BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

Is Your Voice Enough, Alexa? Assessing the Role of Digital Assistant Personality, Modality, and Product Involvement on Consumer Evaluations.

Navn:	Henrik Christensen, Stian Eriksen

Start:	15.01.2021 09.00
Finish:	01.07.2021 12.00

Is Your Voice Enough, Alexa? Assessing the Role of Digital Assistant Personality, Modality, and Product Involvement on Consumer Evaluations.

Program:

MSc in Strategic Marketing Management

BI Norwegian Business School

Supervisor:

Carlos Velasco

Date:

01.07.2021

Acknowledgments

We would like to display the utmost gratitude to key people who have been invaluable in the process of developing this Master's Thesis. Thank you, Carlos Velasco, our supervisor. You are truly an attentive, supportive, understanding, and knowledgeable Professor who has given much of your time to provide us with guidance throughout this period. We are sincerely thankful for your patience and assistance.

Thank you, Elia Gatti, UX researcher at Amazon. Your ideas and tips gave us the foundation to develop an even better research design and assisted us in creating personalities for our experiment in a much better way than what we could have achieved on our own. Additional feedback received through our discussions were also highly appreciated and provided us with a more nuanced understanding.

Finally, we would like to thank those who have had to witness our endeavor firsthand. Thank you to our families and especially, thank you to our girlfriends. You have been miraculously patient and understanding of our priority of working with this Master's Thesis day and night over the last six months. We are truly grateful for your support.

For us, this has been a long and strenuous, yet extremely rewarding journey. Now that it is over, we are proud to say that what is in front of you, is the final result of our Master's Thesis. It marks the end of two challenging years of our education, which have given us insights, knowledge, and skills that we are thankful for and will carry with us in the future.

Thank you all. Happy reading.

Henrik Christensen & Stian Eriksen

Abstract

This study dives into the relatively unexplored topic of digital assistants and the new phenomena of voice shopping. Specifically, we were interested in the different characteristics of said assistants on purchase intention. Given their relevance in this context, we also studied trust and perceived risk. We conducted a three-way between-participants experimental design with factors personality (social vs. intellectual), product involvement (low vs. high involvement), and modality (voice vs. voice and visual). With digital assistants and voice interactions being increasingly used in society, we also introduce a novel concept named "Need for Voice" (NFV) to determine whether there are idiosyncratic differences among individuals in their proclivity to enjoy and be affected by voice interactions. Through an online experiment (n = 641) we find, as expected, that low involvement products have a higher purchase intention than high involvement products. We did not find evidence in our data to suggest that personality or modality influence purchase intention. We do find initial evidence in support of a multidimensional NFV construct, however it requires additional validity checks and future testing. Our findings show that there is high risk associated with voice shopping, and that familiarity and previous experience is crucial to increase the rather low purchase intention. At this moment in time, we recommend manufacturers and third-party companies to be aware that selling low involvement products beyond the "reordering"-category might be challenging, and that normalization of the digital assistant as a sales channel will likely be of assistance. Regarding high involvement products, companies should be careful to sell these through voice shopping, as evidence in this study shows that these types of products have the lowest purchase intention. Finally, manufacturers should be cognizant of the possibility that a screen showcasing visual information in terms of a picture and text might not assist the user in voice shopping, and that future research is needed to evaluate its effect.

Table of Contents

Introduction

In the new era of technology, consumers are rapidly being introduced to cutting-edge innovations that have a significant impact on everyday life. From manual and physical to automatic and digital; people are experiencing a change in how their surroundings operate, even in their own home. Artificial intelligence (AI) and the Internet of Things (IoT) have become popular terms and are increasing in pervasiveness for the common individual. These two buzzwords have laid the foundation for the creation of the smart home, the home where everything fluidly operates the way you want it – without even having to leave the couch. Objects connected to the internet, also known as IoT objects, are fulfilling every command you make. From having your coffee poured and the shower heated to a certain temperature in the morning, to dimming the lights and playing romantic music when you come home with your loved one on a Friday evening. The advantages of IoT are "... evident and compelling" as it enables reduced costs and energy consumption as well as it increases individuals' value in terms of both entertainment and comfort (Alaa et al., 2017, p. 55). The possibilities are endless, and at this moment in time, we are only facing the end of the beginning, where the worldwide number of IoT-connected devices is stipulated to 43 billion in 2023, with predicted spending of $1.100 billion (Dahlqvist et al., 2019; Valishery, 2021). IoT and AI will thus continuously develop and expand, where experts consider this technology as essential as electricity and refrigeration (Austin, 2019).

The emergence of IoT objects is truly revolutionizing, where billions of such devices can be found in all shapes and forms across the world (Novak & Hoffman, 2019). The smart TV in your living room you yell at to change the channel, the sports watch you use to measure your heart rate and keep GPS-tracking while running, the Phillips Hue lights that simulate the colors of the northern lights in your living room, and the Tile-brick that lets you know the exact destination of where you lost your keys. However, there is one IoT object which is ubiquitous and can interconnect with all other IoT objects while serving the purpose of being your trusted companion – the digital assistant.

The most popular digital assistants that have hit the market are Amazon's Echo, Google Home, Alibaba's Tmall AliGenie, and Apple HomePod. These assistants are created not only to stand as a bridge to control other IoT objects but also contain an AI which is always there for you. Through voice commands, the digital assistant can answer literally any question which has an existing answer online, thus functioning as a provider of information. Moreover, beyond controlling other devices in your home and answering questions, digital assistants offer new opportunities for purchasing, such as ordering everything from groceries to electronics by using your voice, which is known as 'voice shopping' (Sun et al., 2019). The functionality of the digital assistant is thus multi-faceted, which appears to be appreciated by its users. This is reflected by the estimated sales growth of digital assistants from 3.5 billion in 2019 to 8 billion by 2023 (Brill et al., 2019). Furthermore, the appreciation is manifested by the fact that one year after the release of Alexa, more than 500.000 users had said "I love you" to the device (Risley, 2015). It is possible that some were simply testing Alexa's response, but as a gradually increasing number of users interact with, and are making purchases through their devices (eMarketer, 2020), it appears that people are developing a relation to digital assistants as their trusted partner. This can be strengthened by the fact that people tend to unconsciously develop intimate relationships with AI, as voices provide anthropomorphic cues (Nass et al., 1994).

Despite the optimistic sales estimation, the use of digital assistants for purchases has not developed as rapidly as previously assumed. In 2019, Taylor et al. (2019) estimated that approximately 70% of consumers would replace their visits to stores, dealers and banks with digital assistants within three years. Now, amid this period, we know that this is not the reality. However, we are seeing an annual increase, with reports based on the historical growth over the last few years estimating that 11,8% of all US digital buyers will purchase at least once through a digital assistant within 2021 (eMarketer, 2020), and that 24% of all users on a global basis had reportedly done so in 2018 (Tennant, 2018). The delayed improvement in voice shopping is partially due to the singular-sensory nature of digital assistants resulting in users being hesitant to trust the device and thus defer purchases (Munz & Morwitz, 2019). In fact, multi-sensory experiences are deemed highly influential as it is found that consumers' experiences

are derived from stimulation through multiple senses (Ackerman et al., 2010), where auditory stimuli alone are more cognitively challenging than being exposed to textual information (Munz & Morwitz, 2019). Voorveld & Araujo (2020) support this claim with their findings, where the persuasion of text-based product recommendations on smartphones was much greater than through voice presented by digital assistants. Building on the sensory nature of digital assistants, eMarketer (2020) states that:

> ... device-makers haven't fully gained users' trust. Potential buyers are still concerned when it comes to secure payments and privacy. The absence of screens on many smart speaker models is an added aspect of this problem - people often want to see products before a purchase.

From a non-user perspective, Rhee & Choi (2020) and Mari (2019) postulate that this very aspect of the absence of visual information represents a challenge to marketers and e-commerce. Conveniently, Whang & Im (2021) explain through their research that digital assistants are expected to eventually present visual information, either through connection to other devices or design changes, such as the already existing Amazon Echo Show or Google Home Hub.

Adoption of digital assistants is proven by Kowalczuk (2018), Wagner et al. (2019), Moriuchi (2019), and Martin et al. (2015) to be partly due to ease-of-use, convenience, and enjoyment, where digital assistants are mainly used for utilitarian purposes, such as checking weather forecasts, playing music or controlling other devices (Lopatvoska et al., 2019; Mclean & Osei-Frimpong, 2019; PwC, 2018). These reasons for adoption are relevant also in voice shopping, as consumers use the assistants for habitual purchases, which do not require much thought or long conversations, whereas more complicated or high involvement purchases are done through a smartphone or on the computer (Moriuchi, 2019). Hence, voice shopping does not cannibalize other purchase channels and has in fact positive spillover effects on the PC channel, where consumers spend 19,5% more money when adopting voice shopping (Sun et al., 2019). With such uplifting findings, this then raises the question as to which products are in fact purchased through digital assistants. Sun et al. (2019) found that products that do not

require active search or comparison, have low substitutability, or high purchase frequency are often purchased by users. Examples of such products are milk and coffee (Sun et al., 2019), and are categorized as typical products for reordering

The fields of digital assistants and voice shopping are relatively new and have not amassed a great abundance of research (Sun et al., 2019), making it interesting to investigate further. Previous studies of importance to these topics have examined why consumers adopt digital assistants (Mclean & Osei-Frimpong, 2019; Kowalzuk, 2018), how anthropomorphization affects user-assistant relationships (Wagner et al., 2019), auditory sense on trust and brand affect (Poushneh, 2021a), how trust is evolved (Pitardi & Marriott, 2021), customer satisfaction (Brill et al., 2019), consumer decision making (Dellaert et al., 2020), digital assistant personality (Lopatovska et al., 2019; Lopatovska et al., 2020; Lopatovska, 2020; Poushneh, 2021b), the future of voice shopping (Klaus & Zaichkowsky 2020; Klaus & Zaichkowsky 2021), implication of digital assistants in an e-commerce environment (Moriuchi, 2019), consumers' purchase and search behaviors (Sun et al., 2019), purchase behavior when presented with auditory or visual information (Munz & Morwitz, 2019), voice shopping and its effect on brands (Mari, 2019), product relevance and customer satisfaction in voice shopping (Carmel, 2019), recommendation of products moderated by involvement and personalities (Rhee & Choi, 2020), and recommendation of products through digital assistants versus websites moderated by personality style (Whang & Im, 2021).

Our study aims to contribute to the aforementioned fields and to assist both marketeers and manufacturers in gaining insights into consumer evaluations, specifically purchase intention, trust, and perceived risk, on voice shopping. This research consists of three moderators, which are deemed relevant to study as proposed by Whang & Im (2021), Rhee & Choi (2020), Klaus & Zaichkowsky (2021) and Lopatvoska et al. (2020). These are modality, involvement and personality. Elaborating, we know that visual aid helps decision-making (Munz & Morwitz, 2019) and that products that are categorized as habitual or reordering are mostly purchased (Sun et al., 2019; Moriuchi, 2019). However, we wish to not only see how visual information affects purchases compared

to voice, but how visual and voice combined influences purchase intention compared to voice alone in an AI setting.

Furthermore, we will examine product involvement, where we separate between two products (high vs. low involvement) that are not categorized as reordering. In this way, we can evaluate whether individuals would be experimental and seek out products that are slightly outside of what they are used to, and if visual properties would assist with such a purchase. The third addition to this study is personality. Digital assistants come with several forms of voices and personalities, where we follow a pervasive theoretical framework of personalities in connection to social and task-oriented interactions, related to high and low product involvement. Finally, marketing literature has over the last two decades introduced scales that capture individual differences in sensory impressions, such as Need for Touch (NFT) (Peck & Childers, 2003) and Need for Smell (NFS) (Dörtyol, 2020). We wish to build on this literature in the context of digital assistants. Therefore, we introduce the novel, exploratory scale of "Need for Voice" (NFV) to determine if there are idiosyncratic differences in enjoyment of, and susceptibility to, voice interactions.

Through this research, we contribute in several ways. First, we systematically uncover aspects regarding the consumer in the digital age of smart homes and thus provide additional knowledge in the fields of IoT and marketing. Second, we obtain a deeper understanding of how visual and voice information combined affect purchase intention, trust, and perceived risk. Third, we get to evaluate how products outside of habitual ordering, within high and low involvement, fares in the landscape of voice shopping. Fourth, we examine how personality traits dichotomized within social and task-oriented interactions influence purchases and how it interacts with other manipulations. Finally, we introduce the construct of NFV and determine whether it might have merit for future research. Our results provide theoretical implications, as well as managerial implications, which are directed towards both manufacturers of digital assistants and third-party companies. In sum, this research assists the field of digital assistants and voice shopping, while also developing a steppingstone for future research. We begin by diving deeper into the respective areas of the literature review.

Literature Review

Digital Assistants

In newer times, technology has partially or fully replaced several tasks traditionally performed by humans. Among examples of such technology are warehouse robots, educational robots, and service robots (Aeschlimann et al., 2020; Davenport et al., 2020). The most recent "robot" is the digital assistant, also known as a smart assistant, virtual assistant, or voice assistant. In essence, this is a speaker with embedded software allowing it to perform tasks by interpreting oral commands. Once the digital assistant is awakened by the user saying the key phrase, for example, "Hey Google" or "Alexa", it records the oral command and processes it in order to respond or perform the task requested by the user (Hoy, 2018).

Digital assistants are on the rise with high adoption among consumers. Today, approximately 4,2 billion digital assistants are in use globally, with this number predicted to double by 2024 (Holst, 2021). Unlike warehouse robots and agricultural robots, digital assistants are not developed to replace human labor, but rather to simply function as the personal assistant to an individual, relieving one's workload. Studies show that most people use the digital assistants for simple tasks, such as replacing the traditional search engine, keeping the calendar, checking the latest news, and playing music (PwC, 2018; Dellaert et al., 2020). However, manufacturers and companies have seen the possibility of moving beyond these basic tasks, thus evaluating new usage areas. This has resulted in the number of services compatible with oral commands being subject to rapid growth (Hoy, 2018). This does not just pertain to an increase in oral commands related to interconnectedness with other IoT objects in the home, but also to a new purchase environment where these assistants are increasingly guiding consumers' purchase decisions (Dellaert et al. 2020).

Surveys indicate that voice shopping will reach $40 billion in 2022 (Perez, 2018), where this type of shopping might make the process easier, more personalized, and hence more convenient (Bentahar, 2018; Klaus & Zaichkowsky, 2021; Tennant, 2018). Dellaert et al. (2020) explain that most shopping relates to reordering, but that the

assistant has potential for more interactive tasks and may serve as a partner for purchase decisions rather than mere order takers. Dawar (2018) illustrates this type of relationship in a futuristic scenario where the digital assistant is "every" individual's best friend. As such, it is a conversational partner and purchase optimizer based on the individual's habits and needs. Although this scenario might not take form just yet, the number of product searches by voice commands is increasing, showcasing that digital assistants can also be used as marketing tools (Smith, 2020a). The emergence of digital assistants in the e-commerce space changes how marketers and companies must consider the new purchase environment, and with it, how the digital assistant is perceived by users as a strengthening factor to purchase.

Assistant Personality

Digital assistants are perceived through voice interaction and can through this communication style express their identity and personality to a considerable extent (de Vries et al., 2009; Gartner et al., 2002). These shine through as anthropomorphic cues, which are important for relationship building (Araujo, 2018) and the behavioral intent of the user (Wagner et al., 2019). As the market of digital assistants is expected to grow in numbers and advance in abilities in future years, developers are working on identifying the right social characteristics and assistant personalities (Poushneh, 2021b). For instance, Alexa can alter her response to reflect tones and moods such as happiness and empathy (Schlosser, 2019), showing great emotional intelligence. Furthermore, it is shown that a more conversational Alexa is met with greater engagement from the user (Poushneh, 2021b). Developing the right personality is thus crucial as it affects buyer-seller relationships (Barrick & Mount, 1991), purchase intention (Poddar et al., 2009), and brand equity (Aaker & Fournier, 1995; Aaker, 1996). In fact, Bickmore & Cassell (2005, p. 11) report that "... interfaces without a face" are rated less engaging and entertaining, and so for a screenless digital assistant to engage the user, it needs to show its personality through voice. Such findings support the relevance of knowledge about the effect of personality traits in digital assistants, in

which personalities are revealed through all interactions with consumers (Plummer, 2000; McCracken, 1989).

Rosenberg et al. (1968) showed through their research a multi-faceted approach to personalities when they asked undergraduates to sort 64 personality traits into categories. Through multidimensional scaling and subsequent analyses, the authors identified two primary dimensions: social good-bad and intellectual good-bad. For the purpose of our research, the bad continuum for each dimension will not be discussed. The social-good and the intellectual-good, however, will be pursued further and will partly stand as the basis for the personalities of our research. First, social-good relates to honesty, tolerance, helpfulness, sincerity, humor, good-nature, happiness, popularity and sociability. Second, the intellectual-good relates to a personality that is determined, skillful, intelligent, serious, important, assertive, and persistent. The descriptions of social-good and intellectual-good overlap with the warmth-competence model by Fiske et al. (2007). Here, the authors state that warmth captures friendliness, helpfulness, sincerity, trustworthiness, and morality. Competence, in their model, refers to the perceived ability, intelligence, skill, creativity, and efficacy. These two dimensions have been deemed relevant as Fiske et al. (2007), Asch (1946), and Kelley (1950) explain that the warmth dimension has the greatest influence on total first impressions and thus provides information that encapsulates judgments of personality impressions that are the most quickly identified. This is supported by Cemalcilar et al. (2018), who found that warm personality traits such as openness and agreeableness are positive predictors of attraction based on first impressions, both in face-to-face and digital interactions.

Competence, in Fiske et al.'s (2007) model, came in second in how quickly people identified these impressions. As such, marketing and branding literature suggests that products should solicit high warmth and competence (Lopatovska, 2020) and it is the digital assistants' ability to signal warmth and competence that enhances trust by the user (Dellaert et al., 2020). Conveniently, the warmth-competence model has been tested by Loptavovska (2020) in a digital assistant context, where Google Home, Alexa, Microsoft Cortana, and Apple's Siri were applied. Answers to three types of

utterances were assessed, and the results showed that regardless of the four dimensions of warmth and competence, the digital assistants were rated high on the competence dimension. This is an important finding, proving that the digital assistant is innately competent as developed by its manufacturer, which is likely to be necessary for further usage by consumers. In terms of ranked importance of warmth and competence, however, Nauts et al. (2014) find that warmth does not always have primacy over competence, and that intelligence (a competence-related trait) was at least as important, and sometimes more important in shaping impressions. Hence, this shows that context will have a say in which of these personalities that means most to the user. This coincides with a study by Goetz et al. (2003) who found that users expect the personality of robots and AI to match the task context. Goetz et al. (2003) referred to serious and playful personalities that were to reflect a high and low involvement task. A similar endeavor has been conducted by Zhou et al. (2019) who examined how personality types of AI job interviewers would affect applicants' trust in these assistants. The results of both studies show that users trust assistants more in a high-risk task (high involvement) when the AI has a serious, assertive personality compared to a warm, cheerful personality.

Another example of how context affects personalities is researched by Chattaraman et al. (2019), who found that a social-oriented interaction style, entailing a warmer approach (informal, emotional support, small talk) was more effective in situations where participants found the task easier compared to those who found the task difficult. Those participants who found it more difficult were better aided by a task-oriented interaction style resembling a more competent and intelligent approach (formal, on-task dialog to achieve functional goals). Finally, Rhee & Choi (2020) tested users' attitudes towards product recommendations by digital assistants who represented two different personalities: social role (informal) versus a secretary-type agent (formal). They found that the information provided by the social role for low involvement products was sufficient and that users had positive attitudes towards this personality. There were no significant results between the formal agent and personalized recommendations of high involvement products, but the authors did witness that consumers required a heightened need for message quality.

In our research, where varying levels of involvement and risk are applied, we consider, based on previous argumentation, that context will decide on the user's preference for a certain type of personality. As such, we divide between two personalities that are deemed fit for this purpose. On one hand, the social personality, taking its inspiration from Rosenberg et al. (1968), Kelley (1950), Asch (1946), and Fiske et al. (2007), can be seen as a warm, sociable interaction style filled with chit chat, encouragement and greetings (Chattaraman et al., 2019). It is known that a social personality creates positive affective responses (Pelsmacker et al., 2013) and builds rapport with users (Lucas et al., 2014). The latter aspect is further studied by Bickmore & Cassell (2005), who explain that the purpose of social dialogue is not to fulfill certain aims in itself, but rather to put interpersonal goals in focus and task-orientation in the background. In other words, it is about caring and compassion towards the counterparty in the best interest of both parties.

On the other hand, the personality of intellectual will take inspiration from the same authors, as well as Goetz et al. (2003), Rhee & Choi (2020), Whang & Im (2021), and Zhou et al. (2019). This can be categorized as a more functional, task-focused interaction style, with the purpose of goal-oriented interaction (Chattaraman et al., 2019; Whang & Im, 2021). The two personalities will in the words of Zhou et al. (2019) reflect a friend (social) that you can imagine being close with and a counselor (intellectual) that you have not met before. Furthermore, it is important to mention that both personalities are deemed competent, as aligned with the research of Lopatovska (2020). Competence is also a dimension that is not mutually exclusive with other personalities, as an individual can be both social and competent. The personalities in a digital assistant are fully recognized based on communication that takes on oral cues, also known as voice interaction, which will be discussed next.

Voice Interaction

The process of communication between users and digital assistants happens entirely through voice, both from the user to the assistant and vice versa. As voice is presumably the most natural human communication modality (Pagani et al. 2019) and given that a

large amount of consumer interactions is handled without a human agent (Schneider, 2017), it stands as a promising field of investigation (Pagani et al., 2019; Krishna, 2019). Voice is defined by Frühholz & Belin (2018, p. 9) as:

> … an acoustic signal … registered and auditorily perceived mainly by conspecifics, and is detected, rated, and potentially classified as a distinctive vocal auditory object or as a distinctive voice feature depending on its specific voice quality compared to other auditory objects.

To clarify, our research looks mainly at voice, and not speech, where Belin (2018) explains that voice differs from speech in that it is the modality that carries speech. The two are highly connected, however, as human voice regions of the brain respond strongly to speech (Fecteau et al., 2005; Perrodin & Petkov, 2018). With voice being a natural modality for communication between conspecifics, it contains rich information in both socially relevant and person-related contexts (Belin, 2018). Furthermore, it is found to be one of the strongest anthropomorphic cues in human-machine interaction and can foster trust, affinity, and pleasure (Lee & Nass, 2004; Qiu & Benbasat, 2009; Whang & Im, 2021). This enables the user to create strong parasocial relationships with a medium, such as the digital assistant, which elicits voice interactions that are perceived as human (Whang & Im, 2021). However, as voice interaction requires one talker and one listener, the communication cues through prosody (pitch, duration, rhythm, energy) were recommended by Skantze (2016) to be implemented in digital assistants. This is aligned with Wagner et al. (2019), postulating that the more humanlike the conversation can be, the better. With Alexa now starting to show emotions through voice, such as disappointment and enthusiasm (Haselton, 2019), customer satisfaction has increased by 30%, where consumers show that they favor human-like voices compared to synthetic (Schwartz, 2019). This goes against the uncanny valley theory proposed by Masahiro Mori, explaining that the more realistic a robot becomes, the more uncanny and creepy it looks (Duffy, 2003). This theory was, however, developed with facial recognition in mind and might not be equally relevant in the context of voices.

Voice interaction between users and assistants presents challenges but also huge opportunities (Platz, 2017). Platz (2017) explains that voice interaction solves many existing issues for consumers in an easier way, such as turning on the alarm or asking questions. Voice interaction can therefore be considered beneficial when the user deems the action to be convenient (Moriuchi, 2019). But where does the complexity of the task hinder this convenience? PwC (2018) conducted a survey where 76% would rather purchase through an online store than through a digital assistant, showing that this might be the limit for many. This is likely to be based on previous usage, where digital assistants, for the moment, have their flaws and do not always achieve what the user desires, hence losing dependability (Raphael, 2019). Also, privacy concerns as well as limited knowledge of, and trust in, the digital assistants' capabilities stand as constraints to further usage (Clark 2019; PwC, 2018). The main reason, however, might be that digital assistants offer a restricted set of products inside a product category, where the information is limited compared to more sensorial enriched devices (Mari, 2019), and that comparisons between them are challenging (Munz & Morwitz, 2019). With that being said, PwC (2018) did find that 80% who did purchase through a digital assistant were satisfied, and that these individuals were more greatly inclined to purchase with the assistant in the future. This aligns with how the adoption of technology is heavily dependent on habit and experience (Venkatesh et al., 2012), and shows an upward trend for voice shopping.

