive service encounter, in terms of *loyalty*, *trust*, *service quality* and *satisfaction* will spill-over to the brand being advertised. Brand managers should hence be aware of what voice-platforms their brands are being mentioned in, and assess the extent to which this platform is able to deliver on the four service variables.

Similarly, the study establishes the benefits of having a human-like voice, which have positive effects on the brand attitudes. Hence brand managers need to be aware which voice is used on the applications that are mentioning their brand.

Moreover, the benefits of personalizing the encounter through using customer data is seen on the effect on brand loyalty. Customers will be more likely to return to the brand if the voice assistant have made a personalized recommendation to them. The benefits of personalizing the dialogue of a voice assistant is hence greater for brands rather than for the sales-platform itself. Lastly, no direct effect was found on purchase intentions, indicating the necessity of establishing customers buying intents before their interaction with the voice assistant.

7.4 Critique and limitations

Although careful considerations were made to enhance the generalizability, internal validity and data quality of the study, results must be interpreted in the light of inherent limitations of the research.

7.4.1 Limitations to ecological validity

With regards to the ecological validity, one can raise concerns regarding the choice of having respondents listen to an audio recording and answer the survey from the perspective of the interaction heard in the recording. The method only studies a hypothetical scenario, whereas theory from human-robot interaction research studies the behaviors and actions related to interacting with a robot (Dautenhahn, 2007). As a result, the study measures attitudes and intentions, which does not have to correlate with behaviors (Hoyer & MacInnis, 2008). Although this delimitation was made to enhance the feasibility of the study, the results are still valuable for practice, as measuring attitudes is common practice within marketing research (Spears & Singh, 2004)

7.4.2 Limitations to internal validity

A further methodological critique is the choice of dialogue as an independent variable. In contrast to voice human-likeness, personalization of dialogue is a much more multi-faceted concept (Abidi, 2002). Firstly, it is inherently difficult to measure the concept of personalization through a standard stimulus audio recording that is the same for all participants. It is possible that our results would be different if the manipulation were adapted towards each independent respondent, however such research was not practically feasible. However, the use of social language, friendliness and avoiding scripted service roles are dimensions are also part of the concept of personalization (Surprenant & Solomon, 1987), and through careful pre-testing were able to ensure that the different dialogues showed significant internal validity in terms of perceived personalization.

With respect to the design of the questionnaire, it is possible that answers relating to the voice assistant were misinterpreted and answered as general opinions about all voice assistants and not about this specific interaction. However, the sample size should cancel out individual difference in attitudes and significant differences should only be with regards to the manipulation. Although the questionnaire was pre-tested 8 times, and measures showed adequate Cronbach's alpha. It is still difficult to know whether some questions were misinterpreted and did not measure what they intended to.

7.4.3 Limitations to external validity

Similarly, the study only tested the context of purchasing toothpaste, a low-involvement routine purchase. The study did not test different contexts of interaction with a voice assistant, the effect of different product categories nor the effect on another brand. While the results suggest that a low-involvement purchase would follow a similar path, they cannot be generalized towards another context, product category nor brand. However, the choice of product category was backed up both by pre-test 2 and previous research (Kim et al., 2018), hence we believe this is a springboard for future research.

Although *MTurk* has been argued to lead to diverse and high-quality data (Buhrmester et al., 2011; Paolacci et al., 2010), respondents were mainly located in the US and India. Despite diverse demographic variables, it is difficult to generalize the responses towards other nationalities and non-*MTurk* users.

7.5 Further research

As a general comment, further academic works in the nascent field of voice assistant research is needed. This thesis is a first building stone towards a broader understanding of the voice assistant phenomena but is not enough to holistically and thoroughly apprehend the opportunities, challenges and specificities of conversational commerce through voice assistants. Further research topics are therefore multiple.

First, it would be interesting to conduct a study on conversational commerce in an interactive, real-life setting, with the researcher in the room with the participant. Such new methodology would answer to several of the limitations of this study, such as having a real scenario instead of a hypothetical one and active involvement of the respondent in the experimental research.

Second, further research should be conducted to study voice assistant attitudes in other contexts, such as a request for information. Likewise, more research is needed to understand brand attitudes in other conversational commerce contexts, such as in a voice request regarding a high-involvement purchase. It would be particularly interesting to test if, in these new contexts, dialogue personalization would prove relevant and be significant.

Third, more academic work is needed to identify further mediating and moderating relationships explaining overall attitude towards brand advertised through the voice assistant. This research studied the moderating effect of the subjective variable *personal innovativeness*, however looking through an objective lens at previous interactions with a voice assistant could likely influence attitudes towards the device. Likewise, research on further variables influencing attitude towards the voice assistant would be welcome.

Eventually, additional exploration of brand recommendation and brand advertisement should be done in the future. Voice assistants provide a new and exceptionally well context-blended touchpoint for companies and provide countless opportunities for brand marketing and relationship-building.

8. References

- Aaker, J.L. 1997, "Dimensions of brand personality", *Journal of Marketing Research*, vol. 34, no. 3, pp. 347-356.
- Abidi, A. 2002, *La personnalisation sur Internet: un essai de conceptualisation*, ATER, Centre de Recherche en Gestion.
- Agarwal, R. & Prasad, J. 1998, "A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology", *Information Systems Research*, vol. 9, pp. 204-224.
- Arnould, E.J. & Price, L.L. 1993, "River Magic: Extraordinary Experience and the Extended Service Encounter", *Journal of Consumer Research*, vol. 20, no. 1, pp. 24-45.
- Schuller, B., Rigoll, G. & Lang, M. 2003, "Hidden Markov model-based speech emotion recognition", *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, pp. II.
- Ball, D., Coelho, P. & Vilares, M. 2006, "Service Personalization and Loyalty", *JOURNAL OF SERVICES MARKETING*, vol. 20, no. 6, pp. 391-403.
- Bänziger, T. & Scherer, K.R. 2005, "The role of intonation in emotional expressions", *Speech Communication*, vol. 46, no. 3, pp. 252-267.
- Bartneck, C., Kulic, D., Croft, E. & Zoghbi, S. 2009, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots", *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71-81.
- Bellini, N. & Convert, L. 2016, "The Concierge. Tradition, Obsolescence and Innovation in Tourism", *Symphony*, , no. 2, pp. 17-25.
- Bentahar, A. 2018, December 3, 2018, "How Voice Search Is Changing Shopping", *Forbes*, online. Available <https://www.forbes.com/sites/forbesagencycouncil/2018/12/03/how-voice-search-is-changing-shopping/>, consulted May 2, 2019.
- Biocca, F., Harms, C. & Burgoon, J.K. 2003, "Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria", *Presence: Teleoperators and Virtual Environments*, vol. 12, no. 5, pp. 456-480.
- Bitner, M.J., Booms, B.H. & Tetreault, M.S. 1990, "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents", *Journal of Marketing*, vol. 54, no. 1, pp. 71-84.
- Björkman, I. & Kock, S. 1995, "Social relationships and business networks: The case of Western companies in China", *International Business Review*, vol. 4, no. 4, pp. 519-535.
- Bloch, P.H. 1981, "An Exploration Into the Scaling of Consumers' Involvement With a Product Class", *NA - Advances in Consumer Research*, vol. 08, no. 1, pp. 61-65.
- Bloemer, J.M. & Kasper, H.D.P. 1995, "The complex relationship between consumer satisfaction and brand loyalty", *Journal of Economic Psychology*, vol. 16, no. 2, pp. 311-329.
- Braiker, B. 2018, January 12, 2018-last update, "Ad lib: What really works in voice – And why Google is smarter than Amazon", *AdAge*, Online. Available: <http://adage.com/article/podcasts/ad-lib-fresh-digital-s-doug-robinson-works-voice/311928/>, consulted on March 15, 2019.

- Brexendorf, T.O., Mühlmeier, S., Tomczak, T. & Eisend, M. 2010a, "The impact of sales encounters on brand loyalty", *Journal of Business Research*, vol. 63, no. 11, pp. 1148-1155.
- Bruner, G. 2016, *Marketing Scales Handbook: Multi-Item Measures for Consumer Insight Research*, 8th edn, Createspace Independent Publishing Platform, Carbondale.
- Bryman, A. & Bell, E. 2015, *Business Research Methods*, 4th edn, Oxford University Press.
- Buhrmester, M., Kwang, T. & Gosling, S. 2011, "Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?", *Perspect Psychol Sci*, vol. 6, no. 1.
- Burnham, T., Frels, J. & Mahajan, V. 2003, "Consumer Switching Costs: A Typology, Antecedents, and Consequences.", *Journal of the Academy of Marketing Science*, vol. 31, no. 2, pp. 109-126.
- Chaudhuri, S. & Terlep, S. 2018, February 27, 2018, "The Next Big Threat to Consumer Brands (Yes, Amazon's Behind It)", *The Wall Street Journal*, online. Available: <https://www.wsj.com/articles/big-consumer-brands-dont-have-an-answer-for-alexa-1519727401>. Consulted April 9, 2019.
- Cheetham, M., Suter, P. & Jäncke, L. 2011, "The Human Likeness Dimension of the "Uncanny Valley Hypothesis": Behavioral and Functional MRI Findings", *Front Hum Neurosci*, vol. 5, no. 126.
- Chérif, E. & Lemoine, J. 2019, "Anthropomorphic virtual assistants and the reactions of Internet users: An experiment on the assistant's voice", *Recherche et Applications en Marketing (English Edition)*, vol. 34, no. 1, pp. 28-47.
- Choi, Y.H. & Choo, H.J. 2016, "Effects of Chinese consumers' relationship benefits and satisfaction on attitudes toward foreign fashion brands: The moderating role of country of salesperson", *Journal of Retailing and Consumer Services*, vol. 28, pp. 99-106.
- Chow, S. & Holden, R. 1997, "Toward An Understanding Of Loyalty: The Moderating Role Of Trust", *Journal of Managerial Issues*, vol. 9, no. 3, pp. 275-298.
- Chung, H., Iorga, M., Voas, J. & Lee, S. 2017, "Alexa, Can i Trust You?", *Computer*, vol. 50, no. 9, pp. 100-104.
- Cialdini, R.C. 1993, *Influence: Science and practice*, 3rd edn, Harper Collins, New York.
- Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V. & Cowan, B.R. 2019, "What Makes a Good Conversation?: Challenges in Designing Truly Conversational Agents", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* ACM, Glasgow, Scotland UK, May 04 - 09, 2019, pp. 475.
- Coelho, P.S. & Henseler, J. 2012, "Creating customer loyalty thorough service customization.", *European Journal of Marketing*, vol. 46, no. 3, pp. 331-356.
- Cowan, B.R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D. & Bandeira, N. 2017, ""What can i help you with?": Infrequent users' experiences of intelligent personal assistants", *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2017*.
- Cronin, J.J. & Taylor, S.A. 1992, "Measuring Service Quality: A Reexamination and Extension", *Journal of Marketing*, vol. 56, no. 3, pp. 55-68.