Klaus & Zaichkowsky (2020, p. 393) are highly supportive of voice interactions in regard to purchases and explain that the three main reasons as to why consumers would delegate shopping to digital assistants are "... convenience and ease of use with voice, feelings of control with voice, and positive emotion with voice". All three factors relate to the importance of voice between consumer and assistant and is supported by Simms (2019), who mentions that voice shopping is quicker than any other modality and can provide a frictionless experience for the user. Klaus & Zaichkowsky (2020) also highlight that voice can increase the feeling of power and control, as well as aiding consumers with developing positive emotions, such as enjoyment, which is one of the strongest indicators of this technology's adoption and behavioral intention (Kowalczuk, 2018; Wagner et al., 2019). Power and control are in connection to the

fact that the digital assistant does not disagree nor spread negativity, where interactive voice control is limitless and thus creates a friend to rely on, engaging the user in fulfilling interactions (Klaus & Zaichkowsky, 2020). However, we do wonder whether all individuals feel that these interactions are equally fulfilling.

Voice as input is processed by our auditory sense, and automatic processing of sounds happens without conscious attention and is finely tuned by experience, where cognitive and sensorial processing are melded together (Kraus & Slater, 2016). As such, it is known that perceptual impressions of voice are largely subjective, and varying considerably, for each specific listener (Frühholz & Belin, 2018; Belin, 2018; Belin et al., 2004). Furthermore, voice provides cues to identify individuals and emotions (Whitehead & Armony, 2018; Belin et al., 2004), where the assessment of emotions is based on prosody (Anikin & Persson, 2016). However, the detection and processing of emotional information and identification also differ significantly between individuals (Belin, 2018). For instance, females are shown to be more prone to identify emotional valence when perceiving voice (Schirmer et al., 2005). Whitehead & Armony (2018) explain that these idiosyncratic perceptions must be considered in light of sex, degree of emotional intelligence, level of neuroticism and anxiety, as well as cultural background. Hence, the authors explain that while a specific stimulus can be irrelevant to one, it can indeed be very meaningful to another. For example, the understanding of personal characteristics such as trustworthiness or visualization of the speaker (e.g., over the phone) appears to differ markedly from person to person (Belin et al., 2004), where two people can hear the same voice but interpret it differently (Krishna, 2019). Belin (2018) conceptualizes all these differences by using the term ‘voice cognition’, as a way of understanding auditory cognitive abilities, including speech perception, which allows extracting information from vocal sounds like a particular sound category. With the emergence of digital assistants, which do not communicate with any other cues than the voice itself, perhaps the subjective perceptions of voice might have an impact on the relationship that can be developed.

Need for Voice

Based on the above discussion on the possible effect of voice in AI-purchase situations and how voice can be subjectively interpreted, we wish to introduce a novel NFV-scale to explore whether individuals find varying levels of enjoyment and susceptibility towards voice interactions. Consequently, this allows us to evaluate whether varying degrees of NFV has an influence in our model. Just like Peck & Childers (2003) and Nuszbaum et al. (2010) found that the higher NFT the greater likelihood there is for a person to elicit positive feelings from touching an object, we propose that higher NFV correlates with more positive feelings towards a voice-controlled interaction. NFV does not pertain to an individual's ability to formulate speech, nor how talkative a person is. Moreover, the construct is not meant to stand as mutually exclusive to other senses. In other words, being high on NFV does not necessarily deter the general need for haptic or visual information, as our senses always operate together, never independently (Knoeferle & Spence, 2021). However, an individual with high NFV might more easily settle for interactions through voice with AI, seeing that the person is more prone to enjoy voice interactions. In such regard, auditory stimulation is a decisive part of the NFV construct. The importance of auditory stimuli in marketing, for example, is highlighted by Knoeferle & Spence (2021) among others, as increased knowledge about sensory stimuli and consumer reactions to stimuli can offer new ways of marketing and creating wholesome consumer experiences.

The development of the NFV-scale is an utmost exploratory process and bases itself on a quite complex topic. Hence, it is convenient to attempt defining what we wish to measure. Voice has two sides to its existence, the production of voice and the perception of voice (Frühholz & Belin, 2018). Traditionally, these are treated as two different dimensions and can be studied as such (Kreiman & Geratt, 2018). However, in newer literature, the abilities to produce and perceive voice are considered to be closely integrated (Kraus & Slater, 2016). To understand each dimension, one must therefore understand how they are related and how they contribute to each other's function and structure (Kreiman & Geratt, 2018). To specify, "... perception cannot be understood independently from production and acoustics any more than production can be understood without attention to the listener" (Kreiman & Gerratt, 2018, p. 12). We

would argue that perception of voice, the auditory aspect, is of greatest importance as it relates directly to determining an individual's reaction to voice. However, production must also be captured to some extent, as it covers the other side of the 'coin'. Also, as voice is traditionally referred to as a carrier of a message from a speaker to a listener (Denes & Pinson, 1993), it hints at a relational aspect, which both production and perception must cover in combination. Voice, a vehicle for speech, also conveys nonverbal information such as age, identity, and emotions (von Kriegstein et al., 2003). These are elements that are important and are thus captured in some of the items in the scale. It is, however, not our scope to capture elements beyond these dimensions.

The NFV-scale has taken inspiration from NFT- (Peck & Childers, 2003), NFS- (Dörtyol, 2020) and Need for Cognition (NFC) (Cacioppo & Petty, 1982) -scales, as well as additional literature within the field of voice. The NFT- and NFS-scales are directly measuring single-sensory stimuli, tactile and olfactory, respectively. Audition is the sense measured in NFV, however, as audition entails numerous dimensions of sound, the construct of NFV has been limited to voice only. The scale is developed based on previous research that has found individual differences and aspects of both production and perception of voice. This pertains to authors prevalent in a marketing setting to cognitive brain research. Cacioppo & Petty (1982), Peck & Childers (2003), and Dörtyol (2020) started the creation of their scales through a pool of opinions or arguments, based on previous theory, experiments, and in-depth interviews. Other authors such as Richins & Dawson (1992) started the development of their materialism scale with interviews, but where they also drew from social critics and previous theory. The results, by all authors, were then used in respective factor analyses to categorize different aspects of the construct where reliability and construct validity checks were executed. Factor analysis is also conducted as part of the pretest in this study, but other checks beyond the reliability check of Cronbach's Alpha are not. Items used for the factor analysis are developed from previous studies, theories, and findings within voice and auditory interaction, where the aforementioned authors are looked to for guidance. Dörtyol (2020) obtained a 99-item scale, whereas Peck & Childers (2003) started out with a 50-item scale, Richins & Dawsons (1992) with a 30-item scale and Cacioppo & Petty (1982) used a 45-item scale before conducting their respective factor analyses.

The interval is thus relatively flexible in terms of providing a starting point that meets relevant construct checks, although this would naturally vary depending on the construct of interest.

It is important to convey that the NFV-scale, as with other scales, must be empirically tested to keep its merit if used in future research. However, Peck & Childers (2003, p. 433) refer to the following in their article on the NFT-scale: "... recently, Citrin, Stem, Spangenberg, and Clark developed a six-item scale to measure the need for tactile input. Although not formally defined as such…". Hence, we see that the exploratory introduction of a "need for"-scale can be tested without the empirical evidence, as Citrin and colleagues had done before Peck & Childers (2003) developed it further. The purpose of this research is thus to take the first step and discover whether susceptibility to voices may influence the use of digital assistants and examine whether such a scale should in fact be tested more comprehensively for future use. The items for the NFV-scale can be found in Appendix 1.

As elaborated on in previous paragraphs, the construct of NFV is about susceptibility and enjoyment towards voices in general. Hence, NFV is not about the voice of the product that is to be purchased, as most products are "dead" and not anthropomorphized through voice. In opposition, other "need for"-scales are entirely product-centric. For the NFS- and NFT-scales, Dörtyol (2020) and Peck & Childers (2003) use items where the sensorial modality is measured in connection with e.g., purchase intention of a certain product. Although we also wish to have this marketing spin on the NFV-scale, this is inherently difficult as products themselves do not utter voices. Hence, the items inspired by the already existing "need for"-scales are modified to an extent where the measurement of pre-purchase sensorial input is connected to a salesperson and not the product itself. It should be noted that a potential issue with these types of items is that they do not distinguish between whether the listener would like to purchase a product due to the pleasantness of a voice, or due to the product information that the voice of that individual carries. The NFV-scale is therefore developed to measure a general "need for", which can be used in many cases and not just purchases. However, based on the findings of Peck & Childers (2003), Dörtyol (2020), and Cacioppo & Petty

(1982) through their respective scales, there is merit in believing that an individual that is high on NFV is likely to elicit positive feelings, gain trust, and thus have a favorable intention to purchase through a digital assistant. On a closing note, NFV is not the only aspect that might contribute to explain the effect of mechanisms influencing purchase intention, as perceptions of voice can be affected by information provided by other stimuli, such as vision (Kraus & Slater, 2016).

Visual Information in Combination with Voice

We are constantly exposed to stimuli around us, from waking up in the morning until falling asleep at night. It can be sun rays creating a certain light through your window, the sound of birds outside, the smell and taste of fresh coffee, and the haptic feeling of a warm mug. Such everyday experiences are all speaking directly to our senses, and the focus on how these impacts consumers have gained much traction in marketing. In a commercial context, sensory marketing is defined as "... marketing that engages the consumers' senses and affects their perception, judgment, and behavior" (Krishna, 2012, p. 332). When two or more sensory stimuli directed at different senses are present at the same time, such as with the coffee example, it is referred to as multisensory experiences (Krishna, 2019). On this subject, it is found that multisensory exposures can improve both the isolated and the total effect of the different stimulus, thus strengthening the overall experience through providing additional value and information (Krishna, 2012; Krishna, 2019; Russell, 2002, Quittner et al. 1994; Bulkin & Groh, 2006).

In its nature, digital assistants are single-sensory devices, where voice is the only stimuli appealing to the sense that is audition. However, providing users with additional sensory stimuli such as the visual elements of a screen might prove to influence users in an AI setting. As postulated by Simms (2019), voice shopping can lead to challenges with retrieving sufficient amounts of information, and there is more speed in visually browsing on the computer. A fundamental theory on this matter is dual coding theory, which describes the individual processing and storing of auditory and visual information in two distinct brain systems (Paivio, 1971). Paivio (1971) explains that information from an auditory cue is stored in a different place than information from a

visual cue. However, they can be used simultaneously to create a holistic impression when retrieving the information. Regarding these two senses, Krishna (2012) explains that vision is found to be the dominant sense and can not only strengthen but also override the total sensory impression. This is proved through the McGurk effect, where the recipient ends up hearing what she sees through visual information and not what she actually hears through auditory stimulus (Peynircioğlu et al., 2017). However, Bulkin & Groh (2006) state that combining vision with other senses offers several benefits regarding the accuracy of sensory perception and understanding. The reason for this is that voice and vision are spatially interacted. In other words, they occur in the same space as one event, in which the two senses fulfill one another (vision can detect information that auditory perception does not capture, and vice versa) (Bulkin & Groh, 2006). Russell (2002) elaborates on this by explaining that voice tells the story in a context that is created by the visual element, in which the level of congruency between the two stimuli influences the power of persuasion. On that note, the greater the perceived congruency between voice and visual perception, the stronger is the total impression of the multisensory experience (Russell, 2002). This is in line with research on the identification of voices and senders, where seeing the face of the person speaking increases the receiver's ability to recognize the voice later on (Yarmey, 1993; Belin, 2011). In other words, adding the modality of visual to voice stimulus strengthens the total impression of the experience by maximizing information detected by the receiver.

When it comes to using digital assistants, Pitardi & Marriott (2021) explain that consumers are hesitant to purchase through the device due to not having the opportunity to see the product in question. Despite Klaus & Zaichowsky (2020) suggesting that service providers should adjust their marketing efforts from visual channels (e.g., TV, computers, phones) to auditory channels through digital assistants, eMarketing (2020) states that the absence of visual information reduces consumers' trust in the interaction and particularly in purchase situations. Even though the interaction of audio and vision is researched quite extensively as generic sensory topics, there is no prominent research with clear findings on voice and visual combined in an AI-setting accounting for voice-controlled purchases.

Involvement & Perceived Risk

Involvement as a construct has been researched extensively, where marketing literature suggests that involvement can be regarding advertisements (Krugman, 1966), products (Howard & Sheth, 1969; Hupfer & Gardner, 1971; Malär et al., 2011), or with purchase decisions (Clarke & Belk, 1979). In each of these domains, measuring involvement is done slightly differently, but all culminates in the fact that the advertisement, product, or decision must be relevant for the individual. Hence, high involvement means high personal relevance and importance (Greenwald & Leavitt, 1984; Zaichkowsky, 1985). A more formal definition of involvement is made by Zaichkowsky (1985, p. 342): "A person's perceived relevance of the object based on inherent needs, values, and interests".

A person's involvement can be seen as a continuum stretching from low to high, and concerns all forms of decision-making, as it is "... a motivational factor that affects the cognitive effort individuals expend on a problem" (Verplanken & Svenson, 1997, p. 40). As such, varying involvement levels have a significant impact on purchase decision behavior and processes (Novak et al., 2000; Liu et al., 2020). To illustrate, Fennis & Stroebe (2016) explain that low involvement decisions are characterized as relatively automatic and require little information search. High involvement decisions, on the other hand, require consumers to go into great depth before making a choice. Consider buying a soft drink, intrinsically low involvement (Drossos et al., 2014), compared to a new car, which is considered one of the highest involvement products (Richins & Bloch, 1986). While the soft drink is bought at an instant in a store, the new car will for most people require a careful and deliberate thought and search process, where various choice alternatives will be considered (Verplanken & Svenson, 1997). As found in the literature, product categories have inherently different involvement levels, which are determined based on product importance, such as level of harm, product cost and/or the length of commitment to the product (Bloch & Richins, 1983). Liu et al. (2020), for example, explain that prior studies indicate that digital and durable products with a high monetary value, complex functionality, and a long lifetime are generally in the category of high involvement. They further mention that non-durables and products such as books, groceries, CD's and other consumables are classified as

low involvement for the consumer (Liu et al., 2020). Richins & Bloch (1986) explain that this categorization of involvement can be classified further depending on a temporal dimension, however, this will not be considered here.

As with several topics within marketing, such as the brand resonance model (Keller, 2013), involvement can also be divided into cognitive and affective factors. The cognitive refers to the "think" factor, and the affective to the "feel" factor (Drossos et al., 2014). Zaichkowsky (1994) explains that cognitive involvement entails personal relevance based on the functional, utilitarian performance of the stimuli, while affective involvement is associated with the personal relevance based on feelings, emotions, and moods emerging from the stimuli. For our purpose, testing utilitarian products, we consider cognitive involvement as the most relevant, yet both dimensions are important and could occur together (Park & Young, 1986). As this research looks at products and purchase intention through digital assistants, product involvement and its purchase decision will be discussed in the following.

Product involvement has been dichotomized into enduring and situational involvement (Hong, 2015). Richins & Bloch (1986) explain that enduring involvement represents an ongoing concern with a product that lasts over a long period of time, independent of the purchase decision. Situational involvement occurs only in specific situations, is highly related to the purchase decision, and diminishes slowly after the purchase (Richins & Bloch, 1986). One can imagine the difference between buying a smartphone and having an enduring involvement with this smartphone over a year, compared to the potentially ambivalent purchase decision when the smartphone is bought. To elaborate, Venkatraman (1989, p. 230) defines situational involvement as the "... degree to which a consumer is motivated to avoid the negative outcomes of a product purchase …". As such, situational involvement is highly similar to purchase decision involvement, understood as interest and concern an individual bears upon a purchase decision (Mittal, 1989). In the following, it is this situational involvement and the 'interest and concern' with the purchase decision that are appropriate for our research.

The importance of involvement has started being evaluated in the context of digital assistants and purchase environments (Rhee & Choi, 2020; Klaus & Zaichkowsky, 2021). This is appropriate as involvement moderates purchase intention, where this is influenced by a medium's inherent limitation to convey sufficient information to support highly cognitively involved decision making (Drossos et al., 2014). As such, voice shopping is celebrated for making it easier to purchase low involvement products (Mari, 2019), where low cognitive involvement has a positive relationship with ease-of-use (Smith et al., 2013). However, the information delivery through voice interaction between users and digital assistants has limitations, perhaps especially in terms of quantity, the processing time of information for the user, information discrepancy, and visual representation of products - complicating high involvement purchases (Carmel, 2019; Mari, 2019; Whang & Im, 2021). An example of this is posited by Klaus & Zaichkowsky (2020), who explain that voice interactions are on a single temporal dimension. Communication with digital assistants about the same subject over a longer period of time limits the users' possibility of comparing alternatives and remembering previous information due to consumers' limitation of retaining information in short-term memory (Bjork, 1970). This aspect is important to consider in a purchase decision with situational involvement, as individuals are unlikely to obtain much product information before purchasing, which increases uncertainty and makes it necessary to evaluate risk in the purchase decision (Bateman & Valentine, 2019; Venkatraman, 1989).

This naturally shifts the discussion to perceived risk, or uncertainty, which are deemed as equivalent constructs (Taylor, 1974), where perceived risk is a consequence of involvement level (Delgado-Ballester & Munuera-Aleman, 2001; Venkatraman, 1989). This is well explained by Dholakia (2001, p. 1343): "Perception of risk makes the importance of the purchase, and through that means, the importance of the product class, salient to the consumer". Kim et al. (2008, p. 546) refer to perceived risk in an online context as "... a consumer's belief about the potential uncertain negative outcomes from the online transaction". In this research, we focus on perceived risk, or uncertainty, mostly with the outcome of a purchase decision, as this requires the acquisition and handling of information to mitigate (Taylor, 1974). However,

uncertainty with the consequences, where the user might experience a mistake/loss (Taylor, 1974), is also present and is often where the individual defers the decision.

To highlight uncertainty with outcome and consequences pertaining to purchase decisions, there are two main aspects of risk that we deem important. These are the perceived risk inherent with the product class (Dholakia, 2001), and the perceived risk associated with the notion of uncertainty with the medium (Bianchi & Andrews, 2012; Andrews & Boyle, 2008). These two risks correlate positively when making a decision. For example, consider purchasing a computer, a high involvement product (Stewart et al., 2019; Liu et al., 2020), through a digital assistant. The uncertainty regarding the outcome could be "if I buy this product, how can I be certain it arrives on time?" or "if I buy this product, how can I be certain it is the computer for me?", whereas uncertainty with the consequences might be "if I buy this product, and it gets delivered to the wrong address or does not perform as specified, how would that affect my situation?". Several authors have highlighted different dimensions of perceived risk with a purchase decision, such as financial/economic, performance, physical, social, product category, delivery, psychological, and time loss (Jacoby & Kaplan, 1972; Kaplan et al., 1974; Taylor, 1974; Peter & Ryan, 1976; Stone & Grønhaug, 1993; Hong, 2015; Bhatnagar et al., 2000). Simply, a high involvement product will be rated high on some or all of these dimensions, is thus associated with high perceived risk, and requires more effortful and informational-driven decision-making (Stewart et al. 2019). As such, the computer might have high financial and performance risks inherent to the product. However, the risk with the medium becomes heightened due to the risk of the product. These can be delivery risk (Hong, 2015), transaction risk (security and reliability of the transaction) (Biswas & Biswas, 2004), or product risk (inability to physically inspect the product) (Bhatnagar & Ghose, 2004).

Taylor (1974) explains that perceived risk can be mitigated by handling and acquiring information. In terms of acquiring and handling information that lays the ground for decision making, one can consider the elaboration likelihood model (ELM) by Petty et al. (1983), which has been verified and highly used in research. These authors postulate that high involvement products require central cues and strong, persuasive arguments

relevant for the individual, while peripheral cues, such as entertaining, repetitious messages are sufficient for low involvement products. In terms of e-commerce, one major part of the escalated perceived risk is due to higher uncertainty, as the medium cannot provide an examination nor visualization of a product (Hong, 2015; Mari, 2019). Hence, strong, persuasive cues can thoughtfully be difficult to render. Extrapolating the research on online shopping to voice shopping, which seems suitable albeit more challenging for voice, information asymmetry and concerns regarding delivery can pose increased perceptions of risk and reduce purchase intention (Verhagen et al., 2006; Hong, 2015). There are thus several aspects that potentially hinder the user from gaining the needed information to make a sound decision. Finally, not trusting the information or the lack of information that is being delivered can heighten perceived risk. Pitardi & Marriott (2021) highlight that the omission of trust in digital assistants is one of the main barriers to voice shopping. This is important, as trust is necessary for a high involvement purchase (Delgado-Ballester & Munuera-Aleman, 2001).

Trust

In psychology and marketing literature, trust is a well-researched subject. Morgan & Hunt (1994) define trust as the existence of certainty in the integrity and reliability of another party, such as a person or a company, in this case, a digital assistant. This definition is further elaborated on by Rousseau et al. (1998, p. 395), who define trust as "... a psychological state composing the intention to accept vulnerability based on expectations of the intentions or behavior of another". For digital assistants in specific, trust is highly connected to the humanlike presence of the device (Klaus & Zaichkowsky, 2021) and is a key element with human-technology relationships, as "... systems that are not trusted are not used" (Schaefer et al., 2016, p. 393). Pitardi & Marriott (2021) find that the prominent antecedents of trust in digital assistants are social attributes, social presence, and social cognition, in other words, its personality and perceived competence. Here, personality is an important element for creating the relationship between the user and the device, as well as reflecting the competence of the device to conduct different tasks. Overall, the concept of trust is central for

establishing and preserving transactional relations, such as consumer relationships, commitment to the relations, decrease risk and uncertainty between parties, as well as improving consumer willingness-to-pay, and adoption (Wang et al., 2015; Bejou et al., 1998; Kaasim & Abdullah, 2010; Dinev & Hart, 2006). Building on these findings and the definition of Rousseau et al. (1998) in the context of digital assistants, trust appears to be a decisive factor of purchase decisions, as users need to be able to rely on the medium and its product recommendation as well as its capabilities to complete the purchase correctly (Hsiao et al., 2019).

On that note, previous research (Mayer et al., 1995; Rempel et al., 1985; Martin et al., 2015; Jarvenpaa et al., 1999) has found that perceived risk and trust are intercorrelated. Blomqvist (1997) explains that under the traditional economic assumption of perfect information, there is no perceived risk and thus no trust, but only rational calculation. However, as this assumption is not the reality but simply describes a utopian situation, there will be a minimum level of uncertainty and therefore an inherent need for trust. As such, trust is the key to accept a situation or decision that entails risk (Blomqvist, 1997). Blomqvist (1997) further explains that providing consumers with more information should reduce uncertainty and therefore also the importance of trust. For manufacturers of digital assistants, this implies that increasing the amount of information (e.g., combining the modalities of voice and visual) can mitigate the perceived negative risks in terms of outcome and consequences by users who do not fully trust the device. The need for information is particularly necessary in purchase decisions of increased involvement, where the effect of trust on subsequent consumer behavior such as purchase intention becomes more prominent (Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021). Through familiarity with the device, however, the user is likely to experience reduced uncertainty (Miyazaki & Fernandez, 2001), which provides a greater possibility to establish trust (Lee & Turban, 2011), which in turn is a key aspect in building a strong relationship between the user and the digital assistant (Delgado-Ballester & Munuera-Alemán, 2001).

At the current point in time, there appears to be conflicting findings amongst the relatively little body of literature on trust in the context of digital assistants. Even

though Dellaert et al. (2020) find that users, in general, tend to trust their digital assistants, there are no findings on whether this will make users rely on the devices for making important decisions, complete high involvement purchases, and carry out complex tasks. Furthermore, Cowan et al. (2017) and Luger & Sellen (2016) propose that users do not trust voice commands to a digital assistant such as sending a text message or dial a phone call, as they expect visual feedback before executing the task. According to Lopatovska et al. (2020), however, people do in fact blindly trust their digital assistants in other situations than those proposed by Cowan et al. (2017) and Luger & Sellen (2016). This pertains to marketing, recommendations, and information related to voice search (e.g., asking about the weather). In these cases, users are not critical to neither the information nor the service provided, not to mention the intention of the assistant (Lopatvoska et al., 2020). Further, as Cassell (2001) proposes that trust is in fact a significant aspect in establishing a relationship between humans and technology, it appears to be a vital construct of study in the context of digital assistants.

In this regard, Foehr & Germelmann (2020) find four paths to how consumers develop trust in digital assistants. These rely on (1) friends and family as a reference point to trust, (2) by using partners as proxy for relation-based trust, (3) transferring the trust from the technology producer to the digital assistant itself, and (4) anthropomorphization where consumers establish close relationships to digital assistants based on the perceived personality of the device. The three former paths evolve around familiarity and social connections, while the latter path builds trust in direct interaction with the device. Consumer interactions with digital assistants might lead to perceptions of commonality, where Novak & Hoffman (2019, p. 222) state that "... consumers may be more likely to trust an anthropomorphized object because, since it is like us, it must share our motives and goals". This is in line with Wagner et al. (2019) and Reeves & Nass (1996) who find that the greater fit there is between user personality and device personality, the greater is the chance that the user will like the device and be influenced by it. However, it is not given that trust in digital assistants is a consequence of similarity reflected through anthropomorphism. If that in fact is the case, then that trust derived from anthropomorphism might not be particularly strong compared to trust based on experience with the device in question (Culley &

Madhavan, 2013). Tegmark (2017) supports this by stating that the central element on the path to the development of trust is common goals and values rather than anthropomorphizing the personality of the device. This is in line with the concept of value similarity (Earle & Cvetkovich, 1995), in which people tend to trust counterparties of similar salient values as a means to reduce the risk that the counterparty will act differently from oneself (Siegrist et al., 2000).