- Cronin, J.J., Brady, M.K. & Hult, G.T.M. 2000, "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments", *Journal of Retailing*, vol. 76, no. 2, pp. 193-218.
- Cyr, D., Hassanein, K., Head, M. & Ivanov, A. 2007, "The role of social presence in establishing loyalty in e-Service environments", *Interacting with Computers*, vol. 19, no. 1, pp. 43-56.
- Dautenhahn, K. 2007, "Methodology & Themes of Human-Robot Interaction: A Growing Research Field", *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, pp. 15.
- Dodds, W.B., Monroe, K.B. & Grewal, D. 1991, "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations", *Journal of Marketing Research*, vol. 28, no. 3, pp. 307-319.
- Duffy, B.R. 2003, "Anthropomorphism and the social robot", *Robotics and Autonomous Systems*, vol. 42, no. 3-4, pp. 177-190.
- Edlund, J., Gustafson, J., Heldner, M. & Hjalmarsson, A. 2008, "Towards human-like spoken dialogue systems", *Speech Communication*, vol. 50, no. 8-9, pp. 630-645.
- Elkins, A.C. & Derrick, D.C. 2013, "The sound of trust: Voice as a measurement of trust during interactions with Embodied Conversational Agents", *Group Decision and Negotiation*, vol. 22, no. 5, pp. 897-913.
- Epley, N., Waytz, A. & Cacioppo, J.T. 2007, "On Seeing Human: A Three-Factor Theory of Anthropomorphism", *Psychological review*, vol. 114, no. 4, pp. 864-886.
- Fink, J. 2012, "Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction", *Proceedings of the 4th International Conference on Social Robotics*, eds. S.S. Ge, O. Khatib, J. Cabibihan, R. Simmons & M. Williams, Springer, China, October 29-31, 2012, pp. 199-208.
- Finstad, K. 2010, "Response interpolation and scale sensitivity: evidence against 5-point scales", *Journal of Usability Studies*, vol. 3, no. 5, pp. 104-110.
- Ford, M. & Palmer, W. Pers 2018, "Alexa, are you listening to me? An analysis of Alexa voice service network traffic", *Personal and Ubiquitous Computing*, vol. 23, no. 1, pp. 67-79.
- Gartner 2016, October 3, 2016, "Gartner says worldwide spending on VPA-enabled wireless speakers will top \$2 billion by 2020", *Gartner Newsroom*, online. Available: <https://gtnr.it/2IRJEIL>, consulted March 15, 2019.
- Gefen, D. 2002, "Customer Loyalty in E-Commerce", *Journal of the Association for Information Systems*, vol. 3, pp. 27-51.
- Gefen, D. & Straub, D.W. 2004, "Consumer trust in B2C e-Commerce and the importance of social presence: Experiments in e-Products and e-Services", *Omega*, vol. 32, no. 6, pp. 407-424.
- Geuens, M. & De Pelsmacker, P. 2017, "Planning and Conducting Experimental Advertising Research and Questionnaire Design", *Journal of Advertising*, vol. 46, no. 1, pp. 83-100.
- Goldberg, M.E. & Gorn, G.J. 1987, "Happy and Sad TV Programs: How They Affect Reactions to Commercials", *Journal of Consumer Research*, vol. 14, no. 3, pp. 387-403.

- Gong, L. 2008, "How social is social responses to computers? The function of the degree of anthropomorphism in computer representations", *Computers in Human Behavior*, vol. 24, no. 4, pp. 1494-1509.
- Grohmann, B. 2009, "Gender Dimensions of Brand Personality", *Journal of Marketing Research*, vol. 46, no. 1, pp. 105-119.
- Grove, S., Fisk, R. & John, J. 1983, "The dramaturgy of service exchange: an analytical framework for service marketing.", *Emerging perspectives on services marketing*, American Marketing Association, pp. 45.
- Hamman, B.M. 2006, "Two voices: Social presence, participation, and credibility in online news", *Journalism electronic theses and dissertations*, University of Missouri-Columbia, available at <https://mospace.umsystem.edu/xmlui/handle/10355/4640>, consulted May 12, 2019.
- Hande, S.S. 2014, "A Review on Speech Synthesis an Artificial Voice Production", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 3, pp. 8046-8053.
- Hawes, J.M., Mast, K.E. & Swan, J.E. 1989, "Trust Earning Perceptions of Sellers and Buyers.", *Journal of Personal Selling & Sales Management*, vol. 9, no. Spring, pp. 1-8.
- Hess, T., Fuller, M. & Campbell, D. 2009, "Designing interfaces with social presence: Using vividness and extraversion to create social recommendation agents", *Journal of the Association of Information Systems*, vol. 10, no. 12, pp. 889-919.
- Hoy, M.B. 2018, "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants", *Medical reference services quarterly*, vol. 37, no. 1, pp. 81-88.
- Hoyer, D. & MacInnis, J. 2008, *Consumer behavior*, 5th edn, Independence, KY: South-Western Cengage.
- Hu, J.C. 2018, December 20, 2018, "People still aren't into buying things through their smart speakers", *Quartz*, online. Available: <https://bit.ly/2GJfnsb>, consulted April 30, 2019.
- Hunt, S. 2010, *Marketing Theory: Foundations, Controversy, Strategy, and Resource-advantage Theory*, 3rd edn, Routledge.
- Hwang, Y. 2009, "The impact of uncertainty avoidance, social norms and innovativeness on trust and ease of use in electronic customer relationship management", *Electronic Markets*, vol. 19, no. 2-3, pp. 89-98.
- Iribarren, M. 2018, December 28, 2019-last update, "H&M Home Enables Voice Shopping with Google Assistant", *Voicebot.AI*, online. Available: <https://bit.ly/2M13srC>, March 30, 2019.
- Jacobsen, D. 2002, *Vad, hur och varför? Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*. First edition edn, Studentlitteratur, Lund.
- Jacoby, J. & Kyner, D.B. 1973, "Brand Loyalty Vs. Repeat Purchasing Behavior", *Journal of Marketing Research*, vol. 10, no. 1, pp. 1-9.
- Jiang, L., Hoegg, J., Dahl, D. & Chattopadhyay, A. 2010, "The Persuasive Role of Incidental Similarity on Attitudes and Purchase Intentions in a Sales Context", *Journal of Consumer Research*, vol. 36, no. 5, pp. 778-779.

- Kantar Worldpanel 2017, 2017-last update, "Leading fast moving consumer goods (FMCG) brands worldwide in 2017, by household reach (in million Consumer Reach Points)", *Statista*, online. Available: <https://www.statista.com/statistics/718304/fmcg-brands-reach/>, May 2, 2019.
- Karlsberg, J. 2016, *Reception, Reception, Reception. The effect of receiver context on advertising effectiveness*, Stockholm School of Economics.
- Kim, D., Park, K., Park, Y., Ju, J. & Ahn, J. 2018, "Alexa, Tell Me More: The Effect of Advertisements on Memory Accuracy from Smart Speakers", *PACIS 2018 - Opportunities and Challenges for the Digitized Society: Are We Ready?*, ed. AIS, AIS, , 26 June - 30 June, pp. 204.
- Kinsella, B. 2018, August 6, 2018, "The Information Says Alexa Struggles with Voice Commerce But Has 50 Million Devices Sold – One of Those Seems Right", *Voicebot.AI*, online. Available: <https://voicebot.ai/2019/03/15/5-percent-of-dutch-households-adopt-smart-speakers-in-just-4-5-months-google-home-is-the-leader/>, consulted May 7, 2019.
- Kinsella, B., March 25, 2019, "5 Percent of Dutch Households Adopt Smart Speakers in Just 4.5 Months, Google Home is the Leader", *Voicebot.AI*, online. Available: <https://voicebot.ai/2018/08/06/the-information-says-alexa-struggles-with-voice-commerce-but-passes-50-million-devices-sold-one-of-those-seems-right/>, consulted May 12, 2019.
- Kiseleva, J., Williams, K., Jiang, J., Awadallah, A.H., Crook, A.C., Zitouni, I. & Anastasakos, T. 2016, "Understanding user satisfaction with intelligent assistants", *CHIIR 2016 - Proceedings of the 2016 ACM Conference on Human Information Interaction and Retrieval*, pp. 121.
- Krugman, H.E. 1983, "Television program interest and commercial interruption", *Journal of Advertising Research*, vol. 23, no. 1, pp. 21-23.
- Lau, J., Zimmerman, B. & Schaub, F. 2018, "Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers", *Journal Proceedings of the ACM on Human-Computer Interaction - CSCW*, vol. 2, no. CSCW, pp. Article 102.
- Lee, K.-. & Nass, C. 2005, "Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices", *Media Psychology*, vol. 7, no. 1, pp. 31-45.
- Lemmink, J.G.A.M. & Bloemer, J.M.M. 1992, "The importance of customer satisfaction in explaining brand and dealer loyalty", *Journal of Marketing Management*, vol. 8, no. 4, pp. 351-363.
- Leung, F. & Cheung, C. 2004, "Consumer Attitude toward Mobile Advertising", *AMCIS 2004 Proceedings*, pp. 331.
- Levin, M.R. & Lowitz, J.N. 2019, February 5, 2019, "Smart Speaker Market Takes Off in Holiday Quarter", *Consumer Intelligence Research Partner*, online. Available: <https://bit.ly/2SSntqg>, consulted March 15, 2019.
- Li, H., Daugherty, T. & Biocca, F. 2002, "Impact of 3-D Advertising on Product Knowledge, Brand Attitude, and Purchase Intention: The Mediating Role of Presence", *Journal of Advertising*, vol. 31, no. 3, pp. 43-57.
- Ligas, M. 2004, "Personalizing Services Encounters", *Services Marketing Quarterly*, vol. 25, no. 4, pp. 33-51.

- Lu, J., Yao, J.E. & Yu, C. 2005, "Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology", *The Journal of Strategic Information Systems*, vol. 14, no. 3, pp. 245-268.
- Lynn, A. & Lynn, M. 2003, "Experiments and Quasi-Experiments: Methods for evaluating Marketing Options.", *The Cornell Hotel and Restaurant Administration Quarterly*, vol. 44, no. 2, pp. 75-84.
- Malhotra, N. 2010, *Review of marketing research*. Armonk, N.Y.: M.E.
- Marascuilo, L.A. & Levin, J.R. 1970, "Appropriate Post Hoc Comparisons for Interaction and Nested Hypotheses in Analysis of Variance Designs: The Elimination of Type IV Errors", *American Educational Research Journal*, vol. 7, no. 3, pp. 397-421.
- McKnight, D.H., Choudhury, V. & Kacmar, C. 2002, "Developing and validating trust measures for e-commerce: An integrative typology", *Information Systems Research*, vol. 13, no. 3, pp. 334-359.
- Merriam-Webster Unknown last update, "Dialogue (definition)", *Merriam-Webster*, online. Available: <https://www.merriam-webster.com/dictionary/dialogue>, consulted 2019, May 6, 2019.
- Meuter, M., Ostrom, A., Roundtree, R. & Bitner, M. 2000, "Self-Service Technologies: Understanding Customer Satisfaction with Technology-Based Service Encounters", *Journal of Marketing*, vol. 64, pp. 50-64.
- Meyersohn, N. 2019, April 2, 2019, "Walmart partners with Google for voice-assisted grocery shopping", *CNN Business*, online. Available: <https://edition.cnn.com/2019/04/02/business/walmart-voice-order-grocery-pickup-amazon-alexa/index.html>, consulted April 9, 2019].
- Mitchell, A.A. & Olson, J.C. 1981, "Are product attribute beliefs the only mediator of advertising effects on brand attitude?", *Journal of Marketing Research*, vol. 18, no. 3, pp. 318-332.
- Mittal, B. & Lassar, W.M. 1996, "The role of personalization in service encounters", *Journal of Retailing*, vol. 72, no. 1, pp. 95-109.
- Moore, M. 2018, January 11, 2018, "Voice assistants set to revolutionize commerce and become a dominant mode of consumer interaction in the next three years", *Capgemini*, online. Available: <https://bit.ly/2DB04Pg>, consulted April 9, 2019.
- Moorman, M., Willemsen, L.M., Neijens, P.C. & Smit, E.G. 2012, "Program-Involvement Effects on commercial Attention and Recall of Successive and Embedded Advertising", *Journal of Advertising*, vol. 41, no. 2, pp. 25-38.
- Morgan, R.M. & Hunt, S.D. 1994, "The Commitment-Trust Theory of Relationship Marketing", *Journal of Marketing*, vol. 58, no. 3, pp. 20-38.
- Mori, M. 1970, "Bukimi no tani The uncanny valley", *Energy*, vol. 7, no. 4, pp. 33-35.
- Moriuchi, E. 2019, "Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty", *Psychology and Marketing*, vol. 36, no. 5.
- N. Lubold, E. Walker & H. Pon-Barry 2016, "Effects of voice-adaptation and social dialogue on perceptions of a robotic learning companion", *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 255.