Regardless of how trust is developed, once the user has attained a sufficient amount of trust in the device, Pitardi & Marriott (2021) suggest that this might prove to positively influence purchase intention in the context of voice shopping. Furthermore, Delgado-Ballester & Munuera-Alemán (2001) and Klaus & Zaichkowsky (2021) address the need to study trust in high involvement contexts, as that is when the effects of trust truly become apparent. Therefore, it appears expedient to connect digital assistants and voice shopping to the concept of trust and doing so in the presence of new variables not included in previous research.

Purchase Intention

Just like involvement level is an antecedent of trust (Mayer et al., 1995; Rempel et al., 1985), trust is an antecedent of purchase intention (Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021; Ha et al., 2014; Yoon, 2002; Lu et al., 2016). Beyond the use of digital assistants as helpers to carry out everyday tasks at home, the devices can function as personal shoppers. Today, most purchases conducted through digital assistants are low-involvement routine purchases such as reordering of groceries (Hoy, 2018). However, there is still an unfulfilled potential for high involvement products from the manufacturers' and third-party companies' point of view, and a general lack of research on high involvement purchases through AI (Delgado-Ballester & Munuera-Alemán, 2001; Klaus & Zaichkowsky, 2021). Moreover, Moriuchi (2019) found that people who purchase through digital assistants use the device more in general compared to users who do not purchase through the device. Thus, examining the construct of purchase through voice-controlled AI appears to be of great importance in order to understand the conditions leading to improved purchase intention.

Purchases made through digital assistants are what Lu et al. (2016) refer to as social commerce. This construct entails social technologies, community interactions, and commercial activities (Lu et al., 2016), meaning that consumers need to experience the element of human presence, in this case through anthropomorphism, when interacting with the digital assistant. The concept is based on Social Presence Theory (SPT), proposing that IT-enabled human-like interaction (Pavlou et al., 2007) and personalized greetings (Gefen & Straub, 2004) evoke the feeling of the counterparty being socially present (Lu et al., 2016), consequently developing parasocial relationships such as with digital assistants (Whang & Im, 2021). According to Dellaert et al. (2020, p. 2), digital assistants "... have the potential to be more interactive, include experiential service purchases (e.g., which restaurant to visit), and serve as partners in decision dialogs rather than mere order takers". They can display empathy with the current emotional mood of the user, and for instance suggest a low-calorie food item and emphasize this fact in the oral response when being aware that the user seeks to lose weight (Dellaert et al., 2020). Thus, purchase intention may also be influenced by the types of alternatives provided by the assistant and the emphasis on certain attributes of the alternatives. Dellaert et al. (2020) suggest that this is particularly relevant in cases of low autonomy. This is in line with Castelo et al. (2019) and Newman et al. (2020) who find that users in fact prefer AI over humans for conducting objective tasks. On one hand, based on the above discussion, the treatment condition of low involvement in our involvement-variable might prove to influence purchase intention positively, as low involvement interactions do not require extensive thinking and decision making, thus leaving the user with less autonomy. On the other hand, high involvement purchases entail a more cognitively demanding thought-process, requiring higher autonomy and thus complicating the purchase decision.

Within SPT, as social presence is perceived through the sense of human interaction, warmth, and the sensitivity of the medium (Rice & Case, 1983), digital assistants serve to fill the alleged void of lack of human and social cues in traditional online shopping (Lu et al., 2016). However, Poushneh (2021a, p. 7) states that "... customers would not like to encounter a surprise when they are serious about taking a purchase action". Thus, a digital assistant simply providing social presence is not a sufficient motivator

of purchase intention without the user trusting the assistant to be a reliable companion. This is in line with the findings of Blomqvist (1997) regarding the effect of increased information on the importance of trust. Following this, Lu et al. (2016) call for the phenomenon of social commerce to be subject to greater attention in future studies. Accordingly, introducing purchase intention as the main dependent variable in our model allows us to study whether users' intention to purchase through digital assistants is influenced by the treatment conditions, and if so, under which circumstances users experience the greatest purchase intention in voice shopping.

Hypotheses Development

The scope of this research is to examine how purchase intention, our main dependent variable, is affected when moderated by modality (voice vs voice and visual), personality (social vs. intellectual) and product involvement (high vs low). We also measure trust and perceived risk, which are dependent variables, but treated also as covariates. Eight hypotheses are linked to this framework. Finally, a covariate and independent variable is the novel concept of NFV, which relates to users' susceptibility and proclivity to gain enjoyment toward voice interactions. Two research questions are developed for this construct, where Table 1 displays an overview of all hypotheses and research questions.

In our research, we do not look at voice alone, but at how personality shines through the voice (Campbell & Pennebaker, 2003). Here, we divide between two personalities, one that is warm and socially oriented, and one that is intellectual and task-oriented. It is recognized in previous research that digital assistants and machines should follow a warm and competent blueprint (Lopatovska, 2020), as this builds rapport with the user (Lucas et al., 2014), creates engagement (Poushneh, 2021b), and enhances trust (Dellaert et al., 2020; Qiu & Benbasat, 2009). Trust is known as an antecedent to reduce risk and thus increase purchase intention (Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021; Ha et al., 2014; Yoon, 2002; Kim et al., 2008), and is developed when the interaction between user and digital assistant feels human (Klaus & Zaichkowsky, 2021). As such, a warm personality is known to generally create more engagement and have positive perceptions (Pelsmacker et al., 2013; Niculescu et al.,

2013) than an intellectual personality that is not as conversational (Chattaraman et al. 2019). The first main effect is thus:

H1: *A digital assistant portraying a social personality will lead to higher purchase intention compared to an intellectual personality.*

Regardless of personality, however, voice interaction has its limitations. This regards information discrepancy, the processing time of information, and quantity (Carmel, 2019; Mari, 2019). Klaus & Zaichkowsky (2021) postulate that humans might have issues comparing information in the marketplace through voice alone. Hence, providing users with a more wholehearted experience providing additional sensorial stimuli can increase the persuasiveness of the message and increase trust due to the effect of multiple sources (Lee & Nass, 2004). Furthermore, it can strengthen the experience, providing additional information and value (Krishna, 2012; Krishna, 2019; Russell, 2002, Quittner et al. 1994; Bulkin & Groh, 2006). This is essential as product information is found to be important for consumers in online shopping (Cho & Sagynov, 2015). Thus, we rely on the second modality of vision combined with voice to provide additional information, which should increase trust and reduce risk (Blomqvist, 1997; Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021), having a positive impact on purchase intention. Therefore, the second main effect is:

H2: *Combining voice and visual information will increase purchase intention compared to voice stimuli only.*

Even though voice shopping is not prevalent for the time being, it is predicted to grow substantially in the coming years (Tennant, 2018; Klaus & Zaichkowsky, 2020; eMarketer, 2020). Those who do enjoy purchasing through digital assistants, however, mostly purchase low involvement products for the time being (Hoy, 2018; Mari, 2019) as it is found to have increased convenience and ease-of-use (Tennant, 2018; Klaus & Zaichkowsky, 2020). The vast literature on the topic of involvement postulate that low involvement products do not require much acquisition of information regarding the

product as the perceived risk of outcome and consequences of the purchase is of low relevance to the individual (Taylor, 1974; Verplanken & Svenson, 1997). This logical sentiment is opposite for high involvement products, as the outcome of the decision can have highly relevant consequences. Hence, the individual wishes to ascertain the necessary information to reduce uncertainty with their decision (Taylor, 1974). The division of low and high involvement follows the ELM by Petty et al. (1983), where only information deemed as peripheral cues are necessary for the former, but central and important arguments are paramount for the latter. Based on this, the third main effect is as follows:

H3: *Low involvement products will have a higher purchase intention compared to high involvement products.*

Even though socially-oriented personalities that are rated high on warmth and competence have positive effects on the user and affect trust-building (Keeling et al., 2010; Chattaraman et al., 2019), research has shown that people expect the faced personality to be congruent with the task context (Goetz et al., 2003). Resembling propositions have been found by other authors such as Zhou et al. (2019), Chattaraman et al. (2019), Rhee & Choi (2020), and Keeling et al. (2010) who propose that the interaction must match the product/service, as social interaction can raise suspicion about retailer motives or not provide the right benefits. In our research, we vary the different situations based on the level of involvement and perceived risk among products and the medium itself. As high involvement products require more information search and elaborate decision making (Zaichkowsky, 1985; Taylor, 1974), these types of products need central cues with strong arguments (Petty et al., 1983). Hence, an individual who perceives higher amounts of risk with the purchase decision, and requires information to reduce uncertainty (Taylor, 1974; Blomqvist, 1997; Drossos et al., 2014), should be more inclined to be assisted by an intellectual personality. This is due to its communication form matching the task context and providing essential information through prosody and arguments that are helpful for the purchase decision. This is required seeing that perceived risk immobilizes the decision-making as the outcome can have negative consequences that are of high self-relevance

(Zaichkowsky, 1985). Having trust in the AI interface is significant for both high and low involvement purchases (Klaus & Zaichkowsky, 2021), but this is especially true for the former (Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021), and should be provided by the intellectual, task-oriented personality as it matches the sought-after benefits from the user (Chattaraman et al., 2019). The same amount of trust is not necessary for low involvement products where a warm and social personality should suffice to deliver the peripheral cues and needed information (Petty et al., 1983; Mari, 2019). Hence, when the personality matches the task context, this should increase persuasion power and trust, and thereby purchase intention. The first two-way interactions are:

H4: *For high involvement products, a digital assistant portraying an intellectual personality will increase purchase intention compared to a digital assistant with a social personality.*

H5: *For low involvement products, a digital assistant portraying a social personality will increase purchase intention compared to a digital assistant with an intellectual personality.*

Adding visual as a multisensorial aspect to this setting should provide additional information to the user giving more persuasion power (Lee & Nass, 2004; Krishna, 2012; Krishna, 2019; Russell, 2002, Quittner et al. 1994; Bulkin & Groh, 2006). As individuals require more information in high involvement decisions, where many want to see the product before purchase (Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021), the added sensorial stimuli should increase trust, reduce perceived risk and assist the purchase decision (Mayer et al., 1995; Rempel et al., 1985). Low involvement products require less information, where visual information might not be necessary as voice has proven to aid sufficiently for low involvement products (Mari, 2019). However, as there is a difference between enduring and situational involvement and the inherent need for information (Venkatraman, 1989), and as a purchase decision does entail a certain level of risk, it appears likely that voice

and visual information combined will be helpful also for low involvement products. Thus, the final two-way interaction is as follows:

H6: *For both high and low involvement products, combining voice and visual information will increase purchase intention compared to voice stimuli only.*

Personalities can be identified by visual and oral cues (Rhee & Choi, 2020; Whitehead & Armony, 2018). However, the visual information provided in our study is not developed to manifest any such personality, but rather to provide added information related to the purchase. We have also not found any support in the literature that a social versus intellectual personality, combined with the specific modality of voice and visual, will influence purchase intention. This then falls back to the positive effect of a social, warm personality (Dellaert et al., 2020), which is found to create trust (Keeling et al., 2010) and is thought to be more beneficial in general than an intellectual personality (Chattaraman et al., 2019; Zhou et al., 2019). However, without any form of involvement to moderate the interaction, this assumption falls in line with H1, the main effect, and no two-way interaction is expected between personality and modality.

The final hypotheses will take all constructs into account, and thus potential three-way interactions are presented. As mentioned, high and low involvement products differ in perceived risk (Venkatraman, 1989) and hence also the information that is required to make a decision due to uncertainty (Taylor, 1974; Blomqvist, 1997). With high involvement products, more robust and strong arguments are needed, as the individual is more cognitively involved (Drossos et al., 2014) compared to low involvement products (Petty et al., 1983). We assume, based on literature, that visual information will increase the level of information the user receives, increasing trust and reducing perceived risk (Mayer et al., 1995; Rempel et al., 1985; Drossos et al., 2014), compared to voice-only due to its limitations (Klaus & Zaichkowsky, 2020; Klaus & Zaichkowsky, 2021; Mari, 2019; Carmel, 2019). This is also based on Drossos et al. (2014) who suggest that involvement and purchase intention is influenced by a medium's ability to convey sufficient information, which is limited in the case of digital assistants.

Furthermore, an individual who is presented to a high involvement product requires elaborate decision-making processes (Zaichkowsky, 1985; Taylor, 1974) where the task context is different from a low involvement product. As aligned with Poushneh (2021b), Goetz et al. (2014), Zhou et al. (2019), Chattaraman et al. (2019), Keeling et al. (2010), and Petty et al. (1983), these situations have proved to be better suited with a personality that meets the expectations of the user, and provide the user with strong, central arguments that can directly assist the user with the decision. In terms of low involvement products, however, the same requirements of argument strength are not needed, and a warm, socially-oriented personality is deemed to be more beneficial than an intellectual, task-oriented personality (Keeling et al., 2010; Dellaert et al., 2020). Hence, trust is increased when the appropriate personality matches the benefits needed by the user (Chattaraman et al., 2019). Regardless, the additional visual information in coexistence with voice should be more beneficial than voice alone in all cases. However, it is particularly necessary in the case of high involvement decision making, in order to cope with greater perceived risk compared to low involvement (Drossos et al., 2014; Taylor, 1974). Based on the argumentation, the three-way hypotheses are developed:

H7: *For high involvement products, a digital assistant portraying an intellectual personality while combining voice and visual information will increase purchase intention, compared to a social personality and/or voice only.*

H8: *For low involvement products, a digital assistant portraying a social personality compared to an intellectual personality, voice stimuli will increase purchase intention, and combining visual information with voice will strengthen this effect.*

Conclusively we have introduced the new concept of NFV. This is an overarching concept that we believe might have equivalent effects as proposed by NFT (Peck & Childers, 2003), NFS (Dörtyol, 2020), and NFC (Cacioppo & Petty, 1982). This entails an enjoyment with voice, increasing the propensity for an individual to operate with voice interactions, as based on a "... preference for the extraction and utilization of

information" (Peck & Childers, 2003, p. 435). As such, this concept is tested within our theoretical framework and as we know that enjoyment stands as a basis for usage (Kowalczuk, 2018; Martin et al., 2015; Wagner et al., 2019), and that extraction of information is necessary to reduce risk and increase trust, we suggest that those high on NFV will have an increased purchase intention. Still, as we have no theoretical basis as a foundation for the effect from NFV, two research questions related to our dependent variables are developed:

RQa: *Those who are high on NFV will find voice interactions more enjoyable, increasing purchase intention compared to those who are low on NFV.*

RQb: *Those who are high on NFV will have an increased level of trust, hence decreasing perceived risk compared to those who are low on NFV.*

Table 1: Overview of Hypotheses & Research Questions.

# Hypotheses / RQ	
Hypothesis 1 Main effect	A digital assistant portraying a social personality will lead to higher purchase intention compared to an intellectual personality.
Hypothesis 2 Main effect	Combining voice and visual stimuli will increase purchase intention compared to voice stimuli only.
Hypothesis 3 Main effect	Low involvement products will have a higher purchase intention compared to high involvement products.
Hypothesis 4 Two-way interaction	For high involvement products, a digital assistant portraying an intellectual personality will increase purchase intention compared to a digital assistant with a social personality.
Hypothesis 5 Two-way interaction	For low involvement products, a digital assistant portraying a social personality will increase purchase intention compared to a digital assistant with an intellectual personality.
Hypothesis 6 Two-way interaction	For both high and low involvement products, combining voice and visual information will increase purchase intention compared to voice stimuli only.
Hypothesis 7 Three-way interaction	For high involvement products, a digital assistant portraying an intellectual personality while combining voice and visual information will increase purchase intention, compared to a social personality and/or voice only.
Hypothesis 8 Three-way interaction	For low involvement products, a digital assistant portraying a social personality compared to an intellectual personality, voice stimuli will increase purchase intention, and combining visual stimuli with voice will strengthen this effect.
Research Question (A)	Those who are high on NFV will find voice interactions more enjoyable, increasing purchase intention compared to those who are low on NFV.
Research Question (B)	Those who are high on NFV will have an increased level of trust, hence decreasing perceived risk compared to those who are low on NFV.

Methodology

Experimental Design

In order to test our hypotheses and research questions, this study utilized a 2 (personality: social vs. intellectual) x 2 (modality: voice vs. voice and visual) x 2 (involvement: high vs low) between-participants design, in which data was gathered through a randomized questionnaire-based online experiment using Qualtrics. Figure 1 displays the visualization of the experimental model. The main dependent variable was purchase intention, while NFV, perceived risk, and trust were tested as dependent variables but were also added as covariates to control for when analyzing the relationship between the independent variables and purchase intention. For both the pretests and the main study, respondents were anonymous and no personal information was collected.

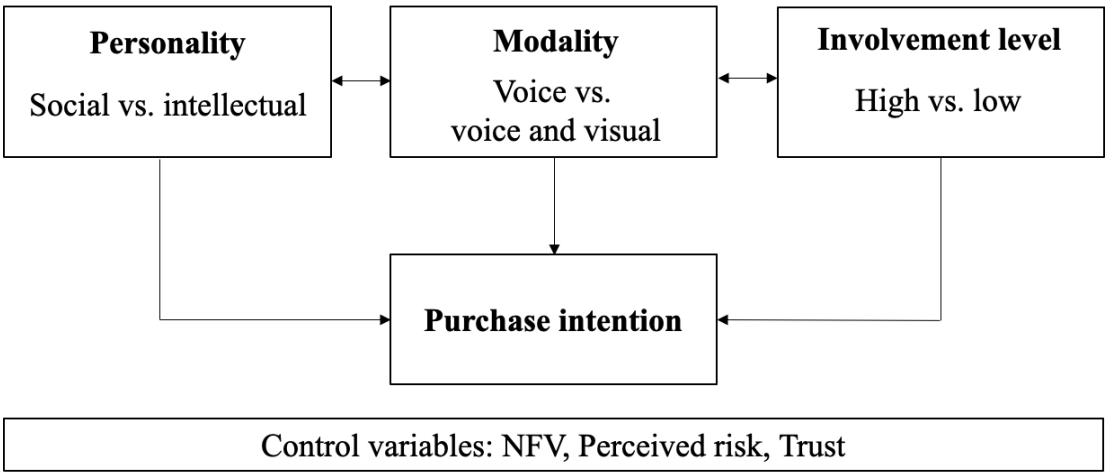


Figure 1: Visualization of Experimental Design.

Pretests

Prior to conducting the main study, we needed to 1) pretest the suggested items of the NFV-scale in order to identify the correct number of items loading on relevant factors, and 2) confirm that the manipulated personalities for the interactions were in fact perceived as intended, while also determining which gender reflected the personalities best. Both pretests were conducted using Prolific Academic (https://www.prolific.co),

an online survey participant database, and participants were from the US only, as we wanted them to be native English speakers from the western world to ensure that no cultural differences would impact the results.

Pretest 1: NFV-scale

We began the construction of NFV by developing items that included voice, both production and perception. The scale was based on the existing scales of NFT, NFC, and NFS, as well as voice-related research. In addition to general sounds, Belin et al. (2004) found that voices also carry sounds beyond speech, such as laughter, cry, screams and moans. However, even though voice can be interpreted as a sound, we found these items to be too distant from the nature of voice-based communication between a user and a digital assistant. We also excluded other general sounds, as the only source of sound in this particular context is the voice from speech originating from the device. As per the discussion of NFV in the literature review, inspiration for items being product-specific have mostly been omitted as products do not utter voices. Yet, a few were included but modified, such as "I like to listen to the pleasant voice of a salesperson even though I have no intention of buying a product", where the original item is: "I like to touch products even if I have no intention of buying them" (Peck & Childers, 2003, p. 432). Thereafter, we ended up with a 34-item scale including items only pertaining to voice (Appendix 1). All items followed a 7-point Likert scale (1 = strongly disagree to 7 = strongly agree) as in the case of NFT (Peck & Childers, 2003). As the purpose of the NFV-scale in this study was to build an initial understanding of the phenomena of voice in the context of digital assistants, and not to stand as an empirically absolute construct, we only conducted one internal validity check.

The survey was completed by 105 respondents who were exposed to the 34 items in a randomized order. Five respondents were excluded from the data, as we deemed their completion time too short for having devoted sufficient attention and answered thoughtfully. This resulted in a total sample of N = 100 (gender: male = 45, female = 53, Other = 2; age: 18-29 = 34, 30-39 = 41, 40-49 = 12, 50-59 = 7, 60+ = 6), where the average completion time was four minutes and 57 seconds, and each participant was paid £1.67. The survey-data was used in an Exploratory Factor Analysis (EFA) with

Maximum Likelihood extraction and oblique (Promax) rotation. The extraction method is one of two methods to obtain "best" results (McDonald, 1985). Oblique rotation eliminates spurious effects and is a more sensible approach, and EFA is deemed suitable when there is no a priori model that the items are based on (Fabrigar et al., 1999). An EFA is also stricter in terms of communalities and factor loadings, showcasing lower numbers, developing the possibility that another outcome could occur compared to a Principal Component Analysis (PCA) (Fabrigar et al., 1999). An issue with EFA, however, is to determine when to stop and be content with the number of common factors extracted (Fabrigar & Wegener, 2011). As such, we follow Fabrigar & Wegener's (2011) four criteria to establish such a stopping point. These are: 1) when the model does a good job accounting for correlation among variables, 2 & 3) a model with fewer factors would do substantially worse, but a model with an extra common factor would not do appreciably better, and 4) all common factors are readily interpretable and can be related to theoretical utility.

With the criteria of Fabrigar & Wegener (2011) in mind, the number of common factors was determined on the basis of theoretical footholds, such as Kaiser's Criterion of Eigenvalue above one, Parallel analysis, Maximum Likelihood tests, the Scree Plot, total variance explained and omission of variables with low communalities, as well as assessing over- and under-factoring and considering that there must be a common theme to all variables that are substantially loading to the same factor (Fabrigar & Wegener, 2011; Mooi et al, 2018). In terms of communalities, it is referred to that everything below .50 is deemed poor (Mooi et al., 2018) while Fabrigar & Wegener (2011) explain that moderately good conditions for EFA entail communalities of .40 to .70 with at least three measured variables loading on each factor, where factor loadings < .40 are deemed poor. The EFA is thus quite an experimental and stepwise process where an intuitive understanding of the factors is appropriate.

Results Pretest 1: NFV-scale

Running the EFA stepwise and following the aforementioned theoretical guidelines resulted in a three-factor solution with three items loading on each factor. Table 2 displays the factors and the loadings of the respective variables. The original 34 items

were thus reduced to nine and based on the content of the items, these factors, which are all related to the perception of voice, are labeled: 1) “Voice & Hedonic Properties” 2) “Inference Based on Voice”, and 3) “Attribution by Voice” (see Appendix 1.2 in relation to Appendix 1.1). This three-factor solution proved to be an adequate solution, also when conducting a PCA as well as Principal Axis Factoring to confirm the EFA (Fabrigar & Wegener, 2011).

Table 2: Structure Matrix EFA.

Items	Factors		
	1	2	3
When I listen to a pleasant voice, it elicits positive emotions	.856		
Listening to a pleasant voice is delightful	.871		
I experience a sense of satisfaction when I hear certain voices	.683		
I tend to be able to assess personal characteristics of the speaker when listening to that person's voice		.678	
I can infer the emotions that a person is expressing through their voice		.746	
I tend to be able to infer how people feel based on their voice		.858	
When I listen to a pleasant voice, it increases my trust in the person who speaks			.580
I tend to desire that a person I talk to has a voice that is congruent with their identity			.777
My attention towards a person increases when their voice appears congruent with their personality			.747

Extraction Method: Maximum Likelihood.

Rotation Method: Promax with Kaiser Normalization.

The data output showed a satisfactory KMO measure of sampling adequacy (.731) and Bartlett’s test ($p < .001$), representing that factor analysis is appropriate. Furthermore, all communalities were sufficient, at a moderately good condition from .400 up to .809

(Fabrigar & Wegener, 2011), except one item ("When I listen to a pleasant voice it increases my trust in the person who speaks") at .375. However, this item showed a decent loading on its respective Factor 3: "Attribution by Voice", indicating that it contributes to a factor of acceptable internal reliability (Malhotra et al. 2010). On the notion of reliability, Factor 1: "Voice & Hedonic Properties" (Cronbach's α = .825) and Factor 2: "Inference Based on Voice" (Cronbach's α = .783) also proved to be reliable, as well as the fact that all the three factors together resulted in a reliable measure of the construct of NFV (Cronbach's α = .775). The factors were also relatively uncorrelated (1 and 2: .268, 1 and 3: .341, 2 and 3: .378), indicating that they do not overlap and thus in fact measure different dimensions of the construct. In total, the factor model accounted for 72.48% of the variance in the data, in which Table 3 displays the explained variance.