- Nass, C. & Brave, S. 2005, *Wired for Speech*, 1st edn, MIT Press, Cambridge, MA.
- Nass, C. & Moon, Y. 2000, "Machines and mindlessness: Social responses to computers", *Journal of Social Issues*, vol. 56, no. 1, pp. 81-103.
- 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*, vol. 7, no. 3, pp. 171-181.
- Nielsen, Hui, R., Delfi, & DLF, 2017, "Market share of selected grocery retailers in Sweden in 2017", *Statista*, online. Available: <https://www.statista.com/statistics/565650/market-share-of-selected-grocery-retailers-in-sweden/>, consulted March 13, 2019.
- Nowak, K.L. & Biocca, F. 2003, "The Effect of the Agency and Anthropomorphism on users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments", *Presence: Teleoperators and Virtual Environments*, vol. 12, no. 5, pp. 481-494.
- OC&C Strategy Consultants 2018, February 28, 2018, "Voice Shopping Set to Jump to \$40 Billion By 2022, Rising From \$2 Billion Today", *Cision PR Newswire*, online. Available: <https://prn.to/2tez9Ix>, consulted March 16, 2019.
- Orehovacki, T., Etinger, D. & Babic, S. 2019, "The antecedents of intelligent personal assistants adoption", Nunes I. (eds) *Advances in Human Factors and Systems Interaction 2018*, Advances in Intelligent Systems and Computing, vol. 781, Springer, Cham. pp. 76-87.
- Orr, J.M., Sackett, P.R. & Dubois, C.L. 1991, "Outlier detection and treatment in I/O psychology: A survey of researcher beliefs and an empirical illustration", *Personnel Psychology*, vol. 44, no. 3, pp. 473-486.
- Paolacci, G., Chandler, J. & Ipeirotis, P. 2010, "Running experiments on Amazon Mechanical Turk", *Judgment and Decision Making*, vol. 5, no. 5.
- Parasuraman, A., Zeithaml, V.A. & Berry, L.L. 1985, "A conceptual model of service quality and its implications for future research", *Journal of Marketing*, vol. 49, no. 4, pp. 41-50.
- Pelsmacker, P.D., Geuens, M. & Anckaert, P. 2002, "Media Context and Advertising Effectiveness: The Role of Context Appreciation and Context/Ad Similarity", *Journal of Advertising*, vol. 31, no. 2, pp. 49-61.
- Porcheron, M., Fischer, J., Reeves, S. & Sharples, S. 2018, "Voice Interfaces in Everyday Life", *CHI '18 Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ed. ACM, ACM, Montreal QC, Canada, April 21-26, 2018, pp. Paper no. 640.
- Pornpitakpan, C., Yuan, Y. & Han, J.H. 2017, "The effect of salespersons' retail service quality and consumers' mood on impulse buying", *Australasian Marketing Journal*, vol. 25, no. 1, pp. 2-11.
- PostNord 2018, October 19, 2018, "Allt fler européer e-handlar och fler och fler handlar från utländska sajter", *PostNord*, online. Available: <https://bit.ly/2DznDKv>, consulted April 30, 2019.
- Preacher, J. & Hayes, A. 2008, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models", *Behavior Research Methods*, vol. 40, no. 3, pp. 879-891.

- Pugh, S.D. 2001, "Service with a Smile: Emotional Contagion in the Service Encounter", *The Academy of Management Journal*, vol. 44, no. 5, pp. 1018-1027.
- Purington, A., Taft, J.G., Sannon, S., Bazarova, N.N. & Taylor, S.H. 2017, "'Alexa is my new BFF': Social roles, user satisfaction, and personification of the Amazon Echo", *Conference on Human Factors in Computing Systems - Proceedings*, pp. 2853.
- PwC 2018, "Prepare for the voice revolution", *PwC US*, online. Available: <https://pwc.to/2G8gGi3>, consulted April 9, 2019.
- Qiu, L. & Benbasat, I. 2008, "Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems", *Journal of Management Information Systems*, vol. 25, no. 4, pp. 145-181.
- Ramirez, R., Mukherjee, M., Vezzoli, S. & Kramer, 2015, "Scenarios as a scholarly methodology to produce 'interesting research'", *Futures*, vol. 71, pp. 70-87.
- Reed, J. & Roskel Payton, V. 1997, "Focus groups: issues of analysis and interpretation", *Journal of Advanced Nursing*, vol. 26, pp. 765-771.
- Reeves, B. & Nass, C.I. 1996, *The media equation: How people treat computers, television, and new media like real people and places*, Cambridge University Press, New York, NY, US.
- Rogers, E. 1983, *Diffusion of Innovation*, 3rd edn, Macmillan Publishing Co., Inc., New York, N. Y.
- Rosengren, S. & Dahlén, M. 2015, "Exploring Advertising Equity: How a Brand's Past Advertising May Affect Consumer Willingness to Approach Its Future Ads", *Journal of Advertising*, vol. 44, no. 1, pp. 1-13.
- Ross, J., Irani, L., Zaldivar, A. & Tomlinson, B. 2010, "Who are the Turkers? Worker Demographics in Amazon Mechanical", *CHI EA*, pp. 2863-2872.
- Salesforce 2018, June 5, 2018, "State of the Connected Customer. Insights from 6,700+ consumers and business buyers on the intersection of experience, technology, and trust", *Salesforce*, online. Available: <https://bit.ly/2yLvvggN>, consulted April 30, 2019.
- Saunders, M., Lewis, P. & Thornhill, A. 2012, *Research Methods for Business Students*. Pearson Education Ltd, Harlow.
- Schröder, M. 2001, "Emotional Speech Synthesis: A Review", *EUROSPEECH 2001 Scandinavia*, Aalborg, Denmark, September 3-7, 2001, pp. 561-564.
- Sciuto, A., Saini, A., Forlizzi, J. & Hong, J.I. 2018, "'Hey Alexa, What's Up?': A Mixed-Methods Studies of In-Home Conversational Agent Usage", *DIS '18 Proceedings of the 2018 Designing Interactive Systems Conference*, ed. ACM, ACM, Hong Kong, China, June 09 - 13, 2018, pp. 857-868.
- Shamdasani, P.N. & Balakrishnan, A.A. 2000, "Determinants of relationship quality and loyalty in personalized services", *Asia Pacific Journal of Management*, vol. 17, no. 3, pp. 399-422.
- Sims, V.K., Chin, M.G., Lum, H.C., Upham-Ellis, L., Ballion, T. & Lagattuta, N.C. 2009, "Robots' Auditory Cues are Subject to Anthropomorphism", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 53, no. 18, pp. 1418-1421.
- Smith, K.T. 2018, "Marketing via smart speakers: what should Alexa say?", *Journal of Strategic Marketing*, DOI: 10.1080/0965254X.2018.1541924

- Söderlund, M. 2010, *Experiment Med Människor*, Liber, Malmö.
- Söderlund, M. 2005, *Mätningar och mått - i marknadsundersökaren värld*, Liber, Malmö.
- Solomon, M.R., Surprenant, C.F., Czepiel, J.A. & Gutman, E.G. 1985, "A role theory perspective on dyadic interactions: The Service Encounter", *Journal of Marketing*, vol. 49, pp. 99-111.
- Spears, N. & Singh, S.N. 2004, "Measuring Attitude toward the Brand and Purchase Intentions", *Journal of Current Issues & Research in Advertising*, vol. 26, no. 2, pp. 53-66.
- Srinivasan, S., Anderson, R. & Ponnnavolu, K. 2002, "Customer loyalty in e-commerce: an exploration of its antecedents and consequences", *Journal of Retailing*, vol. 78, pp. 41-50.
- Sterling, G., February 12, 2019, "Analyst: 8 billion voice assistants by 2023", *Search Engine Land*. Available: <https://searchengineland.com/analyst-8-billion-voice-assistants-by-2023-312035>. Consulted on May 12, 2019
- 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*, vol. 64, no. 1, pp. 43-52.
- Stone, G., Besser, D. & Lewis, L. 2000, "Recall, Liking, and Creativity in TV Commercials: A New Approach", *Journal of Advertising Research*, vol. 40, no. 3.
- Strack, F. 1992, "Order Effects" in *Survey Research: Activation and Information Functions of Preceding Questions*, Context Effects in Social and Psychological Research, Springer, New York, NY.
- Sundar, S.S. 2004, "Loyalty to computer terminals: Is it anthropomorphism or consistency?", *Journal of Behaviour & Information Technology*, vol. 23, no. 2, pp. 107-118.
- Surprenant, C.F. & Solomon, M.R. 1987, "Predictability and personalization in the service encounter", *Journal of Marketing*, vol. 51, no. 2, pp. 86-96.
- Swan, J.E., Bowers, M.R. & Richardson, L.D. 1999, "Customer trust in the salesperson: An integrative review and meta-analysis of the empirical literature", *Journal of Business Research*, vol. 44, no 1. pp. 93-107.
- Sweeney, J.C., Soutar, G.N. & Johnson, L.W. 1997, "Retail service quality and perceived value: A comparison of two models." *Journal of retailing and consumer services*. Vol. 4. No 1. Pp. 39-48
- Szucs, D. & Ioannidis, J. 2017, "When Null Hypothesis Significance Testing Is Unsuitable for Research: A Reassessment", *Frontiers in Human Neuroscience*, vol. 11, pp. 390.
- Thibaut, J.W. & Kelley, H.H. 1959, *The social psychology of groups*, John Wiley, Oxford, England.
- Tunvall, F. 2018, *How conversational AI concepts enrich customer service interactions*, IBM, <https://www.ibm.com/blogs/watson/2018/09/how-conversational-ai-concepts-enrich-customer-service-interactions/>, consulted May 12, 2019
- Turan, A., Tunç, A.Ö. & Zehir, C. 2015, "A Theoretical Model Proposal: Personal Innovativeness and User Involvement as Antecedents of Unified Theory of Acceptance and Use of Technology.", *Procedia - Social and Behavioral Sciences*, vol. 210, pp 43 – 51

- Vincent, J. 2018, March 27, 2018, "Google launches more realistic text-to-speech service powered by DeepMind's AI", *The Verge*. Available: <https://www.theverge.com/2018/3/27/17167200/google-ai-speech-tts-cloud-deepmind-wavenet>, consulted May 2, 2019
- Vishwanath, A. & Goldhaber, G.M. 2003, "An Examination of the Factors Contributing to Adoption Decisions among Late-Diffused Technology Products", *New Media & Society*, vol. 5, no. 4, pp. 547-572.
- Voza, S. 2018, October 24, 2018, "What Retailers Need to Know About Voice Shopping", *Shopify*, online. Available: <https://www.shopify.com/retail/voice-shopping-what-retailers-need-to-know>, consulted April 9, 2019.
- Walters, M., Syrdal, D. & Koay, K. 2008, "Human approach distances to a mechanical-looking robot with different robot voice styles.", *IEEE International Symposium on Robot and Human Interactive*, August 1-3.
- Weisberg, J., Te'eni, D. & Arman, L. 2011, "Past purchase and intention to purchase in e-commerce: The mediation of social presence and trust", *Internet Research*, vol. 21, no. 1, pp. 82-96.
- Wu, P.C.S., Yeh, G.Y. & Hsiao, C. 2011, "The effect of store image and service quality on brand image and purchase intention for private label brands.", *Australasian Marketing Journal*, vol. 19, pp.30-39
- Yang, Z. 2001, "Consumer perceptions of service quality in internet-commerce: Strategic Implications", *Business Premium Collection*, American Marketing Association, vol. 12, pp. 76.
- Zhao, X., Lynch Jr, J. & Chan, Q. 2010, "Reconsidering Baron and Kenny: Myths and truths about mediation analysis", *Journal of consumer research*, vol. 37, no. 2, pp. 197-206.
- Zhuang, J., Mei, T., Hoi, S.C.H., Xu, Y.-. & Li, S. 2011, "When recommendation meets mobile: Contextual and personalized recommendation on the go", *UbiComp'11 - Proceedings of the 2011 ACM Conference on Ubiquitous Computing*, pp. 153.
- Zlotowski, J., Strasser, E. & Bartneck, C. 2014, "Dimensions of anthropomorphism: from humanness to humanlikeness", *Proceedings of the 9th ACM/IEEE Conference on Human-Robot Interaction*, ed. G. Sagerer, ACM, New York, USA, pp. 66.

9. Appendix

9.1 Pre-study 1: Understanding consumers and application developers

9.1.1 Interview guide: Interview with a task-application developer at ICA

Semi-structured interview performed on February 12, 2019 at ICA's HQ in Solna. Face-to-face interview with one a Product Manager of Ica X working directly with their development of voice task applications. Interviewee prepared to stay anonymous.