Table 3: Total Variance Explained.

Component	**Initial Eigenvalues**			**Extraction Sums of Squared Loadings**			**Rotation Sums of Squared Loadings**
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3.331	37.013	37.013	3.331	37.013	37.013	2.615
2	1.809	20.102	57.115	1.809	20.102	57.115	2.487
3	1.383	15.366	72.481	1.383	15.366	72.481	2.390
4	.602	6.688	79.169				
5	.540	6.000	85.169				
6	.462	5.131	90.300				
7	.362	4.018	94.318				
8	.277	3.074	97.392				
9	.235	2.608	100.000				

Extraction Method: Principal Component Analysis.

When deciding on the number of suitable common factors, there is no such thing as one universal true factor solution, but rather a mission to obtain a meaningful number of factors that clarifies the total data in a way that can explain the nature of a concept (Fabrigar & Wegener, 2011). As Fabrigar & Wegener (2011, p. 55) put it; "... it is a decision that is as much theoretical as it is statistical". It should be mentioned, however, that one critique to this pretest, is that even though Fabrigar et al. (1999) explain that a sample of 100 can be adequate under optimal conditions (communalities all higher than .70 with three to five measured variables), that was unfortunately not our case. The sample of 100 was therefore not truly satisfactory per the literature, and should in hindsight have been larger, upwards of 200 participants (Fabrigar & Wegener, 2011;

Mooi et al., 2018). Nevertheless, the three-factor solution proved to fulfill all other aforementioned selection criteria.

Pretest 2: Personality Manipulations

In order to ensure that the manipulations of personalities used in the main study reflected the intended conceptualized personalities of social (social-oriented) and intellectual (task-oriented), we pretested different versions of these two personalities through a series of manipulations of words, pitch, tempo, and rhythm. As previous research has found that most people wish for a female voice in a digital assistant (Shead, 2017; Borau et al., 2021), it was also of interest to determine how a male voice would be rated on social versus intellectual personalities compared to a female voice, as this has not yet been examined (to our knowledge). Before the results are introduced, the theoretical background for the manipulations is presented.

Campbell & Pennebaker (2003) found that the use of words, particularly seemingly meaningless words such as articles and pronouns, reflect the true personality of a person, as these are not as easy to control in free speech compared to controlling which nouns and verbs people use to portray a certain image of themselves. For instance, words such as "except", "but" and "however" show a more cognitive complex personality (intellectual), which in turn appears more trustworthy to the counterparty (Campbell & Pennebaker, 2003). Moreover, having a rich vocabulary and diversity in the use of words is shown to portray intellect (Zhou et al., 2019). On the other side of the scale, saying "I" instead of "we", and speaking directly and impulsively shows an emotional (social) personality (Campbell & Pennebaker, 2003).

Regarding personality types, Chattaraman et al. (2019), Whang & Im (2021), and Rhee & Choi (2019) have previously manipulated social- and task-oriented personalities by altering the formality of language, in which they developed specific sentences that were found to express the two personalities. For instance, a script example of a task-oriented product page is "*I found # product(s). If you like any product, please click on the product to view more details*", whereas the socially-oriented version was "*Great! I found # cool shoes. Hey, if you like any of them, just click on the product image and I*

can show you more details (Chattaraman et al., 2019, p. 321). As depicted through these sentences, a social personality is reflected through informal wording (e.g., using the word "cool"), cheerfulness (e.g., using the word "great"), and fill-words (e.g., using the word "hey" where it does not have an informative function). Pertaining to digital assistants specifically, Chattaraman et al. (2019) and Zhou et al. (2019) found that digital assistants portraying a task-oriented personality should be high on seriousness, intelligence, and entail less chit-chat and thus have a formal conversational style. Furthermore, it must focus on achieving functional aims, not interact beyond simply helping and informing the user, and no social conversation beyond the initial greeting (saying e.g., "hello"). The social personality, however, should be sincere, informal in its wording, and be conversational through small-talk, encouragement, emotional support, and positive expressions (Chattaraman et al., 2019; Bickmore & Cassell, 2005; Yoo et al., 2015).

Beyond the use of words, Schroeder & Epley (2015) and Anikin & Persson (2016) found that pitch variation can also reflect personality. Schroeder & Epley (2015, p. 889) state that "... variance in pitch may reveal the presence of an active or lively mind … enthusiasm, interest and active deliberation, whereas a monotone voice sounds dull and mindless". In other words, pitch variation here concerns vocal glide and changes in intonation. Specifically, Guyer et al. (2019) and Klofstad et al. (2012) showcase evidence that falling intonation and a lowered relative pitch increases perceptions of confidence, competence and trustworthiness. This is conflicting with McAleer & Belin (2019) who found that, specifically for males, a high relative pitch is perceived as more trustworthy than the opposite. In terms of voices with a rising intonation, these voices are deemed more untrustworthy (McAleer & Belin, 2019) but can also serve with the function of being polite, excited or enthusiastic (Warren, 2016). Therefore, our social personality is set to have a higher relative pitch with a rising intonation, whereas our intellectual personality is set to have a lower relative pitch with a falling intonation.

Another aspect of voice reflecting personality is tempo (Dahl, 2010), which is explained as the number of syllables pronounced during a certain time period (Auer et al., 1999). In other words, it pertains to how fast or slow a sentence is said. Auer et al.

(1999) further explain that saliency and relevance (to the context in question) is expressed through a slower tempo, which in this study reflects the intellectual personality and the opposite for the social personality. The last aspect of personality signaled through voice in this study is rhythm (Anikin & Persson, 2016; Auer et al., 1999; Pitt & Samuel 1990; Sapir, 1927). Auer et al. (1999, p. 23) describe rhythm as the "... proportion between the duration (length) of several events (words) or groups of events (phrases) of which the sequence (sentence) is composed". Adding to this, Hildebrand et al. (2020) include the length of speech pauses as an important voice characteristic. Based on the literature of Auer et al. (1999), Pitt & Samuel (1990), and Hildebrand et al. (2020), rhythm is not the same as tempo, as tempo regards the pace of pronunciation of an entire sentence, whereas rhythm regards how the sentence is composed by different sections, words, syllables, and pauses of different stress levels and length. Rhythm is also different from pitch variations, as the latter pertains to how high or low the sound of speech is. Regarding personalities, Sapir (1927) found that formal voices, in our case the intellectual personality, entail less playfulness in variations of rhythm, whereas this is the opposite for voices reflecting a social personality.

In practice, we based our sentences on the research of Chattaraman et al. (2019), Rhee & Choi (2019), Campbell & Pennebaker (2003), and Zhou et al. (2019) for the scripts. As the main study pertained to purchases made through digital assistants, the sentences for the pretest were thus created for a purchase situation. The manipulation of altering pitch variation, tempo, and rhythm was done with one male (Matthew) and one female voice (Joanna), which are two English-US in-built Text-To-Speech (TTS) voices in Amazon Polly (https://aws.amazon.com/polly). We purposely did not use the default voice of Alexa, or other digital assistants, to enable better generalization of the results. The "neural" version of the voices was equipped to enhance quality and lifelikeness, ensuring realistic voices. Based on Speech Synthesis Markup Language (SSML), the prosody was manipulated, where the syntax "conversational" was used for the social personality. When the audio clips were satisfying, pitch variation was slightly tweaked using Audacity, a post-processing tool for sound. In line with Schroeder & Epley (2015) and Bruckert et al. (2010), we increased the variety of tones in the middle and

end of words for the social voice and kept the intellectual voice slightly more monotone and formal. Furthermore, relative pitch was increased for the social personality and decreased for the intellectual personality (Guyer et al., 2019; Klofstad et al., 2012; McAleer & Belin, 2019). Regarding the tempo, we reduced it for the intellectual voice to emphasize salience and relevance (Auer et al., 1999), while increasing the tempo of the social voice for it to be less serious and formal, but rather warmer and more generous (Auer et al., 1999). Lastly, for rhythm, we input more variety in tempo and stress across the different sequences of the social sentences, while reducing variety in these aspects to ensure a less rhythmic voice for the sentences of the intellectual personality (Sapir, 1927; Auer et al., 1999). The audio files can be accessed in Appendix 2. The scripts are presented below:

> *Social:* Hey there, I'm Joanna/Matthew! I am your very own digital assistant. How good to have you here today! How about I show you a product that you have been considering buying? I am very excited, I hope you are too! Don't worry about delivery. Based on your profile, I will have it sent to your closest post office within 3-5 working days, completely free of charge! What do you say? Should I put the product in your shopping cart?
>
> *Intellectual:* Hi, my name is Joanna/Matthew, your digital assistant. You will momentarily be introduced to a product that you have been considering buying. Based on the information that is stored in your profile, this product can be sent to your closest post office within approximately 3-5 working days. There will be no shipping fee. Please, take your time and consider if you want me to add the product to your shopping cart.

Results Pretest 2: Personality Manipulations

The sample consisted of 98 participants (gender: male = 46, female = 51, other = 2, N/A = 1; age: 18-29 = 41, 30-39 = 34, 40-49 = 11, 50-59 = 9, 60+ = 3, N/A = 1), where the response time of two participants from the original sample of 100 was too short for having devoted sufficient attention and answered thoughtfully. These respondents were thus excluded from the data set. The average completion time was four minutes and two seconds, and each participant was paid £0.51. Participants were exposed to four

voices through audio files: 1) female social, 2) female intellectual, 3) male social, and 4) male intellectual. As previously mentioned, the female voice was Joanna, and the male voice was Matthew. The order of exposure was randomized in Qualtrics to eliminate potential order effects. For each audio file, participants were asked to evaluate 13 personality traits in relation to the voices, on a 7-point Likert scale ranging from very low (1) to very high (7). The dimensions were those of Rosenberg et al. (1968), relating to good-social and good-intellectual which reflect social- and task-orientation respectively, as these coincide with the warmth-competence framework of Fiske et al. (2007) and previous measurements of Bickmore & Cassell (2005). In total, we extracted the following personality traits to measure the perceptions of our two personalities: Skillful, serious, intelligent, reserved, cautious, determined, and scientific for the intellectual personality, and good-natured, happy, sociable, warm, helpful, and sincere for the social personality.

The social personality showed a Cronbach's α of .926, whereas the intellectual personality gave a Cronbach's α of .781, which indicates satisfactory results (Malhotra et al., 2010). Table 4 displays that, when comparing the mean scores of the dimensions through a one-way ANOVA, both personalities were perceived as intended for Joanna and Matthew. The highest evaluation for both the social (M = 5.40, SD = 1.089) and intellectual (M = 4.69, SD = .99) personalities were given to Matthew compared to those of Joanna (M_{social}= 4.87, SD = 1.24, $M_{intellectual}$= 4.61, SD = .89). The two personalities of Matthew were significant and more different ($p < .001$) than the two personalities portrayed by Joanna ($p = .674$). Hence, Matthew was able to reflect the social and intellectual personalities more distinctly than what Joanna was able to. This is also the case for differences of social and intellectual traits within each respective personality type. For Matthew, the intellectual personality showed a difference of 1.147 between the intellectual and social items ($p < .001$), whereas the difference was slightly larger for the two sets of traits within the social personality (difference in means = 1.215, $p < .001$). The differences proved to be greater than for Joanna, who had a mean difference of 1.085 for the dimensions of her intellectual personality ($p < .001$) and 0.637 ($p < .001$) for her social personality. As both personalities of both Matthew and

Joanna were different, showing that they were rated as we intended, Matthew stood with the greatest difference, thus more appropriate for being used in the main study.

Table 4: Rating of Personality Traits.

	Matthew				**Joanna**			
Personality	Social		Intellectual		Social		Intellectual	
Rating of personality traits	Social	Intellectual	Social	Intellectual	Social	Intellectual	Social	Intellectual
	5.40	4.18	3.54	4.69	4.87	4.24	3.53	4.61

The social personalities were also deemed relatively high on the intellectual dimensions, showing that these two are not mutually exclusive. This highlights the evolution of such personalities as shown by Chattaraman et al. (2019), Zhou et al. (2019), and Rhee & Choi (2020) where dimensions of intellect shine strongly through a positive and sociable personality while social elements are less present, but not omitted, for an intellectual personality. This is aligned with the aforementioned counselor or secretary personalities (Zhou et al., 2019; Rhee & Choi, 2020), and supports the findings of Lopatovska (2020), who found that digital assistants are perceived as always competent regardless of sociability. In specific, the social personality of Matthew was considered less task-oriented than Joanna's social personality ($M_{Matthew}$= 4.18, SD = .83, M_{Joanna}= 4.24, SD = .96), although both were rated relatively high. The intellectual personalities had social elements which were rated almost equal for both voices ($M_{Matthew}$ = 3.54, M_{Joanna}= 3.53). The difference between the intellectual and social dimensions within all personalities was significant ($p < .001$).

The findings of the pretest are interesting by themselves, as other researchers have postulated that most people prefer a female voice compared to a male (Shead, 2017; Borau et al., 2021). It is also interesting to see how Matthew, providing a male voice, proves to manifest two different personalities that are perceived clearer by the audience relative to the female voice (Joanna). There were no significant differences between age groups and their evaluation of the personalities, but we did find some gender-related differences. There was a main effect of gender ($p < .001$), where males found

the social dimensions of the intellectual personalities of both Joanna ($p < .001$) and Matthew ($p = .010$) to be, on average, higher than females ($M_{Joanna_{Males}} = 3.97$, $M_{Joanna_{Females}} = 3.12$, and $M_{Matthew_{Males}} = 3.92$, $M_{Matthew_{Females}} = 3.21$). This indicates that males might be inclined to find voices as more social, regardless of the gender of the voice they hear. This could be aligned with the fact that men tend to be less sensitive to detect emotional cues received from voice than females (Schirmer et al., 2005).

Main Study

Participants

As with the pretests, participants were recruited through Prolific Academic, resulting in a sample size of N = 641, where each participant received a payment of £1.00. To ensure unbiased responses, previous participants who had completed our pretests were removed from the pool in Prolific. The sample was based on calculations in G*Power, with a partial $\eta^2$of .03, error probability of .03%, and statistical power of .95. The error probability and partial η^2 were based on suggestions from Cohen, where .0099 represents a small effect, a medium effect is .0588 (Richardson, 2011) and .04 is often a good estimate (Brysbaert, 2019). As these calculations proved a sample of 619, we rounded it up to 640 to ensure approximately 80 participants per group. We based our recruitment on the criteria of age, as Auxier (2019) found that those who are 18-29 years of age have adopted most digital assistants, where the age group of 30-49 were second on that list. Furthermore, Kinsella & Mutchler (2018) explain that those who are below 45 are approximately 8.5 times (1,32% versus 0.16% for age groups > 44) more likely to prefer voice shopping than others. Hence, our participants were recruited between 18-44 years of age, as they are the most engaged with digital assistants and are more enthusiastic about voice shopping than those who are older. These criteria were used to ensure that the results were not influenced by a predisposition to having a negative outlook, either in terms of risk or by having technological disadvantage, on digital assistants as a medium for purchases. Finally, the participants were recruited from the US with English as their native language. This is similar to the pretests to keep

consistency, but also due to the information used by Auxier (2019) and Kinsella & Mutchler (2018) stemming from the US market.

Independent Variables

As we conducted a 2 (modality: voice vs. voice and visual) x 2 (personality: social vs. intellectual) x 2 (involvement level: low vs high) between-participants experiment, we needed to manipulate the different treatment conditions for each of the independent variables. This will be discussed further in the following.

Involvement Level

The first manipulation of our study was product involvement (low vs. high). Even though risk factors related to the product and involvement levels are highly causal, these were not to be determined here, as they have been, both specifically and generally, vastly researched. Thus, the chosen products for the main study were derived from the literature. Different authors list for instance beverages, groceries, chocolate, scented candles, light bulbs, candy, magazines, books, coffee, sunscreen, slippers, detergents, toothpaste, deck of playing cards, video-on-demand, vitamins, aspirin, and toilet tissue as low involvement products (Hong, 2015; Radder & Huang, 2008; Akbari, 2015; Stewart et al., 2019; Liu et al., 2020; Drossos et al., 2014; Holmes & Crocker, 1987; Rhee & Choi, 2020; Kaplan et al., 1974). High involvement products are exemplified through clothing, skin lotion, cars, TV's, laptops, smartphones, refrigerators, cameras, cruises, musical festivals, and airline tickets (Hong, 2015; Stewart et al., 2019; Mittal, 1989; Akbari, 2015; Liu et al., 2020; Solomon, 1986; Holmes & Crocker, 1987; Rhee & Choi, 2020; Richins & Bloch, 1986).

To ensure that the chosen products do not have salient characteristics which can cause intricacies such as confounds (e.g., a favorite taste or smell) and that the products are not within the reordering category, the low involvement product for this main study was a deck of playing cards. This is also the product that is rated the lowest on overall perceived risk by Kaplan et al. (1974). The high involvement product we chose was a laptop, which is mentioned as high involvement by many authors (Hong, 2015; Akbari, 2015; Stewart et al., 2019; Liu et al., 2020; Venkatraman, 1989). Furthermore, neither

a deck of cards nor a laptop deviates substantially within their product category in terms of design or other attributes, attenuating the possibility that one participant could favor the product in question over another. On that note, a MacBook was not chosen, as it could deviate too much from the standard laptop due to its design and perhaps cause confounds, such as brand attachment due to recognition of the MacBook.

A final remark relates to the division of hedonic and utilitarian products. Much literature has evaluated hedonic products, products providing experiential consumption entailing fun, pleasure, and excitement (Dhar & Wertenbroch, 2000), to generate a stronger emotional response (Pham, 1998; Malhotra, 2005) than utilitarian products, which entail instrumental and functional goals (Dhar & Wertenbroch, 2000). Bettiga et al. (2020), however, found that equal emotional responses relate to both product types, showcasing that the former might not be the case. Even though that might be true, with examples of hedonic products to be wine, flowers, perfumes, and luxury watches (Venkatraman, 1989; Bettiga et al., 2020), it seems that these come with strong subjective preferences, which could have caused confounds for the main study. Thus, utilitarian products, such as a laptop (Hirschmann & Holbrook, 1982) and a deck of cards are likely to be more neutrally perceived.

Personality

The second manipulation of the experiment was the personalities and interaction styles of the digital assistant. Based on the pretest, Matthew, the male voice, was evaluated most favorably on both personalities (social and intellectual) and was therefore chosen for the main study. As previously, Amazon Polly was utilized to create the voices. The voices were set to “neural”, where the syntax of “conversational” was added for the social personality. The scripts for the main study were longer than in the pretest with the purpose of informing participants about the products of exposure, but the nature of the wording in the scripts was similar as previously. Pitch variation, tempo, and rhythm were manipulated in Amazon Polly through SSML and went through slight post-processing in Audacity, as with the pretest. Hence, we followed the same procedure of including increased relative pitch, more pitch variation, higher tempo, and more variations in rhythm for the social voice, and decreased relative pitch, less pitch

variation, lower tempo and less rhythm for the intellectual voice. Both voices provided equal information, but simply angled differently, in accordance with the pretest and previous research. In total, four different audio files were created to reflect social and task-oriented personalities with high and low involvement products. See Appendix 3 for audio files and scripts.

Modality

The third and last variable of manipulation was modality, that is exposure to the voice message of a digital assistant, or both the voice message and visual information (picture and text) showcasing the product in question. The latter was to represent the existence of a screen attached to the assistant, providing additional information through an added sensory stimuli. The picture was complemented by text, specifically information about product features through listwise bullet points, simply for the purpose of making the scenario realistic and similar to ordinary online shopping. We are aware that combining picture and text disables the ability to examine which of the two elements influence purchase intention and to what extent, at least with the present 2 x 2 x 2 design. However, it is, as mentioned, purposely done to ensure a realistic purchase scenario in the case of digital assistants.

Pictures of the laptop and the deck of cards (see Appendix 4 for pictures of the products) were carefully selected to mitigate potential visual noise. These were also chosen to ensure that potential effects on purchase intention due to visual information were caused by visualization of the products and the text related to the products, not other aspects of the pictures that could have caused confounds. Both pictures showed neutral, unbranded products that fit the description of product features communicated by either voice or voice and text combined and nothing more beyond that. The logo and “made in Taiwan”-stamp on the playing cards were photoshopped out of the picture to remove any chance of branding- or country-of-origin effects. The laptop was ordinary, colored black, slim, and was equipped with modern hardware, which was confirmed through multiple searches on electronic stores online. We chose to explicitly communicate the price of the products to ensure control over this aspect. The prices were based on real average US prices, leaving the average price of the laptop at $700

(Smith, 2020b) and the deck of playing cards at $5, according to various online searches. We could have chosen not to price the products and allow participants to intuitively consider their own internal reference prices, however, that could have resulted in an ambiguous understanding of the product and hence affected purchase intention.

In order to expose participants to the stimuli, we created videos containing a picture of a digital assistant, as users in real-life situations would naturally see the assistant when physically interacting with it. The voice files were then added on top of the video. For the modality of voice and visual, we included the pictures of the laptop and the deck of cards on top of the voice file and the picture of the digital assistant. These pictures appeared in the video at the time they were first mentioned by the voice, accompanied by textual information on product features in listwise bullet-points, chronologically following the order of information from the script voiced by the digital assistant. Appendix 5 provides the link to all eight videos.

Dependent Variable

As the dependent variable of the study was purchase intention, participants were to report their individual purchase intentions upon exposure to the treatment conditions. Purchase intention is synonymous with purchase probability (Armstrong et al., 2000) and thus, purchase intention was measured through a purchase probability scale as it is found to provide substantially better prediction of purchase behavior (Day et al., 1991; Wright & Macrae, 2007). Therefore, the purchase probability question and scale from Juster (1966) was utilized. The original question by Juster (1966, p. 670) is: *"During the next (6, 12, 24) months, that is, between now and next_, what do you think the chances are that you or someone in the household will buy a_?"*. Table 5 shows the descriptive 11-point item scale ranging from 0-10, used by Juster (1966). Each descriptive item represents an additional 10% probability of purchasing. Thus, answering for instance item 9: *"almost certain to buy"*, is translated into a 90% purchase probability. We used the same measurement scale, due to this being generic and not dependent on product nor sales channel, while the question itself was modified to this particular experiment. In detail, the time period and whether others in the

household would purchase the product was omitted, leaving us with the following question: "*What do you think are the chances that you will purchase this product through the digital assistant?"*

Table 5: 11-item purchase probability scale (Juster, 1966, p. 670).

10	Absolutely certain to buy	10
9	Almost certain to buy	9
8	Much better than even chance	8
7	Somewhat better than even chance	7
6	Slightly better than even chance	6
5	About even chance (50-50)	5
4	Slightly less than even chance	4
3	Somewhat less than even chance	3
2	Much less than even chance	2
1	Almost no chance	1
0	Absolutely no chance	0

Control Variables

As previous experience, habit and familiarity affect the use and adoption of technology (Venkatesh, 2012) by increasing trust in the medium (Lee & Turban, 2001), decreasing perceived risk (Miyazaki & Fernandez, 2001), and increasing purchase intention (Lindh et al., 2020), these were necessary to control for in the main study. Other endogenous variables that were controlled for were age and gender. To evaluate interactions based on our framework, NFV, perceived risk, and trust were controlled for as conditional values.

Need for Voice

NFV contained the nine items from the EFA based on the data from the pretest (Appendix 1.2). Evidence has found individual differences in terms of preference for sensory modality-specific information (Peck & Childers, 2003), and thus we may capture these differences by evaluating participants' individual NFV. We purposely chose to conduct this at the beginning of the online experiment prior to exposing participants to treatment conditions, as we sought to avoid bias of participants being influenced by the manipulations of the study. For instance, if measuring NFV post-treatment, we could have experienced the response of participants to be biased by being exposed to treatment conditions that they liked or disliked just moments before.

Perceived Risk

As opposed to purchasing goods in physical stores, online shopping represents an added risk of using the medium in question (Bianchi & Andrews, 2012). This relates to the security and reliability of the transaction (Biswas & Biswas, 2004) and is the kind of risk that was measured in this study. The most researched sub-dimensions of risk within the context of online purchases and the medium are product performance risk, product risk, financial risk, delivery risk, and transactional risk, where measurement items are suggested by Hong (2015), Biswas & Biswas (2004), Foresythe et al. (2006), Kim et al. (2008) and Bhatnagar & Ghose (2004) among others. Instead of measuring each of these risks separately, as perceived risk is not the main focus of this study, Bianchi & Andrews' (2012) scale of perceived risk in the context of online shopping was used. The foundation for the scale was originally developed by Jarvenpaa et al. (1999), then modified and used by Andrews et al. (2007) and also applied in Bianchi & Andrews (2012). The four items were all measured on a 1-7 Likert scale ranging from strongly disagree (1) to strongly agree (7), as used by Bianchi & Andrews (2012). As regular online shopping is different from voice shopping, Table 6 shows that the scale was modified to fit the specific context of digital assistants. Finally, a benefit of using this scale is that it captures, to some extent, all the sub-dimensions mentioned previously but where we allow for ambiguity in which sub-dimensions relate to the two last items.