Introduction and briefing

- Presentation of the thesis, interviewers and interview purpose
- Explanation of the interview procedure: duration, recording, confidentiality, structure

Background

- Could you introduce yourself and your position and department at ICA?

ICA's voice app

- Could you introduce yourself and your position and department at ICA?
- How does this app work and what does it offer?
- How did you develop the app and on what features did you focus? Why?
- What are customers using the ICA voice app for? What do you want customers use the voice app for?
- How do you see future developments for the ICA voice app?

9.1.2 Interview guide: Focus group with consumers

This focus group took place on February 15, 2019 at the Stockholm School of Economics and gathered six participants. Both thesis authors were present, one moderating the focus group while the other took notes.

Introduction and briefing

- Presentation of the thesis and interviewers
- Explanation of the focus group procedure: recording, confidentiality, duration

Discussion themes

- Smart speakers:
 - o What do you know about them and what is your opinion

- What do you associate with smart speakers?
- What is your attitude towards smart speakers?
- Use:
 - If you own a smart speaker – what do you use it for?
 - If you don't own a smart speaker – what usage could you see yourself having of it?
- Conversational commerce:
 - Have you ever shopped through a smart speaker? If yes, how was your experience like?
 - If you haven't shopped through a smart speaker – what could you see yourself buying through the device?
 - What are the important features and criteria, in your opinion, when shopping through a smart speaker?
 - Describe an ideal voice shopping interaction with a smart speaker

9.2 Pre-study 2: Brand choice

Introduction

You will be presented with 5 brands and asked about your familiarity and attitude towards these. Please answer as carefully yet intuitive as possible.

Thank you for your participation!

Core questions

[Logo of the brand]


How well do you know this brand?

- 1) I am very familiar with the brand
- 2) I have strong preconceptions about the brand
- 3) I have previously interacted with the brand (e.g purchased, seen an advertisement, tried the product)

What do you think about this brand?

- 1) Negative-Positive
- 2) Dislike-Like
- 3) Unfavorable-Favorable

Illustration from Qualtrics in original layout



. How well do you know this brand?

	Strongly Disagree				Strongly Agree		
	1	2	3	4	5	6	7
I am very familiar with the brand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have strong preconceptions about the brand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have previously interacted with the brand (e.g. purchased, seen an advertisement, tried the product)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

What do you think about this brand?

My general attitude towards this brand is							
Negative	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Positive
Dislike	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Like
Unfavorable	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Favorable

9.3 Pre-study 3: Test of independent variables

9.3.1 Voice manipulation pre-study questions

Introduction

Hello!

You will hear a hypothetical conversation with a Voi2ce Assistant similar to Google Home and Amazon Alexa. You will listen to a recording and afterwards be asked 3 questions about your experience. Imagine that you take on the role as the person speaking with the voice assistant in the recording.

Thank you for your participation!

Recording

Listen carefully to the following dialogue. Imagine that you are the person speaking with the voice assistant.

Questions

Imagine that you are the person you heard speaking with the voice assistant in the recording. The following questions address how personalised the conversation seems to be.

1) The conversation was very personalized

- 2) The voice assistant knows me very well
- 3) The voice assistant has stored many data points about my purchase behavior

9.3.2 Dialogue manipulation pre-study questions

Introduction

Hello!

You will hear a hypothetical conversation with a voice assistant similar to Google Home and Amazon Alexa. You will listen to a recording and afterwards be asked 3 questions about your experience. Imagine that you take on the role as the person speaking with the voice assistant in the recording.

Thank you for your participation!

Recording

Listen carefully to the following dialogue. Imagine that you are the person speaking with the voice assistant.

Questions

Did the voice of the voice assistant sound like a human?

- 1) The voice assistant sounded like a human
- 2) The conversation sounded like two humans speaking to each other during a sales encounter
- 3) During the recording I forgot that I was listening to a robot speaking

9.4 Voices ranked by levels of human-likeness

Voice 1 – V1 – Voice Victoria, <http://onlinetonegenerator.com/voice-generator.html>

Voice 2 – V2 – Google Home mini as of April, 2019

Voice 3 – V3 – Voice assistant impersonated by Sarah Wiggins

9.5 Dialogues ranked by levels of personalization

The dialogues used in the pre-study and those used in the main study are the same. All experiences and pre-studies testing different levels of voice human-likeness are using dialogue 2 – medium personalization.

9.5.1 Dialogue 1 – D1 – low personalization

User: Hey, I want to buy toothpaste.

Voice assistant (VA): There are 3 popular options of toothpaste: whitening, herbal, or for sensitive teeth. Choose one or listen to more options?

User: Whitening.

VA: Colgate Sensation White is on special offer of 2 for 1. Confirm to be added to your basket.

User: I confirm.

VA: From where should the order be placed?

User: From a nearby supermarket.

VA: The product is available at Best Price and Big Deal. From which store should the order be placed?

User: Best Price.

VA: The order can be picked-up from Best Price, or it can be home-delivered. Choose one option.

User: Home-delivery.

VA: Colgate Sensation White from Best Price has just been ordered. It will be home-delivered in the next 48 hours. Standard payment solution has been charged.

9.5.2 Dialogue 2 – D2 – medium personalization

User: Hey, I want to buy toothpaste.

VA: Sure! Your latest grocery list includes “buy whitening toothpaste”. Do you want to buy whitening toothpaste?

User: Yes.

VA: Okay. In this category, I found that Colgate Sensation White is on special offer of 2 for 1. Should I add it to your basket?

User: Sounds good.

VA: Okay, I’ll order 2 Colgate whitening toothpastes for you. Where do you want to order it from?

User: A nearby supermarket.

VA: Best Price is your nearest supermarket. Should I order it from there?

User: Yes.

VA: Okay, Colgate Sensation White from Best Price. Do you want to have it home-delivered?

User: Yes.

VA: Okay. I have just ordered Colgate Sensation White from Best Price for you. It will be home-delivered in the next 48 hours. Your purchase has been charged on your preferred payment solution.

9.5.3 Dialogue 3 – D3 – high personalization

User: Hey, I want to buy toothpaste.