Table 6: Modified Scale for Perceived Risk.

Items	Authors
I feel safe making purchases through a digital assistant using my credit card *	Bianchi & Andrews (2012); Andrews (2007); Jarvenpaa et al. (2000)
I feel safe allowing my personal information to be kept by the digital assistant purchases *	
Compared with other ways of making purchases, I think using the digital assistant is more risky	
There is too much uncertainty associated with using a digital assistant to make purchases	

* Reversed items

Trust

As explained in the literature review, trust is an important antecedent to accept risk and consequently form a purchase intention (Martin et al. 2015, Delgado-Ballester & Munuera-Aleman, 2001; Pitardi & Marriott, 2021; Ha et al., 2014; Yoon, 2002; Kim

et al., 2008). Even though purchase intention is the dependent variable of the study, trust is a necessary means in understanding the mechanisms in the influence of personality type, modality, and involvement level on purchase intention. Thus, in order to examine participants' level of trust in the digital assistant after exposure to treatment conditions, we utilized the scale of Pitardi & Marriott (2021), displayed in Table 7. All items were measured on a 1-7 Likert scale ranging from strongly disagree to strongly agree, as originally used by Pitardi & Marriott (2021).

Table 7: Modified Scale for Trust.

Items	Authors
I feel that the digital assistant made truthful claims I feel that the digital assistant was trustworthy I believe what the digital assistant told me	Pitardi & Marriott (2021); Chattaraman et al. (2019); Ye et al. (2019); Hassanein & Head (2007)

Pitardi & Marriott (2021) collected these items from Chattaraman et al. (2019), Ye et al. (2019), and Hassanein & Head (2007), in which their analysis found the items to have a Cronbach's α of 0,838, indicating a high item reliability (Malhotra et al., 2010). There have previously been developed multiple scales measuring the construct of trust in commercial settings, such as Purwanto et al. (2020) for in-phone assistants (e.g., Apple's Siri and Samsung's Bixby), Yagoda & Gillan (2012) for less advanced robots, and Martin et al. (2015) for online shopping. As opposed to these scales, Pitardi & Marriott's (2021) scale is in fact used to measure trust in relation to subsequent consumer behavior for digital assistants in specific, thus making it highly suitable for this study.

Online Experiment Procedure

The online experiment, accompanied by a questionnaire, followed a randomized between-participants design to remove the potential risks of fatigue, order, and learning effects, which would most likely have occurred with our research in a within-subject design, where those risks stand as prevalent (Gravetter & Forzano, 2016). Participants were exposed to only one type of personality (social vs. intellectual), modality (voice vs. voice and visual), and involvement level (high vs. low) through applying the

counterbalancing feature in Qualtrics, which ensured that each group had an approximately equal number of participants. Furthermore, we added quotas (each group = 80) in Qualtrics, to strengthen this effect. In total, the sample was divided into eight different groups, where Table 8 displays the treatment conditions for all groups.

Table 8: Between-Participants Treatment Conditions.

	Personality	**Modality**	**Involvement**
Group 1	Social	Voice	Low
Group 2	Social	Voice	High
Group 3	Social	Voice + visual	Low
Group 4	Social	Voice + visual	High
Group 5	Intellectual	Voice	Low
Group 6	Intellectual	Voice	High
Group 7	Intellectual	Voice + visual	Low
Group 8	Intellectual	Voice + visual	High

As with all types of data collection, experiments also represent certain disadvantages such as lack of realism and motivational differences (Vercruyssen & Hendrick, 2011), whereas online surveys have the weaknesses of lack of monitoring and respondent control (Burns et al., 2016). For the former, we tried to manipulate both aspects, where an imagined scenario was presented pre-treatment as part of a story to set the stage (Webster & Sell, 2014). As imagination is vivid, it can increase the degree of realism as it is proved by Klatzky (2010) and Krishna et al. (2016), among others, that individuals can picture themselves in the set scenario. On one hand, Webster & Sell (2014) explain that the essence of this story should be mentioned three times to ensure that participants capture the message. On the other hand, we found a third addition to be cumbersome as the story was relatively short, and it was therefore mentioned twice. For the latter, we added an attention check amid the questions related to NFV, and a soundcheck before the treatment to ensure that the participants in fact had functioning sound, which was required to hear the voice manipulation. Finally, the online

experiment was kept short ($M_{response\ \ time}$= 5 min 18s), as this helps to relieve fatigue effects (Webster & Sell, 2014).

The online experiment follows the logic of Burns & Bush (2008) pertaining to questionnaires. In specific, Burns & Bush (2008) refer to two aspects: the introduction and the actual flow of questions. The introduction should contain five functions which are to identify the surveyor, indicate the purpose of the survey, explain how respondents were selected, request for participation, and assess whether an individual fits to be a respondent in the particular study (Burns & Bush, 2008). The second and last functions were omitted as pre-qualifications were made in Prolific Academic. What Burns & Bush (2008) describe as the flow of questions is regarding the sequence of questions. In sum, the sequence of questions should follow this order: warm-up questions (generates interest), transition (notifies that the questions will change), complicated and difficult-to-answer questions, and lastly classification and demographic questions (Burns & Bush, 2008). Specifically, the questionnaire began with questions pertaining to familiarity, previous use, and potential ownership, where responding yes to these opened more in-depth questions regarding the frequency of use and potential purchases completed. This was simply to control for participants' previous experience, familiarity, and habits with digital assistants. The next part of the survey contained the nine NFV-items, in which the aforementioned attention check was added amongst the questions. Following this, participants were presented with the imagined scenario to create the context before randomly being exposed to one of the eight treatment conditions. In the set scenario, the participants were told that they were interested in purchasing a certain product, and that their personal and credit card information were stored in the digital assistant. These aspects were included to ensure that participants could imagine wanting the product of exposure, and that there was no ambiguity concerning payment nor delivery. Upon exposure, the next aspect was the question of purchase intention, prior to the two measurement scales of trust and perceived risk, before closing off the survey with questions covering participant's demographics. All the questions were grouped by subjects, not leading, and we attempted to make them as unambiguous as possible, following the guidelines of Kelley

et al. (2003). Furthermore, to reduce order-effects, questions regarding NFV, perceived risk and trust were randomized. See Appendix 6 for the full main study questionnaire.

Results & Analysis

Upon conducting the online experiment, the data was analyzed through multiple statistical procedures using IBM SPSS Statistics 27. First, the construct of NFV was analyzed to confirm the factors from the pretest and examine the findings related to the construct itself. Second, our sample characteristics are presented together with descriptives, frequencies, and mean ratings of the dependent variable and relevant covariates. Third, the results from the main model are presented for hypotheses and research question testing. Fourth and last, findings beyond the hypotheses are presented through the full model in an exploratory analysis before moving on to the general discussion.

Need For Voice

To confirm how our three-factor solution performed in the main study with a larger N, confirmatory factor analysis was conducted. Table 9 shows that factor one was similar as in the EFA, whereas item four loaded on factor number three and item seven loaded on two factors. The explained variance was reduced to 67.06 from 72.48, but KMO was markedly increased to .832 from .731. The construct revealed a satisfactory Cronbach's α of .822. However, as two items get even poorer communalities and as they do not load as hoped, further work will be needed to further study these items and other items in the NFV scale. With that being said, we did encounter the fascinating finding that NFV in fact appears to be a phenomenon. It turned out that the mean was in general very high (M = 5.407, SD = .67), where surprisingly only one individual was rated below 3 on the scale (see Figure 2 for a histogram depicting NFV-scores). According to our analysis, the median was close to the average (Median = 5.44) showcasing that most participants were relatively high on NFV.

Table 9: Structure Matrix CFA.

Items	Factors		
	1	2	3
When I listen to a pleasant voice, it elicits positive emotions	.820		
Listening to a pleasant voice is delightful	.711		
I experience a sense of satisfaction when I hear certain voices	.641		
I tend to be able to assess personal characteristics of the speaker when listening to that person's voice			.537
I can infer the emotions that a person is expressing through their voice		.674	
I tend to be able to infer how people feel based on their voice		.912	
When I listen to a pleasant voice, it increases my trust in the person who speaks		.557	.504
I tend to desire that a person I talk to has a voice that is congruent with their identity			.733
My attention towards a person increases when their voice appears congruent with their personality			.773

Extraction Method: Maximum Likelihood.

Rotation Method: Promax with Kaiser Normalization.

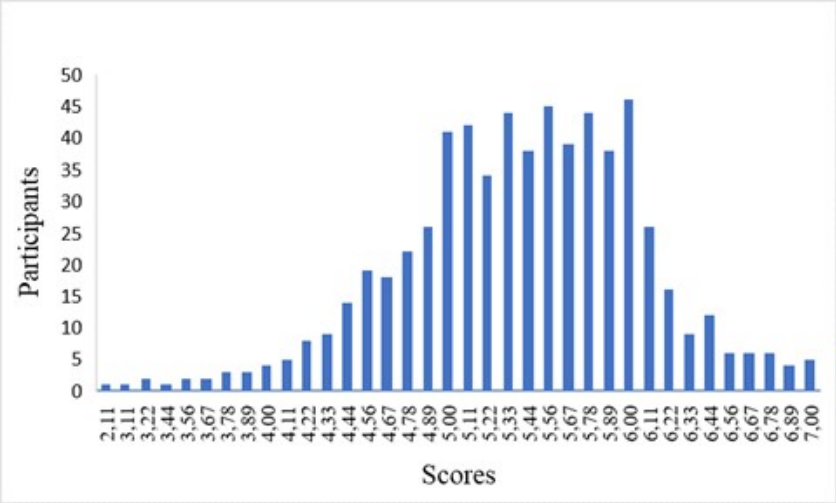


Figure 2: Histogram of NFV Scores.

Sample Characteristics and Descriptive Statistics

The sample of N = 641 was fairly evenly distributed among the eight different treatment conditions, as well as gender and age groups. This resulted in the approximate average age of 27.08, thus reflecting a rather young consumer with high technological familiarity. An initial ANOVA showed that there were no significant differences between groups of neither age nor gender. Table 10 displays the sample, as distributed into the eight treatment conditions with the respective gender, age, purchase intention, trust, and perceived risk per group.

The mean purchase intention across all groups was rather low (M = 4.47, SD = 2,84) on the 11-item scale. Pertaining to treatment conditions, the greatest purchase intention was observed for participants in Group 5 "voice, intellectual personality, low involvement" (M = 5.26), and Group 7, "voice and visual, intellectual personality, low involvement" (M = 5.19). The lowest purchase intention was for Group 6, "voice, intellectual personality and high involvement" (M = 3.84), and Group 8, "voice and visual, intellectual personality, high involvement" (M = 3.86), thus signaling a negative influence of high involvement products on purchase intention. Further, a post-hoc test revealed that there were significant differences between Group 5 and Group 6 (p = .031), and Group 5 and Group 8 (p = .037). For the treatment conditions entailing the social personality, the mean purchase intentions were quite similar for both low and high involvement products. These treatment conditions showed no significant differences among each group nor across the four other groups.

Examining trust ratings, the mean trust was consistent across all treatment conditions and in total quite high (M = 5.27, SD = 1.15) on the 7-item scale. The lowest trust was found for participants of Group 1, "voice, social personality, low involvement" (M = 4.88, SD = 1.36), whereas the highest trust was found for Group 8, "voice and visual, intellectual personality, high involvement" (M = 5.58, SD = 1.09). In general, there was reported slightly greater trust for all groups of the intellectual personality compared to the social personality. Group 1 and Group 8 were also significantly different from each other in trust ratings (p = .004). Looking at perceived risk, the ratings were quite

similar across all groups with no significant differences, where the mean risk (M = 4.42, SD = 1.37) was in the medium range of the 7-item scale.

The introductory questions of the study (Appendix 6) disclosed that 74% of participants own a digital assistant, with 30-35-year-olds being the group with the highest degree of ownership. In total, 91% (n = 583) of the sample reported having previous experience in interacting with digital assistants, thus assuring that the vast majority was already familiar with assistants and did not need to be educated on the nature of the devices. Among this group, the mean frequency of interaction was once a week (frequency = 5.02 of 8, where 8 = several times a day), with the largest group being participants who use the assistant several times a day (24%, n = 140 of participants with previous interaction experience). Despite high ownership and previous interaction, only 12.64% (n = 80) of the participants reported that they have previously purchased through a digital assistant. This number was similar for all age groups (approximately 16-20% for each group), except for the age group 18-24, where only 5.1% had previously purchased. This low level of familiarity with purchases through digital assistants might be an influential factor in the aforementioned low purchase intention among participants. Among those of previous purchase experience through digital assistants, the mean purchase frequency was between once a month to once every other week (M = 3.56 of 8, where 8 = several times a day). 26.3% (n = 21) of these participants used the assistant to make purchases once a week to several times a week, thus indicating that only a small portion of the sample can be categorized as frequent buyers.

Table 10: Between-Participants Frequencies and Descriptives.

Treatment conditions	N	Gender**			Age***	Purchase intention	Trust	Risk
		Male	Female	Non-binary				
Group 1: *Voice, Social, Low involvement*	82	42	38	1	M = 27.60	M = 4.35 (SD = 3.09)	M = 4.88 (SD = 1.36)	M = 4.61 (SD = 1.40)
Group 2: *Voice, Social, High involvement*	81	43	36	2	M = 26.88	M = 4.38 (SD = 2.83)	M = 5.21 (SD = 1.12)	M = 4.30 (SD = 1.40)
Group 3: *Voice + Visual, Social, Low involvement*	81	40	40	1	M = 27.28	M = 4.43 (SD =2.86)	M = 5.04 (SD = 1.27)	M = 4.33 (SD = 1.37)
Group 4: *Voice + Visual, Social, High involvement*	79	48	30	0	M = 27.08	M = 4.47 (SD = 2.81)	M = 5.25 (SD = 1.18)	M = 4.55 (SD = 1.48)
Group 5: *Voice, Intellectual, Low involvement* *	81	43	36	3	M = 26.44	M = 5.26 (SD = 2.79)	M = 5.42 (SD = 0.96)	M = 4.25 (SD = 1.31)
Group 6: *Voice, Intellectual, High involvement* *	80	42	34	4	M = 27.16	M = 3.84 (SD = 2.59)	M = 5.38 (SD = 1.15)	M = 4.52 (SD = 1.36)
Group 7: *Voice + Visual, Intellectual, Low involvement*	78	44	32	1	M = 27.04	M = 5.19 (SD = 2.64)	M = 5.42 (SD = 0.92)	M = 4.27 (SD = 1.20)
Group 8: *Voice + Visual, Intellectual, High involvement* *	79	46	32	1	M = 27.12	M = 3.86 (SD = 2.87)	M = 5.58 (SD = 1.09)	M = 4.53 (SD = 1.42)
Total:	641	348	278	13	M = 27.08	M = 4.47 (SD = 2.84)	M = 5.27 (SD = 1.15)	M = 4.42 (SD = 1.37)

* Significant at $\alpha = .05$

** N_Gender N/A = 3

*** N_Age N/A = 1

Effects of Independent Variables on Purchase Intention, Perceived Risk, and Trust

The analyses were based on four-way ANOVAs. The manipulations included involvement, personality, modality, and NFV, while the main dependent variable was purchase intention. Trust and perceived risk were also tested to evaluate our theoretical framework. The construct of NFV was not a part of our hypotheses covering interactions, but as it is highly explorative, it was interesting to see how it could interact with other independent variables. Furthermore, NFV was split into two groups based on the median, where all participants below 5.44 (n = 339) were categorized as "low NFV" and all participants above 5.44 (n = 302) were categorized as "high NFV". Table 11 displays the results and effect sizes, while Appendix 7 contains all interaction plots. Furthermore, Figure 3 visualizes the effect of each independent variable on our main dependent variable, purchase intention, while the effects on trust and perceived risk can be found in Appendix 8.

Table 11: Four-way ANOVAs on Purchase Intention, Trust and Perceived Risk.

Main effects and Interactions	Dependent variable	Sig.	F	Partial η^2
Involvement	Purchase intention	.002*	9.655	.015
Personality	Purchase intention	.571	.322	.001
Modality	Purchase intention	.975	.001	.000
NFV	Purchase intention	.001*	10.332	.016
Personality * Involvement	Purchase intention	.001*	10.814	.017
Personality * Modality	Purchase intention	.943	.005	.000

Personality * NFV	Purchase intention	.024*	5.116	.008
Involvement * Modality	Purchase intention	.786	.074	.000
Involvement * NFV	Purchase intention	.309	1.036	.002
Modality * NFV	Purchase intention	.412	.674	.001
Personality * Involvement * Modality	Purchase intention	.961	.002	.000
Personality * Involvement * NFV	Purchase intention	.489	.478	.001
Personality * Modality * NFV	Purchase intention	.467	.530	.001
Involvement * Modality * NFV	Purchase intention	.646	.211	.000
Personality * Involvement * Modality * NFV	Purchase intention	.170	1.889	.003
Involvement	Trust	.089	2.902	.005
Personality	Trust	< .001*	16.175	.025
Modality	Trust	.275	1.193	.002
NFV	Trust	< .001*	12.042	.019
Personality * Involvement	Trust	.184	1.769	.003
Personality * Modality	Trust	.733	.117	.000
Personality * NFV	Trust	.038*	4.343	.007
Involvement * Modality	Trust	.854	.034	.000
Involvement * NFV	Trust	.640	.220	.000
Modality * NFV	Trust	.209	1.582	.003

Personality * Involvement * Modality	Trust	.487	.483	.001
Personality * Involvement * NFV	Trust	.668	.185	.000
Personality * Modality * NFV	Trust	.025*	5.014	.008
Involvement * Modality * NFV	Trust	.766	.089	.000
Personality * Involvement * Modality * NFV	Trust	.839	.041	.000
Involvement	Perceived Risk	.258	1.280	.002
Personality	Perceived Risk	.504	.447	.001
Modality	Perceived Risk	.956	.003	.000
NFV	Perceived Risk	.601	.273	.000
Personality * Involvement	Perceived Risk	.124	2.378	.004
Personality * Modality	Perceived Risk	.928	.008	.000
Personality * NFV	Perceived Risk	.950	.004	.000
Involvement * Modality	Perceived Risk	.176	1.832	.003
Involvement * NFV	Perceived Risk	.096	2.787	.004
Modality * NFV	Perceived Risk	.814	.055	.000
Personality * Involvement * Modality	Perceived Risk	.308	1.402	.002
Personality * Involvement * NFV	Perceived Risk	.739	.112	.000
Personality * Modality * NFV	Perceived Risk	.223	1.486	.002
Involvement * Modality * NFV	Perceived Risk	.017*	5.757	.009
Personality * Involvement * Modality * NFV	Perceived Risk	.554	.350	.001

* Significant at $\alpha = .05$

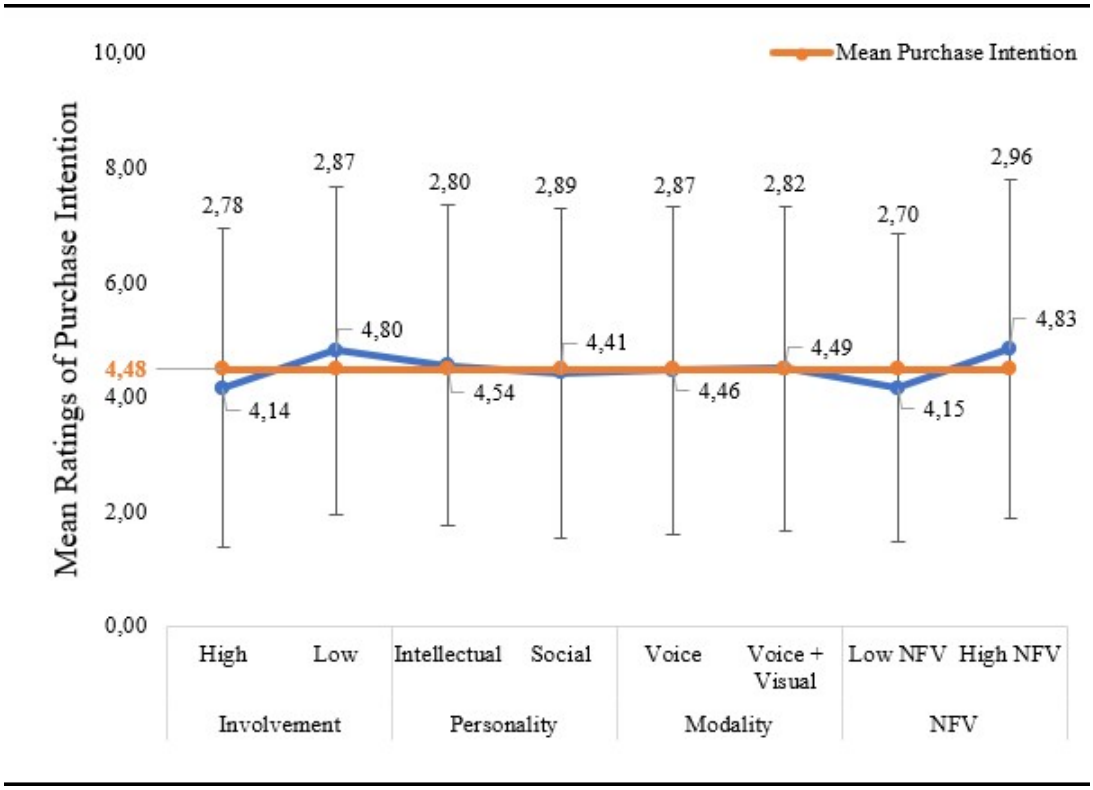


Figure 3: Summary Figure of Independent Variables on Purchase Intention.

When testing purchase intention as the dependent variable (model partial η^2 = .061), there were significant main effects for involvement ($M_{high\ involvement}$= 4.14, $M_{low\ involvement}$= 4.80) and NFV ($M_{low\ NFV}$= 4.15, $M_{high\ NFV}$= 4.83). Moreover, a significant interaction between personality and involvement ($M_{social_{high\ involvement}}$ = 4.43, $M_{social_{low\ involvement}}$ = 4.39, $M_{intellectual_{high\ involvement}}$ = 3.85, $M_{intellectual_{low\ involvement}}$ = 5.23), and NFV and personality was found ($M_{social_{high\ NFV}}$ = 5.03, $M_{social_{low\ NFV}}$ = 3.81, $M_{intellectual_{high\ NFV}}$ = 4.47, $M_{intellectual_{low\ NFV}}$ = 4.62). Testing trust (model partial η^2 = .069) showcased significant main effects for personality (M_{social}= 5.09, $M_{intellectual}$ = 5.45) and for NFV ($M_{low\ NFV}$= 5.13, $M_{high\ NFV}$ = 5.42), and a significant interaction between personality and NFV ($M_{social_{high\ NFV}}$ = 5.35, $M_{social_{low\ NFV}}$ = 4.85, $M_{intellectual_{high\ NFV}}$ = 5.52, $M_{intellectual_{low\ NFV}}$ = 5.39). Furthermore, a three-way interaction between NFV,

modality and personality on trust was found, as well as a three-way interaction between NFV, involvement and modality on perceived risk.

The results supported only one hypothesis related to the main effects, H3, where there was greater purchase intention for low involvement products compared to high involvement products in the context of digital assistants. In terms of interactions, there was no support for any of the stated hypotheses. We did, however, find positive results for the exploratory construct of NFV. Based on our findings, purchase intention and trust increased for those who were high on NFV compared to those who were categorized as low. Hence, RQa was supported and RQb was partially supported. It is notable to witness the many interesting interactions between NFV and other independent variables, showcasing practical implications and confirming the merit of the construct.

Including the Effect of Covariates on Purchase Intention

In order to determine how the covariates affect the model, an ANCOVA was conducted. The dependent variable was purchase intention, factors included our three manipulations, and the covariates were namely perceived risk, trust, NFV, and previous purchase experience (dichotomous, 1 = purchased previously, 0 = has not purchased previously) through digital assistants. Following the literature, risk, trust, familiarity, and previous experience with assistants are known predictors of purchase intention. Additionally, based on the effects of NFT on proclivity to purchase (Peck & Childers, 2003), NFV is also considered to be an important antecedent of purchase intention in the case of digital assistants, and was therefore controlled for. Running the full-model ANCOVA showed that the effect size of the model increased markedly (partial η^2 = .278, $p < .001$), confirming that the covariates are in fact central constructs for explaining purchase intention across all treatment conditions. Table 12 shows that all covariates, except NFV ($p = .062$), are significant predictors of purchase intention, with rather low effect sizes. On that note, the construct of trust proved to be highly reliable (Cronbach's $\alpha = .917$), with a partial η^2 of .079. This was also found for perceived risk, with Cronbach's α of .866 and a partial η^2 of .080.

Table 12: Full Model ANCOVA.

	F	Sig.	Partial η^2
Corrected Model	19.996	.000	.278
Intercept	2.397	.122	.004
Trust	48.644	.000*	.079
NFV	3.498	.062	.006
Perceived risk	49.630	.000*	.080
Previous purchase	8.925	.003*	.015
Personality	.772	.380	.001
Involvement	11.857	.001*	.020
Modality	.147	.702	.000
Personality * Involvement	6.009	.015*	.010
Personality * Modality	.114	.736	.000
Involvement * Modality	.261	.610	.000
Personality * Involvement * Modality	.392	.531	.001

* Significant at $\alpha = 0,05$

Exploratory Analysis

Beyond hypotheses-testing and the full model with covariates, we examined whether there were other findings of practical and/or theoretical interest. First, analyzing the effects of covariates pertaining to familiarity with use of digital assistants provided some interesting findings. On one hand, simply owning a digital assistant does not have any significant effects on trust, purchase intention nor risk. On the other hand, previous experience of interaction, and frequency of interaction, do have certain implications. Testing the dichotomized variable of previous interaction with digital assistants in an ANOVA resulted in a significant finding (p = .028) of greater mean trust (M = 5.30) for those who had previously interacted compared to those who had never interacted before (M = 4.95). Perceived risk and purchase intention, however, was not significant. When dividing into three groups based on frequency of previous interactions (low: 1-3, medium: 4-6, high: 7-8), it was revealed that low-frequency users had significantly greater perceptions of risk (M = 4.82) than medium-frequency users (M = 4.36, p = .003) and high-frequency users (M= 4.15, $p < .001$). Again, purchase intention was not significant, whereas trust was only significant at $\alpha = .10$. Regardless, the interesting

tendency here is the greater reported trust for users of higher frequency, as aligned with findings of Schaefer et al. (2016) who explain that trust is calibrated over time through interactions, and that the two constructs are thus intercorrelated.

Second, even though there was a minority of the sample who had previously purchased through digital assistants, the effect of having done so appears important. In general, having previously completed a purchase accounts for a relatively decent amount of the variation in purchase intention (Partial $\eta^2 = .0473$, $p < .001$). Despite this main effect, there were no interaction effects. The mean purchase intention for these users was significantly higher ($M = 6.00$, $p < .001$) compared to users who had interacted with a digital assistant but never purchased ($M = 4.20$). It is worth noting that we cannot securely draw any conclusions from how the manipulations (e.g., main effects or interactions) affected these participants' purchase intention, as only 80 participants had purchased through a digital assistant before, leaving us with statistically 10 previous purchasers per treatment condition. However, we did witness a tendency regarding the means of the various conditions. There was greater purchase intention for the low involvement product regardless of personality and modality, and a high involvement product fared better with a social personality than an intellectual personality. The findings were albeit not significantly different and therefore require additional research.

Third, beyond purchase intention, trust in the digital assistant was also significantly higher for participants who had previously purchased ($M = 5.55$) compared to those who had not ($M = 5.26$, $p = .039$), reflecting the already established relationship between trust and purchase intention. The same tendency is also clear for risk, where those who had purchased previously reported significantly lower perceived risk ($M = 3.56$) than those who had not ($M = 4.58$, $p < .001$). By dividing those who had purchased before into two groups (infrequent vs. frequent buyers) with a cut-off at 4 (4 = "once every other week"), there were no significant differences between the groups on any of the dependent variables, showcasing that frequency of purchase matters less than simply having purchased at all.

Fourth, our risk construct contains two sub-constructs of two items each, in which one sub-construct is very specific and the other is more generic. As such, we divided the risk construct into the separate sub-constructs of 1) “Credit card use and personal information” and 2) “Using the device”, where conducting a t-test would evaluate whether participants rated the two sub-constructs differently. Consequently, the difference was significant ($p < .001$) with sub-construct 1 resulting in lower risk (M = 4.15, SD = 1.62) than sub-construct 2 (M = 4.70, SD = 1.42). Both constructs still portray a relatively high risk. However, the risk ratings might indicate that users have become more familiar with the input of personal information and credit card use, thoughtfully through experience with online shopping, and that other factors related to the general use of digital assistants for purchases are perceived to contain the most risk.

Finally, we tested how the group that had purchased previously was rated on NFV, to determine whether voice shopping could be affected by a strong proclivity to enjoy voice interactions. The one-way ANOVA showed a significant difference ($p < .001$) among groups, where the group with previous purchase experience was higher on NFV (M = 5.72, SD = .74) versus those who had never purchased through a digital assistant (M = 5.39, SD = .63). This is a notable finding subject to further discussions.

General Discussion

Voice shopping is a new method of purchasing products and services, and albeit non-pervasive at the moment, it is considered to be growing substantially in the future (Tennant, 2018; Klaus & Zaichkowsky, 2020; eMarketer, 2020). The field is nascent, which is illustrated by the relatively limited research published. As such, the objective of this study has been to contribute to the field by measuring purchase intention in the context of voice shopping through digital assistants, where involvement level, modality, and personality have been moderators. Furthermore, trust and perceived risk were also measured, as marketing and consumer literature have proved these to be important antecedents to purchase intention. Finally, we have introduced the concept of NFV to determine whether voice interactions contain idiosyncratic differences in individuals' proclivity to use voice interactions for enjoyment, and hence to gather information.

The results prove that the framework regarding trust, perceived risk, and its effect on purchase intention is supported, also in the context of voice shopping. However, the findings show that purchase intention across groups was relatively low where only one hypothesis, considering the main effect for low involvement products having the highest purchase intention, was supported. Furthermore, involvement had a significant interaction with personality, but we did not expect for this interaction to show that the intellectual personality would contribute to the highest purchase intention for a low involvement product, as it contradicts our hypotheses. In terms of modality, there were no significant differences across the groups that were exposed to the treatment conditions of voice stimuli versus voice and visual information combined. The construct of NFV proved to provide significant findings. Participants high on NFV had greater trust and purchase intention than those who were low on trust. Therefore, RQa is supported whereas RQb is partially supported. Another interesting aspect here is that NFV also had an interaction effect with personality, showcasing that social personality gave higher purchase intention and trust for the group of high NFV, compared to the group that was low on NFV. However, trust was the highest for both groups when interacting with an intellectual personality. Two three-way interactions were revealed

between NFV, personality, and modality when measuring trust, and NFV, involvement, and modality when measuring perceived risk. However, as there are no clear patterns nor theoretical support for these findings, future research is needed to study these interactions further to clarify possible relationships. Finally, we witnessed that participants with familiarity and previous experience with voice shopping had a great effect on trust, perceived risk, and purchase intention in our study, which is aligned with findings of Venkatesh et al. (2012), Lee & Turban (2001), Miyazaki & Fernandez (2001) and Lindh et al. (2020). As such, our ANCOVA displayed that trust, perceived risk, and previous purchase experience explained much of the variation in purchase intention, where NFV was not significant at $\alpha = .05$ when controlling for the covariates.

A notable finding contradictory to previous research is that purchase intentions for both the high and low involvement products (respectively laptop and deck of cards) are almost equal when presented in combination with a social personality. Moreover, comparing purchase intention for the laptop under the two different personalities did not provide any significant differences. We were expecting purchase intention to be markedly lower for the laptop compared to the deck of cards, and particularly for the laptop presented by a social personality as opposed to an intellectual personality. This could indicate that a deck of cards is the only product that was considered to be purchased, and that the act of voice shopping is considered risky, where a low involvement product consequently would require an intellectual personality. However, with the social personality showing less discrepancy in purchase intention across groups, we find it difficult to determine which personality is the most suitable under low and high involvement product scenarios when both personalities are competent and deliver equal product information. Hence, there is ambiguity as to which personality to deploy. This could be due to one personality having attributes of the other. For instance, even the social personality of a digital assistant needs to entail a certain degree of intellect, and vice versa. We then contemplate whether each of the personalities appeared prominent enough in the treatment conditions. Through conversations with Elia Gatti, UX Researcher at Amazon, it was discussed that longer interactions with artificial voices can result in listeners perceiving the voices to be more unnatural to the context. Further, Elia suggested that the unnaturalness could be due to

the participants only experiencing a monologue from the device and not a dialogue. The latter could potentially have resulted in a more nuanced impression of the device personality. Finally, it can be discussed whether the personality (social vs. intellectual) itself might matter less when it includes the aspects of a minimum competence level and equal product information, thus removing the importance of the isolated personalities themselves in the context of involvement levels connected to voice shopping.

The interaction between personality and NFV is interesting, as we found that one affects the other in a varied manner. While an intellectual personality has relatively equal purchase intention for both NFV-groups (high/low), a social personality has a significantly higher purchase intention for those high on NFV. The interaction between NFV and personality on trust depicts that the intellectual personality contributes to the highest levels of trust for both NFV-groups. However, the social personality shows greater trust levels for those who are high on NFV than those who are low, compared to the case of intellectual personality. This might find support with the research of Bickmore & Cassell (2005) who found that social dialogue increases trust for extroverts but had no significant difference for introverts. Hence, it might be the case that those individuals that are high on NFV correlate highly with extroversion. On that note, Ehrenbrink et al. (2017) found that extroverts had a better fit with Google Assistant, which Lopatvoska (2020) found to be higher on warmth than e.g., Amazon's Alexa. These findings may contribute to segmentation strategies, where a social personality might be more appropriate for those who are high on NFV, while an intellectual, or another type of personality not tested in this study, is suitable for those who are low on NFV. This would obviously assist with removing the aforementioned ambiguity.

A difference in NFV was revealed between participants with and without previous experience with voice shopping. Those who have purchased previously proved to have a significantly higher level of NFV compared to the group with no previous purchase experience. Although the difference was not substantial, it might indicate that the sense of enjoyment, through biology or experience, triggered by voice interactions enables

people high on NFV to engage in voice shopping. Also, as with our theoretical framework, those who had purchased previously reported markedly higher trust and purchase intention, as well as lower risk than those who had not purchased before. However, for both groups, the difference in modality had no effect. With the issue of limited information conveyed by the digital assistant in a purchase setting (Mari, 2019) and the assumption that future assistants might come with a screen to assist this process (Whang & Im, 2021), we did expect quite different results. One possible explanation for the lack of support of the hypotheses relates to the novelty of the involvement level, in other words the products in question. We have attempted to test products that go outside of the category of reordering, and therefore beyond consumers' habitual purchasing patterns (Moriuchi, 2019).

Today, digital assistants are mainly used to reorder basic, utilitarian products such as groceries (Dellaert et al., 2021; Sun et al., 2019). As these products are in general low involvement, frequently purchased and relatively unimportant, consumers do not find the motivation to engage substantially in decision making and therefore do so only by simple choice rules (Hoyer, 1984). We advocate that voice shopping therefore aligns with perceived auditory control found by Poushneh (2021a). Poushneh (2021a) explains perceived auditory control as the extent of which consumers feel in control during voice-based interactions with a digital assistant. This particular feeling is found to increase both purchase intention (Lu et al, 2016), trust (Poushneh, 2021a), and willingness to take risks (Rousseau et al., 1998). Hence, perceived auditory control might be present when purchasing a familiar product that is low on substitutability and required search (Sun et al., 2019), thus leaving the user in control of the voice interaction. This could help to explain the hesitance of buying products that are new to oneself, as lack of control might harm purchase intention (Poushneh, 2021a). Assuming this is the case, visual information to display a product that is known to them is perhaps not necessary to assist voice shopping. Moreover, when depicting a product that is new to the user, particularly high involvement products, the user is likely to lack perceived auditory control as it entails risk. This could consequently have a negative effect on convenience, enjoyment and ease-of-use, which are important adoption elements (Kowalczuk 2018; Wagner et al. 2019; Moriuchi 2019; Martin et al. 2015), making

visual information insufficient to assist the user in gaining the feeling of perceived auditory control necessary to purchase.

This then raises the question as to whether digital assistants can be used beyond purchasing known and relatively unimportant products. As our findings depict, previous experience and familiarity with voice shopping increases purchase intention for both the deck of cards and the laptop, which might then coincide with perceived auditory control. Hence, to employ digital assistants as a viable sales channel for consumers to purchase low involvement products beyond reordering, and perhaps even high involvement products in the future, consumers would need experience with voice shopping, and consequently increase control. If this is the case, it would likely be a lengthy process, and the diffusion of voice shopping as a normalized sales channel would be much slower than first anticipated, as was the case of online shopping in general (Andrews & Boyle, 2008).

Limitations and Future Research

This study has several limitations, and upon discussing this research with Elia Gatti, several aspects were mentioned. Overall, this related to the lack of naturalness in relation to the personalities and scripts. In reality, voice shopping includes shorter scripts with shorter sentences, where there is a dialogue between user and assistant. Also, the length of the script in this study enables the text-to-speech voices to be perceived less naturally by the user. A final remark by Elia was that the social personality was deemed more unnatural than the intellectual personality due to its configuration. Hence, the aspect of naturalness might stand as a confound, and could therefore have affected the results. Many of these aspects, however, are connected to the overarching and most obvious limitation to this research, which is that we had an online experiment due to the COVID-19 pandemic. With the online experiment, the use of Prolific Academic to recruit participants stands as a potential limitation, as we are unaware of the degree of diversity in the socioeconomic background of our sample, as well as their motives to complete the survey. With that being said, experiments, in general, face the issue of not necessarily capturing real behavior (Gerber & Green, 2012) and could always stand as a limitation. However, not having the opportunity to

conduct a physical experiment, this hinders realism and naturalness, and can consequently affect the results. Clearly, a digital assistant is built on developing a dialogue between the user and the assistant, whereas the online experiment represented a monologue only. Nevertheless, our study presents the user with a choice of whether to purchase. This is a choice the user would have been given regardless of it being a dialogue. It is, however, likely that more questions would have been asked to the digital assistant, which is something we deprived the participants of. On that note, participants explained that perceived risk was not in connection with what the digital assistant said, but what the participant should ask about and whether the digital assistant would discern right from wrong (e.g., confirming the correct product and price). Such dialogue and the effect of confirmation appears rather important and is something we believe should stand as future research in a voice shopping context.

Our findings prove that purchase intention was generally low for all participants, regardless of involvement level. We found this to be somewhat peculiar considering the literature and general news on the topic, providing information that highlights a bright future within voice shopping that will accumulate billions of dollars over the next few years. If purchase intention for a deck of cards is only at 50% chance and below, several questions arise. First, whether the information on the topic truly provides a genuine picture. Second, whether we could make our participants sincerely imagine that they wanted to buy the product and thus rated only the medium for purchase intention, and that their general liking of the product was not included in that rating. The final question is whether our product type of utilitarian products is too risky. This appears to be correct for laptops, in line with our intention. However, a deck of cards cannot intuitively, nor grounded in literature, be deemed very risky. It seems paradoxical that these two products have equal ratings when presented through a social personality. As previously assessed, we do not possess a clear answer as to why that is the case. Based on this, we believe there is a distinct motivation for researchers to develop a clearer picture as to how individuals truly perceive voice shopping, and if the act of purchasing through a digital assistant is yet to be a suitable method for other products beyond reordering. Perhaps the act of voice shopping simply entails too much risk for it to develop in accordance with optimistic estimates of numerous companies

and reports. Another aspect to further examine is the differences between hedonic and functional products, in and without the context of involvement and visual information, to gain greater insights into how consumers behave when confronted with various purchase decisions.

Per existing literature, the way a digital assistant speaks to you must be in accordance with the situation. Based on previous research, involvement level has been deemed a factor that would require a change in arguments and speech. We found on average, higher purchase intention when the product was introduced with a social personality. However, the greatest intention to buy occurred for the low involvement product with an intellectual personality. Hence, there is ambiguity as to which personality should be used for varying involvement levels. We believe this needs to be researched more extensively for digital assistants to be established as a normalized sales channel like that of online stores today.

The construct of NFV has been introduced here to test whether voice interactions are more enjoyable for some individuals than for others. With our three-factor model, it became clear that there is merit in continuing to develop a measure of the construct, as idiosyncratic differences regarding NFV lead to significantly different results in our dependent variables. After having conducted the analysis and utilized the nine items for the main study, it came to our attention that we, unfortunately, overlooked reversing three of the 34 items before conducting the EFA. One of the reversed items in question led to a four-factor model (see Appendix 9 for the four-factor structure matrix). This model included both perception and production of voice, as fitting with the theoretical framework and coincided with the parallel analysis. It explained slightly less variance (71.55) and KMO (.73), but slightly higher Cronbach's α (.788). Thus, in hindsight, the four-factor solution should perhaps have been our model for the main study as more items could potentially have painted a more nuanced picture of NFV as a construct. However, we still witnessed idiosyncratic differences with a three-factor model as well. We regret not being able to test our four-factor solution, but as what we have tested here is merely a starting point for the construct, we hope that our development and testing of NFV inspire other researchers to examine the construct further, as this would

likely alter the scale regardless of our three- and four-factor solutions. This would provide opportunities to develop a solid foundation for segmentation, and to obtain insights into voice shopping and voice interactions.

Lastly, we would like to mention the construct of perceived risk. Perceived risk has been thoroughly researched in the past, but very few articles contain perceived risk with the medium in the context of digital assistants. As such, we encourage researchers to develop a new perceived risk construct for both usage and voice shopping with digital assistants. In this study, we use the construct by Bianchi & Andrews (2012), where the results show significantly higher risk with the general act of purchasing through a digital assistant than storing credit-card and personal information. Furthermore, Kinsella & Mutchler (2018) found 31.72% of respondents were simply uncomfortable with voice shopping. Hence, there are likely many subdimensions that can be measured to get a clearer picture of what this risk entails. An example is the average price of a product that the digital assistant provides when the user is exploring a product through voice shopping. A user could question this average as (s)he does not know how many or which websites the digital assistant has taken into account before communicating the average price point. Another example is related to risk in connection to security, which is deemed to be high with digital assistants (Cowan et al., 2017) and voice shopping (Voorveld & Araujo, 2020). This type of risk is important in online shopping in general (Andrews & Boyle, 2008), but could perhaps be even more crucial with digital assistants. Even though online shopping, through the standard methods of laptops or smartphones, share similarities in risk dimensions as with voice shopping, it might be that these are just as similar as they differ. Hence, we propose that new subdimensions of perceived risk should be developed as a scale specifically for digital assistants.

Theoretical Implications

This study provides several theoretical contributions. First, it is one of the few studies directly testing purchase intention specifically in the field of digital assistants and not product recommendations or other mediatory constructs that may lead to purchase intention. As such, we have tested several factors to investigate how these affect an

individual's propensity to purchase through a digital assistant. Here, the major focus area has been to test visual information, simulating a screen, where we have added personality changes and utilitarian products to vary involvement levels. Other research articles have tested personality changes and involvement to some extent (Poushneh, 2021b; Rhee & Choi, 2020; Whang & Im, 2021), however visual information and direct measurements of purchase intention are new contributions to the field of digital assistants.

Second, our study highlights an ambiguity as to whether a social or an intellectual personality should be used to present a product during voice shopping. Previous literature shows that demanding tasks or risky purchase decisions require a serious personality that provides strong, relevant arguments void of chit-chat and social comments, as individuals are more skeptical and scrutinizing during such interactions (Chattaraman et al., 2019; Petty et al., 1983; Zhou et al., 2019). Our findings do not support this, but rather indicate that personality type (social vs. intellectual) might be less important than the inherent competence and information provided by the digital assistant. Thus, despite literature explaining that a social, sincere, and creative personality is the most favored (Lopatovska, 2020; Pousneh, 2021b; Rhee & Choi, 2020), we believe it is still uncertain which personality to portray in a voice shopping context, as well the importance of this decision.

Third, pretest 2 disclosed that the male voice (Matthew) was better at manifesting both personalities, particularly the social personality, than the female counterpart (Joanna). Recent research shows that female AI is considered more human than male, and is thus preferred by users (Borau et al., 2021). Borau et al. (2021) adds to the literature that people tend to attribute a higher degree of warmth and friendliness to a female AI-bot than a male, where a female bot is better at capturing emotions (Eyssel & Hegel, 2012; Gustavsson, 2005; Otterbacher & Talias, 2017; Stroessner & Benitez, 2019). Our results thus contradict previous literature. With that being said, much previous research is conducted with physical robots in mind, and the research of Borau et al. (2021) tests, among other things, chatbots, and not digital assistants respectively in their research. It is reasonable to assume that a voice can provide more nuances of information and

develop a different impression than what text can. There is not pervasive literature on gendered voices as in the context of digital assistants, but researchers such as Mitchell et al. (2011) found that both men and females prefer female voices when the voice is synthetic. While this might be true, newly developed voices from e.g., Amazon Polly communicate without synthetism and show impressive levels of realism, perhaps changing people's impressions of the voice. Consequently, the gender of the voice might in fact be of less importance, and how the specific voice in question portrays the personality may prove more important. We question whether the capability of manifesting a certain personality is synonymous with preference from the listener to that personality. However, if warmth and sociability is desired, which much literature confirms, one would probably expect so. Based on this, we lay the possibility that another female voice in Amazon Polly is designed to manifest personalities in a better way than Matthew. However, we also suggest that preference of a certain voice is not connected to the gender, but rather to how one specific voice is configured and thereby perceived by the user.

The final theoretical contribution relates to the construct of NFV. Several "need for"-scales have been developed, with NFS being the latest (Dörtyol, 2020). These all play on idiosyncratic differences where a certain sense is more important to assist in decision-making or develop an overall impression for the individual. For example, the NFT-scale (Peck & Childers, 2003) has proven that some individuals need to touch a certain product before purchasing it. The NFV-scale is developed with an intent to be used within the field of marketing. Although its items do not capture purchase behavior directly, they may indicate purchase behavior based on consumer evaluations found in this study. We expected higher NFV to imply more enjoyment with digital assistants, increasing trust and purchase intention. This is what we found, and if it is the case that the construct holds merit in the future, it can most certainly provide clear segmentation strategies both within the context of general interaction with digital assistants as well as with purchases.

Managerial Implications

Although many hypotheses were not supported, there are several interesting findings that derive from the present work. Moreover, not finding supportive evidence is still a finding on its own and should be expressed. The act of purchasing through digital assistants is novel, which is clearly shown through the data. It seems that individuals do not consider digital assistants to be a reliable or even obvious medium for purchase. This is clear not only for high involvement purchases but also for low involvement purchases. Familiarity and experience prove to increase consumers' purchase intention, but until the digital assistant has become more popular and normalized, manufacturers and third-party companies should not focus on selling anything other than low involvement products through the device. For the time being, we believe that such products should be known to the user, not require active search and be of low substitutability or high purchase frequency, as with the findings of Sun et al. (2019). This builds on the reasons behind the adoption of digital assistants; convenience, ease-of-use, and enjoyment (Kowalczuk, 2018; Wagner et al., 2019; Martin et al., 2015). Typical products that fall within this category are groceries and home supplies (Sun et al., 2019), while products that are deemed inconvenient to purchase through the assistant must be excluded. We believe, however, that other low involvement products beyond the aforementioned category can be sold eventually, and that this will be aided by user experience as well as normalization of the digital assistant as a sales channel. When it comes to high involvement products, our evidence shows that purchase intention for this category is low. Thus, we recommend companies to be careful when wishing to sell these types of products. However, we only tested one high involvement product (a laptop), and therefore, replication studies are required to evaluate the potential of selling high involvement products through voice shopping.

Based on our study, visual information does not provide additional benefits for the user in a purchase setting. Nonetheless, this does not necessarily signify that screenless digital assistants should be dismissed. On one hand, our findings imply that digital assistants are still not ready to be used as a method to purchase products beyond reordering and low involvement, due the uncertainty regarding the general use of the assistant. In other words, a screen does not help, either because voice shopping is not

considered an appropriate method to the user, or because the product category is already well-known to the user, making the screen redundant. On the other hand, it might also signify that a digital assistant is perceived by users to only work with voice interaction and that a screen might transform the digital assistant into a device that is not congruent with the expectations of consumers. Developing a digital assistant with a screen would enable more complex tasks that consequently would require users to be more involved with the device. We propose that a considerable amount of such behavior might not take place, as these tasks are already conducted through laptops, smartphones, and tablets (Moriuchi, 2019). Our recommendation is therefore to not implement screens, visualizing pictures and text as with online shopping, for the purpose of voice shopping at the time being. We rather recommend manufacturers to be cognizant of its potential lack of effect, and to determine through future research, customer data, and feedback, whether a screen and this type of visual information, combined with voice, may be favorable in the context of voice shopping.

Conclusion

Technology enables consumer behavior to change markedly, both in the way information is accumulated and how products are purchased. Throughout the last years, the adoption of digital assistants has risen tremendously (Dellaert et al., 2020), and with it, research agendas have laid their eyes on the topic. The field, however, is still in its infancy and requires more in-depth research to evaluate effects, determine best practices, and figure out how the new landscape of digital assistants should develop. As such, we contribute to the research topic by attempting to connect new variables for future growth. Our study evaluated how purchase intention, perceived risk, and trust were affected when exposed to manipulations of modality, involvement, and personality in the context of voice shopping with digital assistants.