VA: Very good choice. How about whitening toothpaste? It will give you the best smile to rock the day. Do you want me to order some for you?

User: Yes.

VA: Good decision. I see that your usual consumption of toothpaste is two tubes a month. I have found a great deal for you: Colgate Sensation White is on special offer of 2 for 1. Should I add it to your basket?

User: Sounds good.

VA: Great. Your three most recent orders were placed at Best Price supermarket. Do you want me to order your toothpaste from them?

User: Yes.

VA: Good. One last question: I see in your calendar that you are quite busy this week. Do you want to have your toothpaste home-delivered? Best Price can deliver between 18 and 20, when you will be back home from work.

User: Great.

VA: Okay, I have just ordered Colgate Sensation White from Best Price for you. It will be home-delivered in the next 48 hours, between 18 and 20. Your purchase has been charged on your favorite payment solution. Have a great day!

9.6 Main study questionnaire items

Block 1

We are conducting a study on voice assistants. In this survey, you will listen to a short audio recording of an interaction with a voice assistant and thereafter be asked questions about your experience.

All answers will be accepted if:

- All questions are filled out
- Answers to attention tests are correct
- Your time for completing the survey on Qualtrics AND MTurk are approximately the same

- Your time for survey completion is not a significant outlier from the mean of aggregated time to answer the survey

- You have not previously participated to one of our surveys

By continuing the survey, you are agreeing to these terms.

Thank you for your participation!

Block 2

Imagine that you are brushing your teeth and realize that you are now out of toothpaste. You have an intelligent voice assistant at home, similar to that of Amazon Alexa or Google Home. You decide to speak with the voice assistant to order toothpaste. You have no specific toothpaste brand in mind for this order.

Listen to the recording carefully.

You will then be asked questions about the content of the recording as well as your experience.

Block 3

MANIPULATION TESTS

What was the name of the store from which the product was purchased?

- 1) Best Price
- 2) Best Buy
- 3) Big Discount
- 4) Big Deal

When will the product be delivered?

- 1) Within 12h
- 2) Within 24h
- 3) Within 48h
- 4) Within 72h

Block 4-7

Following are questions regarding your experience of using this voice assistant. When answering the questions, imagine that you step into the role of the person that you just heard speaking with the voice assistant in the recording.

Variable	Items
Anthropomorphism (Bartneck, Kulic, and Croft, 2009)	(1) Fake – Natural (2) Machinelike – Humanlike (3) Unconscious – Conscious (4) Artificial – Lifelike (5) Dead - Alive (6) Stagnant - Lively (7) Mechanical - Organic (8) Inert – Interactive (9) Apathetic – Responsive (10) Dislike - Like (11) Unfriendly - Friendly (12) Unkind - Kind (13) Unpleasant - Pleasant (14) Awful - Nice (15) Incompetent - Competent (16) Ignorant - Knowledgeable (17) Irresponsible - Responsible (18) Unintelligent - Intelligent (19) Foolish – Sensible
Social Presence (Qiu & Benbasat, 2009)	Please answer the following statements regarding your experience about the interaction with the voice assistant 1) I felt a sense of human contact in this voice assistant. 2) I felt a sense of personal connection in this voice assistant. 3) I felt a sense of human warmth in this voice assistant. 4) I felt a sense of social connection with this voice assistant. 5) I felt a sense of human sensitivity in this voice assistant.
Service Satisfaction (Coelho & Henselers, 2012)	To what extent do you agree with the following statements - with regards your impression of the voice assistant 1) Overall, I am satisfied with this voice assistant 2) This voice assistant fulfills my expectations 3) The interaction with this voice assistant is ideal
Level of Personalization (Coelho & Henselers, 2012)	1) This voice assistant offered me products and services that satisfy my needs 2) This voice assistants gave me recommendations that I would not have found anywhere else 3) If I stopped using this voice assistant, I would miss out on many good

	product recommendations
Control Question Trust (Coelho & Henselers, 2012)	1) Please press "Neither agree nor disagree" 1) I trust this voice assistant 2) I believe this voice assistant acted in my best interest 3) I believed this voice assistant was sincere and genuine 4) I believe this voice assistant performed its role very well
Service Quality (Coelho & Henselers, 2012)	1) The quality of the service provided by this voice assistant was high 2) This voice assistant gave me valuable recommendations 3) This voice assistant was a reliable source of information 4) This voice assistant gives me clear and transparent information
Customer Loyalty (Coelho & Henselers, 2012) Brand Attitudes (Spears & Singh, 2004) (Bruner, 2009)	1) I would use this voice assistant the next time I am searching for products 2) I would recommend using this voice assistant to friends and family if they were asking for advice What is your attitude towards the brand, Colgate Sensation White, that was recommended by the voice assistant? 1) Negative - Positive 2) Do not like - Like 3) Worse (compared to other brands) - Better (compared to other brands)
Purchase Intentions (Grohmann's 2009) (Dodds, Monroe & Grewal, 1991) Brand Loyalty (Jacoby & Kyner, 1973) (Aaker, 1996)	How likely are you to buy Colgate Sensation White in the near future? 1) Not likely - Likely 2) Not at all probable - Probable 3) I will buy this brand the next time I buy toothpaste To what extent do you agree with the following statements? 1) I would choose this brand over other brands selling toothpaste 2) I would recommend this brand to others 3) I would be willing to pay a higher price for this brand compared with other brands of toothpaste

Block 8

Thank you for your answers so far! The following questions are about your own background - please answer from your own perspective and not based on the recording you have listened to.

Variable	Items
Personal Innovativeness (Bloch, 1981)	1) I have used a voice assistant many times before 2) I am very positive towards using voice assistants in my everyday life 3) I am curious about new technologies 4) I am usually the first to try out new technologies in my social circle 5) I have no concern for companies collecting my personal data 6) I believe I benefit from sharing my data with companies
Demographic Variables - Gender	Which gender do you identify with? 1)Male 2)Female 3)Other 4) Not willing to disclose
Demographic Variables - Age	What is your age?
Purpose of Survey	What do you think was the purpose of this survey?
Feedback	Please let us know if you have any feedback that could help us improve this survey