Previous research declared that high involvement purchase decisions require a task-oriented personality (Chattaraman et al., 2019; Petty et al., 1983), where additional visual information can assist users throughout the purchase process (Mari, 2019). This is expedient as retaining and understanding information through a digital assistant is demanding due to short-term memory limitations (Klaus & Zaichkowsky, 2020; Bjork,

1970). Whang & Im (2021) thus referred to an expectation that new digital assistants would all be delivered with a screen to hinder this shortcoming. However, our findings show that visual information does not increase purchase intention for neither low nor high involvement products. It might be that there is no room for screens on digital assistants for the time being, as screens require extensive user engagement beyond simple tasks. Further, there is no main effect of personality manipulations, and the interaction between personality and involvement contradicts our hypotheses, portraying ambiguity. We believe that one reason for this is that voice shopping is still too premature, hindering the theoretical framework to function as intended, where individuals who lack experience and familiarity with the assistant feel too much uncertainty associated with purchasing products. However, as consumers get more familiar with the digital assistant, the impact on consumer behavior with digital assistants may change (Sun et al., 2019), and perhaps this familiarity enables easier purchases of products that the consumer has not frequently purchased before. As of now, it seems that consumers wish to only purchase products that fall within the category of reordering, as this is habitual and the product is known to them (Sun et al., 2019; Moriuchi, 2019).

A fascinating construct that we have introduced here, which attempts to explore an unknown territory of voice assistants, is NFV, which proved to have significant findings and supported our research questions. Still, this construct needs to be managed in the future to ensure high validity. If it can be proven to show distinct differences between individuals, it would assist companies with clear segmentation strategies, and lay the ground for future research enabling possible new digital assistant tactics. Conclusively, we would like to mention that we find merit in what Mari (2019) claims, which is that previous ways of purchasing online, such as through a laptop, smartphone, or tablet, provide insufficient information to understand the unique nature of the voice shopping method. Hence, as digital assistants are likely to be even more prevalent in the future, voice shopping and general interactions with the device require deeper attention, where there is a need to ascertain how consumer behavior functions, and to develop new theories to match the peculiarities of this technology (Kumar et al., 2016).

Reference list

Aaker, D. A. (1996). *Building strong brands*. New York: Free Press.

Aaker, J., & Fournier, S. (1995). A Brand as a Character, A Partner and a Person: Three Perspectives on the Question of Brand Personality. *Advances in Consumer Research*, *22*(1), 391–395.

Ackerman, J. M., Nocera, C. C., & Bargh, J. A. (2010). Incidental Haptic Sensations Influence Social Judgments and Decisions. *Science*, *328*(5986), 1712–1715.

Aeschlimann, S., Bleiker, M., Wechner, M., & Gampe, A. (2020). Communicative and social consequences of interactions with voice assistants. *Computers in Human Behavior*, *112*, 106466. https://doi.org/10.1016/j.chb.2020.106466

Akbari, M. (2015). Different Impacts of Advertising Appeals on Advertising Attitude for High and Low Involvement Products. *Global Business Review*, *16*(3), 478–493. https://doi.org/10.1177/0972150915569936

Alaa, M., Zaidan, A. A., Zaidan, B. B., Talal, M., & Kiah, M. L. M. (2017). A review of smart home applications based on Internet of Things. *Journal of Network and Computer Applications*, *97*, 48–65. https://doi.org/10.1016/j.jnca.2017.08.017

Andrews, L., & Boyle, M. V. (2008). Consumers' accounts of perceived risk online and the influence of communication sources. *Qualitative Market Research*, *11*(1), 59–75. http://dx.doi.org.ezproxy.library.bi.no/10.1108/13522750810845559

Andrews, L., Kiel, G., Drennan, J., Boyle, M. V., & Weerawardena, J. (2007). Gendered perceptions of experiential value in using web-based retail channels. *European Journal of Marketing*, *41*(5/6), 640–658. http://dx.doi.org.ezproxy.library.bi.no/10.1108/03090560710737660

Anikin, A., & Persson, T. (2016). Nonlinguistic vocalizations from online amateur videos for emotion research: A validated corpus. *Behavior Research Methods*, *49*(2), 758–771. https://doi.org/10.3758/s13428-016-0736-y

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and

company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Armstrong, J. S., Morwitz, V. G., & Kumar, V. (2000). Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting*, *16*(3), 383–397. https://doi.org/10.1016/S0169-2070(00)00058-3

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Auer, P., Couper-Kuhlen, E., & Muller, F. (1999). *Language in Time: The Rhythm and Tempo of Spoken Interaction*. Oxford University Press, Incorporated. http://ebookcentral.proquest.com/lib/bilibrary/detail.action?docID=271791

Austin, P.L. (2019). *What Will Smart Homes Look Like in 10 Years?* Time. https://time.com/5634791/smart-homes-future/

Auxier, B. (2019). *5 things to know about Americans and their smart speakers*. Pew Research Center. https://www.pewresearch.org/fact-tank/2019/11/21/5-things-to-know-about-americans-and-their-smart-speakers/

Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimensions and Job Performance: A Meta-Analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Bateman, C. R., & Valentine, S. R. (2019). Consumers' Personality Characteristics, Judgment of Salesperson Ethical Treatment, and Nature of Purchase Involvement. *Journal of Business Ethics*, *169*(2), 309–331. https://doi.org/10.1007/s10551-019-04312-7

Bejou, D., Ennew, C. T., & Palmer, A. (1998). Trust, ethics and relationship satisfaction. *International Journal of Bank Marketing*, *16*(4), 170–175. https://doi.org/10.1108/02652329810220729

Belin, P. (2018). The «Vocal» Brain: Core and Extended Cerebral Networks for Voice Processing. I S. Frühholz & P. Belin (eds.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception. *British Journal of Psychology*, *102*(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135. https://doi.org/10.1016/j.tics.2004.01.008

Bentahar, A. (2018). *How Voice Search Is Changing Shopping*. Forbes. https://www.forbes.com/sites/forbesagencycouncil/2018/12/03/how-voice-search-is-changing-shopping/?sh=60a3b97e2615

Bettiga, D., Bianchi, A. M., Lamberti, L., & Noci, G. (2020). Consumers Emotional Responses to Functional and Hedonic Products: A Neuroscience Research. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.559779

Bhatnagar, A., & Ghose, S. (2004). Segmenting consumers based on the benefits and risks of Internet shopping. *Journal of Business Research*, *57*(12), 1352–1360. https://doi.org/10.1016/S0148-2963(03)00067-5

Bhatnagar, A., Misra, S., & Rao, H. R. (2000). On Risk, Convenience, and Internet Shopping Behavior. *Communications of the ACM*, *43*(11), 98–105. https://doi.org/10.1145/353360.353371

Bianchi, C., & Andrews, L. (2012). Risk, trust, and consumer online purchasing behaviour: A Chilean perspective. *International Marketing Review*, *29*(3), 253–275. http://dx.doi.org.ezproxy.library.bi.no/10.1108/02651331211229750

Bickmore, T., & Cassell, J. (2005). Social Dialogue with Embodied Conversational Agents. I J. C. J. van Kuppevelt, L. Dybkjær, & N. O. Bernsen (eds.), *Advances in Natural Multimodal Dialogue Systems* (Bd. 30, p. 23–54). Springer Netherlands. https://doi.org/10.1007/1-4020-3933-6_2

Biswas, D., & Biswas, A. (2004). The diagnostic role of signals in the context of perceived risks in online shopping: Do signals matter more on the Web? *Journal of Interactive Marketing*, *18*(3), 30–45.

Bjork, R. A. (1970). Repetition and Rehearsal Mechanisms in Models for Short-term Memory. Normann, D. A. (ed.). *Models of Human Memory*, 307-330.

Bloch, P. H., & Richins, M. L. (1983). A Theoretical Model for the Study of Product Importance Perceptions. *Journal of Marketing*, *47*(3), 69–81. https://doi.org/10.2307/1251198

Blomqvist, K. (1997). The many faces of trust. *Scandinavian Journal of Management*, *13*(3), 271–286. https://doi.org/10.1016/S0956-5221(97)84644-1

Borau, S., Otterbring, T., Laporte, S., & Fosso Wamba, S. (2021). The most human bot: Female gendering increases humanness perceptions of bots and acceptance of AI. *Psychology & Marketing*, 1-17. https://doi.org/10.1002/mar.21480

Brill, T. M., Munoz, L., & Miller, R. J. (2019). Siri, Alexa, and other digital assistants: A study of customer satisfaction with artificial intelligence applications. *Journal of Marketing Management*, *35*(15–16), 1401–1436. https://doi.org/10.1080/0267257X.2019.1687571

Bruckert, L., Bestelmeyer, P., Latinus, M., Rouger, J., Charest, I., Rousselet, G. A., Kawahara, H., & Belin, P. (2010). Vocal Attractiveness Increases by Averaging. *Current Biology*, *20*(2), 116–120. https://doi.org/10.1016/j.cub.2009.11.034

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16. https://doi.org/10.5334/joc.72

Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: Visual and auditory interactions in the brain. *Current Opinion in Neurobiology*, *16*(4), 415–419. https://doi.org/10.1016/j.conb.2006.06.008

Burns, A. C., Veeck, A., & Bush, R. F. (2016). *Marketing research, global edition* (8th ed.). Pearson Education UK.

Burns, A. S., & Bush, R. F. (2008). *Marketing Research, Global ed of 6 revised ed.* (6th ed.). Pearson Education (US).

Cacioppo, J., & Petty, R. (1982). The Need for Cognition. *Journal of Personality and Social Psychology*, *42*, 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Campbell, R. S., & Pennebaker, J. W. (2003). The Secret Life of Pronouns: Flexibility in Writing Style and Physical Health. *Psychological Science*, *14*(1), 60–65. https://doi.org/10.1111/1467-9280.01419

Carding, P. N., Wilson, J. A., MacKenzie, K., & Deary, I. J. (2009). Measuring voice outcomes: State of the science review. *The Journal of Laryngology & Otology*, *123*(8), 823–829. https://doi.org/10.1017/S0022215109005398

Carmel, D. (2019, 13.05). *On the Relation Between Products' Relevance and Customers' Satisfaction in Voice Shopping* [Paper presentation]. The 2019 World Wide Web Conference, San Francisco, California, USA. https://doi.org/10.1145/3308560.3316610

Cassell, J. (2001). Embodied Conversational Agents: Representation and Intelligence in User Interfaces. *AI Magazine*, *22*(4), 67–67. https://doi.org/10.1609/aimag.v22i4.1593

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research (JMR)*, *56*(5), 809–825. https://doi.org/10.1177/0022243719851788

Cemalcilar, Z., Baruh, L., Kezer, M., Kamiloglu, R. G., & Nigdeli, B. (2018). Role of personality traits in first impressions: An investigation of actual and perceived personality similarity effects on interpersonal attraction across communication modalities. *Journal of Research in Personality*, *76*, 139–149. https://doi.org/10.1016/j.jrp.2018.07.009

Chattaraman, V., Kwon, W.-S., Gilbert, J. E., & Ross, K. (2019). Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Computers in Human Behavior*, *90*, 315–330. https://doi.org/10.1016/j.chb.2018.08.048

Cho, Y. C., & Sagynov, E. (2015). Exploring Factors That Affect Usefulness, Ease of Use, Trust, And Purchase Intention in The Online Environment. *International Journal of Management & Information Systems (Online)*, *19*(1), 21–36.

Clark, E. (2019). *Alexa, are you listening? How people use voice assistants*. Clutch. https://clutch.co/app-developers/resources/alexa-listening-how-people-use-voice-assistants

Clarke, K., & Belk, R. W. (1979). The Effects of Product Involvement and Task Definition on Anticipated Consumer Effort. *Advances in Consumer Research*, *6*(1), 313–318.

Comegys, C., Hannula, M., & Váisánen, J. (2009). Effects of Consumer Trust and Risk on Online Purchase Decision-making: A Comparison of Finnish and United States Students. *International Journal of Management*, *26*(2), 295–308.

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*, 1–12. https://doi.org/10.1145/3098279.3098539

Culley, K. E., & Madhavan, P. (2013). A note of caution regarding anthropomorphism in HCI agents. *Computers in Human Behavior*, *29*(3), 577–579. https://doi.org/10.1016/j.chb.2012.11.023

Dahl, D. W. (2010). Understanding the role of spokesperson voice in broadcast advertising. I A. Krishna (ed.), *Sensory Marketing* (1st edition). Routledge.

Dahlqvist, F., Patel, M., Rajko, A., & Shulman, J. (2019). *Growing opportunities in the Internet of Things*. McKinsey. https://www.mckinsey.com/industries/private-equity-and-principal-investors/our-insights/growing-opportunities-in-the-internet-of-things

Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, *48*(1), 24–42. https://doi.org/10.1007/s11747-019-00696-0

Dawar, N. (2018). Marketing in the Age of Alexa. *Harvard Business Review*. https://hbr.org/2018/05/marketing-in-the-age-of-alexa

Day, D., Gan, B., Gendall, P., & Esslemont, D. (1991). Predicting Purchase Behaviour. *Marketing Bulletin, 2*(3), 18-30.

de Oliveira Neto, J. D., Huang, W. D., & Azevedo Melli, N. C. de. (2015). Online learning: Audio or text? *Educational Technology Research and Development*, *63*(4), 555–573. https://doi.org/10.1007/s11423-015-9392-7

de Vries, R. E., Bakker-Pieper, A., Alting Siberg, R., van Gameren, K., & Vlug, M. (2009). The Content and Dimensionality of Communication Styles. *Communication Research*, *36*(2), 178–206. https://doi.org/10.1177/0093650208330250

Delgado-Ballester, E., & Luis Munuera-Alemán, J. (2001). Brand trust in the context of consumer loyalty. *European Journal of Marketing*, *35*(11/12), 1238–1258. https://doi.org/10.1108/EUM0000000006475

Dellaert, B. G. C., Shu, S. B., Arentze, T. A., Baker, T., Diehl, K., Donkers, B., Fast, N. J., Häubl, G., Johnson, H., Karmarkar, U. R., Oppewal, H., Schmitt, B. H., Schroeder, J., Spiller, S. A., & Steffel, M. (2020). Consumer decisions with artificially intelligent voice assistants. *Marketing Letters*, *31,* 335-347. https://doi.org/10.1007/s11002-020-09537-5

Denes, P. B., & Pinson, E. N. (1993). *The Speech Chain* (2nd ed.). W.H. Freeman. Dependence relations and the moderating role of trust: Implications for behavioral intentions in marketing channels. (1995). *International Journal of Research in Marketing*, *12*(2), 157–172. https://doi.org/10.1016/0167-8116(94)00020-O

Dhar, R., & Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. *JMR, Journal of Marketing Research*, *37*(1), 60–71.

Dholakia, U. M. (2001). A motivational process model of product involvement and consumer risk perception. *European Journal of Marketing*, *35*(11/12), 1340–1360. http://dx.doi.org.ezproxy.library.bi.no/10.1108/EUM0000000006479

Dinev, T., & Hart, P. (2006). An Extended Privacy Calculus Model for E-Commerce Transactions. *Information Systems Research*, *17*(1), 61-80,100.

Drossos, D. A., Kokkinaki, F., Giaglis, G. M., & Fouskas, K. G. (2014). The effects of product involvement and impulse buying on purchase intentions in mobile text advertising. *Electronic Commerce Research and Applications*, *13*(6), 423–430. https://doi.org/10.1016/j.elerap.2014.08.003

Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, *42*(3), 177–190. https://doi.org/10.1016/S0921-8890(02)00374-3

Dörtyol, I. T. (2020). Do consumers need to smell? Scale development and validation. *Journal of Sensory Studies*, 1–13. https://doi.org/10.1111/joss.12630

Ehrenbrink, P., Osman, S., & Möller, S. (2017). Google now is for the extraverted, cortana for the introverted: Investigating the influence of personality on IPA

preference. *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, 257–265. https://doi.org/10.1145/3152771.3152799

eMarketer. (2020). *Purchases Via Smart Speakers Are Not Taking Off*. Insider Intelligence. https://www.emarketer.com/content/purchases-via-smart-speakers-are-not-taking-off

Eyssel, F., & Hegel, F. (2012). (S)he's Got the Look: Gender Stereotyping of Robots1. *Journal of Applied Social Psychology*, *42*(9), 2213–2230. https://doi.org/10.1111/j.1559-1816.2012.00937.x

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory Factor Analysis*. Oxford University Press, Incorporated.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2005). Sensitivity to Voice in Human Prefrontal Cortex. *Journal of Neurophysiology*, *94*(3), 2251–2254. https://doi.org/10.1152/jn.00329.2005

Fennis, B. M., & Stroebe, W. (2016). *The Psychology of Advertising* (2. ed.). Routledge.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Foehr, J., & Germelmann, C. C. (2020). Alexa, Can I Trust You? Exploring Consumer Paths to Trust in Smart Voice-Interaction Technologies. *Journal of the Association for Consumer Research*, *5*(2), 181–205. https://doi.org/10.1086/707731

Frühholz, S., & Belin, P. (2018). The Science of Voice Perception. I S. Frühholz & P. Belin (eds.), *The Oxford Handbook of Voice Perception* (s. 2–14). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.1

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*, *32*(6), 407–424. https://doi.org/10.1016/j.omega.2004.01.006

Gerber, A. S., & Green, D. P. (2012). *Field Experiments. Design, Analysis, and Interpretation.* New York: W. W. Norton & Company.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.*, 55–60. https://doi.org/10.1109/ROMAN.2003.1251796

Gravetter, F. J., & Forzano, L.-A. B. (2016). *Research methods for the behavioral sciences* (5th ed.). Cengage learning.

Greenwald, A. G., & Leavitt, C. (1984). Audience Involvement in Advertising: Four Levels. *Journal of Consumer Research*, *11*(1), 581–592.

Gustavsson, E. (2005). Virtual Servants: Stereotyping Female Front-Office Employees on the Internet. *Gender, Work & Organization*, *12*(5), 400–419. https://doi.org/10.1111/j.1468-0432.2005.00281.

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2019). Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion. *Personality and Social Psychology Bulletin*, *45*(3), 389–405. https://doi.org/10.1177/0146167218787805.

Ha, H.-Y., K. Akamavi, R., J. Kitchen, P., & Janda, S. (2014). Exploring key antecedents of purchase intentions within different services. *The Journal of Services Marketing*, *28*(7), 595–606. http://dx.doi.org.ezproxy.library.bi.no/10.1108/JSM-01-2013-0025

Haselton, T. (2019). *Amazon Alexa is about to get more emotional—Here's what it sounds like excited and disappointed.* CNBC. https://www.cnbc.com/2019/11/28/amazon-alexas-voice-is-about-to-get-more-emotional.html

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research*, *121*, 364–374. https://doi.org/10.1016/j.jbusres.2020.09.020

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic Consumption: Emerging Concepts, Methods and Propositions. *Journal of Marketing*, *46*(3), 92.

Holmes, J. H., & Crocker, K. E. (1987). Predispositions and the comparative effectiveness of rational, emotional and discrepant appeals for both high involvement and low involvement products. *Journal of the Academy of Marketing Science*, *15*(1), 27–35. https://doi.org/10.1007/BF02721951

Holst, A. (2021). *Number of voice assistants in use worldwide 2019-2024*. Statista. https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/

Hong, I. B. (2015). Understanding the consumer's online merchant selection process: The roles of product involvement, perceived risk, and trust expectation. *International Journal of Information Management*, *35*(3), 322–336. https://doi.org/10.1016/j.ijinfomgt.2015.01.003

Howard, J. A., & Sheth, J. N. (1969). *The Theory of Buyer Behavior*. John Wiley.

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, *37*(1), 81–88. https://doi.org/10.1080/02763869.2018.1404391

Hoyer, W. D. (1984). An Examination of Consumer Decision Making for a Common Repeat Purchase Product. *Journal of Consumer Research*, *11*(3), 822–829. https://doi.org/10.1086/209017

Hsiao, K., Chuan-Chuan Lin, J., Wang, X., Lu, H., & Yu, H. (2010). Antecedents and consequences of trust in online product recommendations: An empirical study in social shopping. *Online Information Review*, *34*(6), 935–953. https://doi.org/10.1108/14684521011099414

Hupfer, N. T., & Gardner, D. M. (1971). Differential Involvement With Products and Issues: An Exploratory Study. *ACR Special Volumes*, *SV-01*. https://www.acrwebsite.org/volumes/11965/volumes/sv01/SV-01/full

Jacoby, J., & Kaplan, L. (1972). The components of perceived risk. *Advances in Consumer Research*, *3*, 382–383.

Jarvenpaa, S. L., Tractinsky, N., & Saarinen, L. (1999). Consumer Trust in an Internet Store: A Cross-Cultural Validation. *Journal of Computer-Mediated Communication*, *5*(2), 0–0. https://doi.org/10.1111/j.1083-6101.1999.tb00337.

Juster, F. T. (1966). Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design. *Journal of the American Statistical Association*, *61*(315), 658–696. https://doi.org/10.2307/2282779

Kaplan, L. B., Szybillo, G. J., & Jacoby, J. (1974). Components of perceived risk in product purchase: A cross-validation. *Journal of Applied Psychology*, *59*(3), 287–291. https://doi.org/10.1037/h0036657

Keeling, K., McGoldrick, P., & Beatty, S. (2010). Avatars as salespeople: Communication style, trust, and intentions. *Journal of Business Research*, *63*(8), 793–800. https://doi.org/10.1016/j.jbusres.2008.12.015

Keller, L. K. (2013). *Strategic Brand Management. Building, Measuring, and Managing Brand Equity.* Pearson Education Limited.

Kelley, H. H. (1950). The Warm-Cold Variable in First Impressions of Persons. *Journal of Personality*, *18*(4), 431–439. https://doi.org/10.1111/j.1467-6494.1950.tb01260.x

Kelley, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care*, *15*(3), 261–266. https://doi.org/10.1093/intqhc/mzg031

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*, 544-564. https://doi.org/10.1016/j.dss.2007.07.001

Kinsella, B., & Mutchler, A. (2018). *Voice Shopping Consumer Adoption Report*. https://voicebot.ai/wp-content/uploads/2018/06/voice-shopping-consumer-adoption-report-june-2018-voicebot-voysis.pdf

Klatzky, R. L. (2010). Touch: A gentle tutorial with implications for marketing. I *Sensory marketing: Research on the sensuality of products* (s. 33–47). Routledge/Taylor & Francis Group.

Klaus, P., & Zaichkowsky, J. (2020). AI voice bots: A services marketing research agenda. *Journal of Services Marketing*, *34*(3), 389–398. https://doi.org/10.1108/JSM-01-2019-0043

Klaus, P., & Zaichkowsky, J. L. (2021). The convenience of shopping via voice AI: Introducing AIDM. *Journal of Retailing and Consumer Services*, 102490. https://doi.org/10.1016/j.jretconser.2021.102490

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings: Biological Sciences*, *279*(1738), 2698–2704.

Knoeferle, K., & Spence, C. (2021). *Sound in the Context of (Multi)Sensory Marketing*. The Oxford Handbook of Music and Advertising. https://doi.org/10.1093/oxfordhb/9780190691240.013.28

Kowalczuk, P. (2018). Consumer acceptance of smart speakers: A mixed methods approach. *Journal of Research in Interactive Marketing*, *12*(4), 418–431. https://doi.org/10.1108/JRIM-01-2018-0022

Kraus, N., & Slater, J. (2016). Beyond Words: How Humans Communicate Through Sound. *Annual Review of Psychology*, *67*(1), 83–103. https://doi.org/10.1146/annurev-psych-122414-033318

Kreiman, J. (1997). Listening to Voices: Theory and Practice in Voice Perception Research. K. Johnson & Mullenix, J. (eds.), *Talker Variability in Speech Research.* Academic Press

Kreiman, J., & Gerratt, B. R. (2018). Reconsidering the Nature of Voice. I., S., Frühholz & P., Belin (eds.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*(3), 332–351. https://doi.org/10.1016/j.jcps.2011.08.003

Krishna, A. (2019). Note on Sensory Marketing: Shaping Consumer Perception and Behavior. *WDI Publishing, University of Michigan*, 1-16.

Krishna, A., Cian, L., & Sokolova, T. (2016). The power of sensory marketing in advertising. *Current Opinion in Psychology*, *10*, 142–147. https://doi.org/10.1016/j.copsyc.2016.01.007

Krugman, H. E. (1966). The Measurement of Advertising Involvement. *The Public Opinion Quarterly*, *30*(4), 583–596.

Kumar, V., Dixit, A., Javalgi, R. G, & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, *44*(1), 24–45. http://dx.doi.org.ezproxy.library.bi.no/10.1007/s11747-015-0426-9

Latinus, M., & Zäske, R. (2018). Perceptual Correlates and Cerebral Representation of Voices - Identity, Gender, and Age. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Lee, K. M., & Nass, C. (2004). The Multiple Source Effect and Synthesized Speech.: Doubly-Disembodied Language as a Conceptual Framework. *Human Communication Research*, *30*(2), 182–207. https://doi.org/10.1111/j.1468-2958.2004.tb00730.x

Lee, M. K. O., & Turban, E. (2001). A Trust Model for Consumer Internet Shopping. *International Journal of Electronic Commerce*, *6*(1), 75–91. https://doi.org/10.1080/10864415.2001.11044227

Lindh, C., Rovira Nordman, E., Melén Hånell, S., Safari, A., & Hadjikhani, A. (2020). Digitalization and International Online Sales: Antecedents of Purchase Intent. *Journal of International Consumer Marketing*, *32*(4), 324–335. https://doi.org/10.1080/08961530.2019.1707143

Liu, Q., Zhang, X., Huang, S., Zhang, L., & Zhao, Y. (2020). Exploring Consumers' Buying Behavior in a Large Online Promotion Activity: The Role of Psychological Distance and Involvement. *Journal of Theoretical and Applied Electronic Commerce Research; Curicó*, *15*(1), 66–80. http://dx.doi.org.ezproxy.library.bi.no/10.4067/S0718-18762020000100106

Lopatovska, I. (2020). Personality Dimensions of Intelligent Personal Assistants. *Proceedings of the 2020 Conference on Human Information Interaction and Retrieval*, 333–337. https://doi.org/10.1145/3343413.3377993

Lopatovska, I., Griffin, A. L., Gallagher, K., Ballingall, C., Rock, C., & Velazquez, M. (2020). User recommendations for intelligent personal assistants. *Journal of Librarianship and Information Science*, *52*(2), 577–591. https://doi.org/10.1177/0961000619841107

Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., Williams, H., Sorsche, P., Hirsch, D., Li, Q., & Martinez, A. (2019). Talk to me: Exploring user

interactions with the Amazon Alexa. *Journal of Librarianship and Information Science*, *51*(4), 984–997. https://doi.org/10.1177/0961000618759414

Lu, B., Fan, W., & Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: An empirical research. *Computers in Human Behavior*, *56*, 225–237. https://doi.org/10.1016/j.chb.2015.11.057

Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, *37*, 94–100. https://doi.org/10.1016/j.chb.2014.04.043

Luger, E., & Sellen, A. (2016). «Like Having a Really Bad PA»: The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5286–5297. https://doi.org/10.1145/2858036.2858288

Luhmann, N., Davis, H., Raffan, J., Rooney, K., King, M., & Morgner, C. (2017). *Trust and Power*. Polity Press. http://ebookcentral.proquest.com/lib/bilibrary/detail.action?docID=5108561

Malhotra, N. K. (2005). Attitude and affect: New frontiers of research in the 21st century. *Journal of Business Research*, *58*(4), 477–482. https://doi.org/10.1016/S0148-2963(03)00146-2

Malhotra, N. K., Birks, D. K., & Wills, P. (2010). *Marketing research: An applied orientation* (6th European ed.). Pearson Education Limited.

Malär, L., Krohmer, H., Hoyer, W. D., & Nyffenegger, B. (2011). Emotional Brand Attachment and Brand Personality: The Relative Importance of the Actual and the Ideal Self. *Journal of Marketing*, *75*(4), 35–52. https://doi.org/10.1509/jmkg.75.4.35

Mari, A. (2019). *Voice Commerce: Understanding Shopping-Related Voice Assistants and their Effect on Brands*. In: IMMAA Annual Conference. Northwestern University in Qatar, Doha (Qatar). October 4-6, 2019. https://www.researchgate.net/publication/336363485_Voice_Commerce_Understanding_Shopping-Related_Voice_Assistants_and_their_Effect_on_Brands

Martin, J., Mortimer, G., & Andrews, L. (2015). Re-examining online customer experience to include purchase frequency and perceived risk. *Journal of*

Retailing and Consumer Services, *25*, 81–95. https://doi.org/10.1016/j.jretconser.2015.03.008

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.2307/258792

McAleer, P., & Belin, P. (2018). The Perception of Personality Traits from Voices. I S. Frühholz & P. Belin (Red.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

McCracken, G. (1989). Who Is the Celebrity Endorser? Cultural Foundations of the Endorsement Process. *Journal of Consumer Research*, *16*(3), 310–321. https://doi.org/10.1086/209217

McDonald, R. P. (1985). *Factor Analysis and Related Methods*. Taylor & Francis Group.

McGonegal, C. A., Rabiner, L. R., & McDermott, B. J. (1978). Speaker Verification by Human Listeners over Several Speech Transmission Systems. *Bell System Technical Journal*, *57*(8), 2887–2900. https://doi.org/10.1002/j.1538-7305.1978.tb02181.x

McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, *99*, 28–37. https://doi.org/10.1016/j.chb.2019.05.009

Mitchell, W. J., Ho, C.-C., Patel, H., & MacDorman, K. F. (2011). Does social desirability bias favor humans? Explicit–implicit evaluations of synthesized speech support a new HCI model of impression management. *Computers in Human Behavior*, *27*(1), 402–412. https://doi.org/10.1016/j.chb.2010.09.002

Mittal, B. (1989). Measuring Purchase-Decision Involvement: ABSTRACT. *Psychology & Marketing (1986-1998)*, *6*(2), 147.

Miyazaki, A. D., & Fernandez, A. (2001). Consumer Perceptions of Privacy and Security Risks for Online Shopping. *Journal of Consumer Affairs*, *35*(1), 27–44. https://doi.org/10.1111/j.1745-6606.2001.tb00101.x

Mooi, E., Sarstedt, M., & Mooi-Reci, I. (2018). *Market Research. Principal Component and Factor Analysis.* Springer Nature Singapore Pte Ltd.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, *58*(3), 20.

Moriuchi, E. (2019). Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty. *Psychology & Marketing*, *36*(5), 489–501. https://doi.org/10.1002/mar.21192

Munz, K., & Morwitz, V. (2019). Not-so Easy Listening: Roots and Repercussions of Auditory Choice Difficulty in Voice Commerce. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3462714

Nass, C., Steuer, J., & Tauber, E. R. (1994, 24.04). *Computers are Social Actors* [Paper presentation]. The 1994 Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA. https://doi.org/10.1145/259963.260288

Nauts, S., Langner, O., Huijsmans, I., Vonk, R., & Wigboldus, D. H. J. (2014). Forming impressions of personality: A replication and review of Asch's (1946) evidence for a primacy-of-warmth effect in impression formation. *Social Psychology*, *45*(3), 153–163. https://doi.org/10.1027/1864-9335/a000179

Newman, D. T., Fast, N. J., & Harmon, D. J. (2020). When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes*, *160*, 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008

Niculescu, A., van Dijk, B., Nijholt, A., Li, H., & See, S. L. (2013). Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy. *International Journal of Social Robotics*, *5*(2), 171–191. https://doi.org/10.1007/s12369-012-0171-x

Novak, T. P., & Hoffman, D. L. (2019). Relationship journeys in the internet of things: A new framework for understanding interactions between consumers and smart objects. *Journal of the Academy of Marketing Science*, *47*(2), 216–237. https://doi.org/10.1007/s11747-018-0608-3

Novak, T. P., Hoffman, D. L., & Yung, Y.-F. (2000). Measuring the Customer Experience in Online Environments: A Structural Modeling Approach. *Marketing Science*, *19*(1), 22–42. https://doi.org/10.1287/mksc.19.1.22.15184

Nuszbaum, M., Voss, A., Klauer, K. C., & Betsch, T. (2010). Assessing Individual Differences in the Use of Haptic Information Using a German Translation of the Need for Touch Scale. *Social Psychology*, *41*(4), 263–274. https://doi.org/10.1027/1864-9335/a000035

Otterbacher, J., & Talias, M. (2017, 05.03). *S/he's too Warm/Agentic!: The Influence of Gender on Uncanny Reactions to Robots* [Paper presentation]. The 2017 ACM/IEEE International Conference on Human-Robot Interaction, Vienna, Austria. https://doi.org/10.1145/2909824.3020220

Pagani, M., Racat, M., & Hofacker, C. F. (2019). Adding Voice to the Omnichannel and How that Affects Brand Trust. *Journal of Interactive Marketing*, *48*, 89–105. https://doi.org/10.1016/j.intmar.2019.05.002

Park, C. W., & Young, S. M. (1986). Consumer Response to Television Commercials: The Impact of Involvement and Background Music on Brand Attitude Formation. *Journal of Marketing Research*, *23*(1), 11–24. https://doi.org/10.2307/3151772

Pavlou, P. A., Liang, H., & Xue, Y. (2007). Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective. *MIS Quarterly*, *31*(1), 105–136. https://doi.org/10.2307/25148783

Peck, J., & Childers, T. L. (2003). Individual Differences in Haptic Information Processing: The «Need for Touch» Scale. *Journal of Consumer Research*, *30*(3), 430–442. https://doi.org/10.1086/378619

Pelsmacker, P. de, Geuens, M., & Bergh, J. van den. (u.å.). *Marketing communications; a European perspective* (5th ed.). Pearson. https://content.talisaspire.com/binorway/bundles/5bb72e7469df50042b7e59f4

Perez, S. (2018). *Voice Shopping estimated to hit $40+ billion across U.S. and U.K in 2022*. Techcrunch. https://techcrunch.com/2018/03/02/voice-shopping-estimated-to-hit-40-billion-across-u-s-and-u-k-by-2022/?guccounter=1

Perrodin, C., & Petkov, C. I. (2018). Voice-Sensitive Regions, Neurons, and Multisensory Pathways in the Primate Brain. I S. Frühholz & P. Belin (eds.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Peter, J. P., & Ryan, M. J. (1976). An Investigation of Perceived Risk at the Brand Level. *Journal of Marketing Research*, *13*(2), 184–188. https://doi.org/10.2307/3150856

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research*, *10*(2), 135–146.

Peynircioğlu, Z. F., Brent, W., Tatz, J. R., & Wyatt, J. (2017). McGurk Effect in Gender Identification: Vision Trumps Audition in Voice Judgments. *The Journal of General Psychology*, *144*(1), 59–68. https://doi.org/10.1080/00221309.2016.1258388

Pham, M. T. (1998). Representativeness, Relevance, and the Use of Feelings in Decision Making. *Journal of Consumer Research*, *25*(2), 144–159. https://doi.org/10.1086/209532

Pisanski, K., & Feinberg, D. R. (2018). Vocal Attractiveness. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, *38*(4), 626-642. https://doi.org/10.1002/mar.21457

Pitt, M. A., & Samuel, A. G. (1990). The Use of Rhythm in Attending to Speech. *Journal of Experimental Psychology. Human Perception and Performance*, *16*(3), 564–573. https://doi.org/10.1037/0096-1523.16.3.564

Platz, C. (2017). *Voice User Interface Design: New Solutions to Old Problems.* Medium. https://medium.com/microsoft-design/voice-user-interface-design-new-solutions-to-old-problems-baa36a64b3e4

Plummer, J. T. (2000). How Personality Makes a Difference. *Journal of Advertising Research*, *40*(6), 79–83. https://doi.org/10.2501/JAR-40-6-79-83

Poddar, A., Donthu, N., & Wei, Y. (2009). Website customer orientations, Web site quality, and purchase intentions: The role of Web site personality. *Journal of Business Research*, *62*(4), 441–450. https://doi.org/10.1016/j.jbusres.2008.01.036

Poushneh, A. (2021a). Impact of auditory sense on trust and brand affect through auditory social interaction and control. *Journal of Retailing and Consumer Services*, *58*, 102281. https://doi.org/10.1016/j.jretconser.2020.102281

Poushneh, A. (2021b). Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *Journal of Retailing and Consumer Services*, *58*, 102283. https://doi.org/10.1016/j.jretconser.2020.102283

Purwanto, P., Kuswandi, K., & Fatmah, F. (2020). Interactive Applications with Artificial Intelligence: The Role of Trust among Digital Assistant Users. *Foresight and STI Governance*, *14*(2), 64–75. http://dx.doi.org.ezproxy.library.bi.no/10.17323/2500-2597.2020.2.64.75

PwC. (2018). *Consumer Intelligence Series—Prepare for the voice revolution*. https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/voice-assistants.pdf

Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145–181.

Quittner, A. L., Smith, L. B., Osberger, M. J., Mitchell, T. V., & Katz, D. B. (1994). The Impact of Audition on the Development of Visual Attention. *Psychological Science (0956-7976)*, *5*(6), 347–353. https://doi.org/10.1111/j.1467-9280.1994.tb00284.x

Radder, L., & Huang, W. (2008). High-involvement and low-involvement products: A comparison of brand awareness among students at a South African university. *Journal of Fashion Marketing and Management*, *12*(2), 232–243. http://dx.doi.org.ezproxy.library.bi.no/10.1108/13612020810874908

Raphael, J. (2019). *The downfall of the virtual assistant (so far)*. Computer World. https://www.computerworld.com/article/3403332/downfall-virtual-assistant.html

Rathborn, H. A., Bull R. H., & Clifford, B. R. (1981). Voice Recognition over the Telephone. *Journal of Police Science & Administration*, *9*(3), 280-284.

Reeves, B., & Nass, C. (1996). *The Media Equation*. Cambridge University Press.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, *49*(1), 95–112. https://doi.org/10.1037/0022-3514.49.1.95

Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, *109*, 106359. https://doi.org/10.1016/j.chb.2020.106359

Rice, R. E., & Case, D. (1983). Electronic Message Systems in the University: A Description of Use and Utility. *Journal of Communication*, *33*(1), 131–152. https://doi.org/10.1111/j.1460-2466.1983.tb02380.x

Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, *6*(2), 135–147. https://doi.org/10.1016/j.edurev.2010.12.001

Richins, M. L., & Bloch, P. H. (1986). After the New Wears off: The Temporal Context of Product Involvement. *Journal of Consumer Research*, *13*(2), 280–285.

Richins, M. L., & Dawson, S. (1992). A Consumer Values Orientation for Materialism and Its Measurement: Scale Development and Validation. *Journal of Consumer Research*, *19*(3), 303–316.

Risley, J. (2015). *One year after Amazon introduced echo, half a million people have told Alexa, «I love you»*. Geekwire. https://www.geekwire.com/2015/one-year-after-amazon-introduced-echo-half-a-million-people-have-told-alexa-i-love-you/

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. https://doi.org/10.1037/h0026086

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so Different After All: A Cross-Discipline View of Trust. *Academy of Management Review*, *23*(3), 393–404. https://doi.org/10.5465/AMR.1998.926617

Russell, C. A. (2002). Investigating the Effectiveness of Product Placements in Television Shows: The Role of Modality and Plot Connection Congruence on

Brand Memory and Attitude. *Journal of Consumer Research*, *29*(3), 306–318. https://doi.org/10.1086/344432

Sapir, E. (1927). Speech as a Personality Trait. *American Journal of Sociology*, *32*(6), 892–905. https://doi.org/10.1086/214279

Schaefer, K. E., Chen, J. Y. C., Szalma, J. L., & Hancock, P. A. (2016). A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Understanding Autonomy in Future Systems. *Human Factors*, *58*(3), 377–400. https://doi.org/10.1177/0018720816634228

Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport, 16*(6), 635 - 639. https://doi.org/10.1097/00001756-200504250-00024

Schlosser, K. (2019). *Alexa, why so emotional? Because Amazon has created realistic new voice tones for its assistant*. Geekwire. https://www.geekwire.com/2019/alexa-emotional-amazon-created-realistic-new-voice-tones-assistant/

Schneider, C. (2017). *10 Reasons why AI-powered, automated customer service is the future*. https://www.ibm.com/blogs/watson/2017/10/10-reasons-ai-powered-automated-customer-service-future/

Schroeder, J., & Epley, N. (2015). The Sound of Intellect: Speech Reveals a Thoughtful Mind, Increasing a Job Candidate's Appeal. *Psychological Science*, *26*(6), 877–891. https://doi.org/10.1177/0956797615572906

Schwartz, E. H. (2019). *Alexa is Learning to Speak Emotionally*. Voicebot. https://voicebot.ai/2019/11/27/alexa-is-learning-to-speak-emotionally/

Shead, S. (2017). *Report: 1 in 4 people have fantasised about Alexa, Siri, and other AI assistants*. Business Insider. https://www.businessinsider.com/jwt-speak-easy-study-people-fantasised-about-alexa-2017-4

Siegrist, M., Cvetkovich, G., & Roth, C. (2000). Salient Value Similarity, Social Trust, and Risk/Benefit Perception. *Risk Analysis: An International Journal*, *20*(3), 353–362. https://doi.org/10.1111/0272-4332.203034

Simms, K. (2019). *How Voice Assistants Could Change the Way We Shop*. Harvard Business Review. https://hbr.org/2019/05/how-voice-assistants-could-change-the-way-we-shop

Skantze, G. (2016). Real-Time Coordination in Human-Robot Interaction Using Face and Voice. *AI Magazine*, *37*(4), 19–31.

Smith, K. T. (2020a). Marketing via smart speakers: What should Alexa say? *Journal of Strategic Marketing*, *28*(4), 350–365. https://doi.org/10.1080/0965254X.2018.1541924

Smith, M. S. (2020b). *Less is actually less: 5 reasons you should spend more on your next laptop*. Digitaltrends. https://www.digitaltrends.com/computing/dont-be-so-cheap-five-reasons-you-should-spend-more-on-your-next-laptop/

Smith, R., Deitz, G., Royne, M. B., Hansen, J. D., Grünhagen, M., & Witte, C. (2013). Cross-cultural examination of online shopping behavior: A comparison of Norway, Germany, and the United States. *Journal of Business Research*, *66*(3), 328–335. https://doi.org/10.1016/j.jbusres.2011.08.013

Solomon, M. R. (1986). Deep-Seated Materialism: The Case of Levi's 501 Jeans. *Advances in Consumer Research*, *13*(1), 619–622.

Stewart, K., Kammer-Kerwick, M., Auchter, A., Koh, H. E., Dunn, M. E., & Cunningham, I. (2019). Examining digital video advertising (DVA) effectiveness: The role of product category, product involvement, and device. *European Journal of Marketing*, *53*(11), 2451–2479. https://doi.org/10.1108/EJM-11-2016-0619

Stone, R. N., & Grønhaug, K. (1993). Perceived Risk: Further Considerations for the Marketing Discipline. *European Journal of Marketing*, *27*(3), 39–50. https://doi.org/10.1108/03090569310026637

Stroessner, S. J., & Benitez, J. (2019). The Social Perception of Humanoid and Non-Humanoid Robots: Effects of Gendered and Machinelike Features. *International Journal of Social Robotics*, *11*(2), 305–315. https://doi.org/10.1007/s12369-018-0502-7

Sun, C., Shi, Z., Liu, X., Ghose, A., Li, X. & Xiong, F. (2019). The Effect of Voice AI on Consumer Purchase and Search Behavior. *SSRN Electronic Journal, 21 (114*). https://doi.org/10.2139/ssrn.3480877

Sung-Joon Yoon. (2002). The Antecedents and Consequences of Trust in Online-Purchase Decisions. *Journal of Interactive Marketing*, *16*(2), 47–63. https://doi.org/10.1002/dir.10008

Taylor, J. W. (1974). The Role of Risk in Consumer Behavior. *Journal of Marketing*, *38*(2), 54–60. https://doi.org/10.2307/1250198

Taylor, M. P., Jacobs, K., KVJ, S., Shah, H., Girard, S., Buvat, J., Puttur, R., & B., A. (2019). *Smart talk. How organizations and consumers are embracing voice and chat assistants.* Capgemini.

Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence.* Borzoi Book.

Tennant, E. (2018). Voice Shopping Is Revolutionizing E-Commerce. *TWICE: This Week in Consumer Electronics*, *33*(16), 9.

Valishery, L. S. (2021). *Prognosis of worldwide spending on the Internet of Things (IoT) from 2018 to 2023*. Statista. https://www.statista.com/statistics/668996/worldwide-expenditures-for-the-internet-of-things/

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Venkatraman, M. P. (1989). Involvement and Risk: Abstract. *Psychology & Marketing (1986-1998)*, *6*(3), 229.

Vercruyssen, M., & Hendrick, H. W. (2011). *Behavioral research and analysis: An introduction to statistics within the context of Experimental Design.* (4th ed.). Taylor & Francis Group.

Verhagen, T., Meents, S., & Tan, Y.-H. (2006). Perceived risk and trust associated with purchasing at electronic marketplaces. *European Journal of Information Systems*, *15*(6), 542–555. https://doi.org/10.1057/palgrave.ejis.3000644

Verplanken, B., & Svenson, O. (1997). Personal involvement in human decision making: Conceptualisations and effects on decision processes. I R. Crozier, R. Ranyard, & O. Svenson (eds.), *Decision making: Cognitive models and Explanations*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/bilibrary/detail.action?docID=180169

von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal

content. *Cognitive Brain Research*, *17*(1), 48–55. https://doi.org/10.1016/S0926-6410(03)00079-X

Voorveld, H. A. M., & Araujo, T. (2020). How Social Cues in Virtual Assistants Influence Concerns and Persuasion: The Role of Voice and a Human Name. *Cyberpsychology, Behavior, and Social Networking*, *23*(10), 689–696. https://doi.org/10.1089/cyber.2019.0205

Wagner, K., Nimmermann, F., & Schramm-Klein, H. (2019). *Is It Human? The Role of Anthropomorphism as a Driver for the Successful Acceptance of Digital Voice Assistants*, 1386-1395.

Wang, S. W., Ngamsiriudom, W., & Hsieh, C.-H. (2015). Trust disposition, trust antecedents, trust, and behavioral intention. *Service Industries Journal*, *35*(10), 555–572. https://doi.org/10.1080/02642069.2015.1047827

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation.* Cambridge University Press.

Webster, M., & Sell, J. (2014). *Laboratory Experiments in the Social Sciences* (2nd edition). Elsevier Science & Technology.

Whang, C., & Im, H. (2021). «I Like Your Suggestion!» The role of humanlikeness and parasocial relationship on the website versus voice shopper's perception of recommendations. *Psychology & Marketing*, *38*(4), 581–595. https://doi.org/10.1002/mar.21437

Whitehead, J. C., & Armony, J. L. (2018). Amygdala Processing of Vocal Emotions. I S. Frühholz & P. Belin (ed.), *The Oxford Handbook of Voice Perception*. Oxford University Press.

Wright, M., & Macrae, M. (2007). Bias and variability in purchase intention scales. *Journal of the Academy of Marketing Science*, *35*(4), 617–624. http://dx.doi.org.ezproxy.library.bi.no/10.1007/s11747-007-0049-x

Yagoda, R. E., & Gillan, D. J. (2012). You Want Me to Trust a ROBOT? The Development of a Human–Robot Interaction Trust Scale. *International Journal of Social Robotics*, *4*(3), 235–248. https://doi.org/10.1007/s12369-012-0144-0

Yarmey, A. D. (1993). Stereotypes and recognition memory for faces and voices of good guys and bad guys. *Applied Cognitive Psychology*, *7*(5), 419–431. https://doi.org/10.1002/acp.2350070505

Yoo, W., Kim, S. Y., Hong, Y., Chih, M.-Y., Shah, D. V., & Gustafson, D. H. (2015). Patient–Clinician Mobile Communication: Analyzing Text Messaging Between Adolescents with Asthma and Nurse Case Managers. *Telemedicine Journal and e-Health*, *21*(1), 62–69. https://doi.org/10.1089/tmj.2013.0359

Zaichkowsky, J. L. (1985). Measuring the Involvement Construct. *Journal of Consumer Research*, *12*, 13.

Zaichkowsky, J. L. (1994). The personal involvement inventory: Reduction, revision, and. *Journal of Advertising*, *23*(4), 59.

Zhou, M. X., Mark, G., Li, J., & Yang, H. (2019). Trusting Virtual Agents: The Effect of Personality. *ACM Transactions on Interactive Intelligent Systems*, *9*(2–3), 1–36. https://doi.org/10.1145/3232077.

Appendices

Appendix 1: NFV-items.

1.1 Original 34 NFV-items and Inspiration.

Items	Inspiration
Listening to the voice of others can be fun	Peck & Childers (2003)
Using my voice in interaction with objects can be fun (e.g. talking to your car or your plants)	Peck & Childers (2003)
I find myself using my voice with all kinds of objects (e.g. talking to your car or your plants)	Peck & Childers (2003)
When listening to a salesperson who has an unpleasant voice, I am reluctant to purchase	Peck & Childers (2003)
When I listen to a pleasant voice, it elicit positive emotions	Dörtyol (2020)
When I listen to a pleasant voice, it increases my trust in the person who speaks	Dörtyol (2020)
My attention towards a person increases when their voice appears congruent with their personality	Dörtyol (2020)
My decision making can be affected by the voice of the person talking to me (e.g: pleasant voice vs unpleasant voice)	Dörtyol (2020)
My purchase intention increases, when the salesperson I talk to has a voice congruent with the ambience of the store	Dörtyol (2020)
I experience a sense of satisfaction when I hear certain voices	Dörtyol (2020)
I think a person with an unpleasant voice reduces how comfortable I am	Dörtyol (2020)
I find it an effort to use my voice in communication with others	Carding et al. (2009)
I find it an effort to listen to the voice of others	Carding et al. (2009)
Using my voice in interaction with others is easy for me	Klaus & Zaichkowsky (2021)
I elicit positive emotions when I can use my voice in a conversation	Klaus & Zaichkowsky (2021)
I think that talking to someone is more convenient than other forms of communication	Klaus & Zaichkowsky (2021)
I tend to be able to assess personal characteristics of the speaker when listening to that person's voice	Kreiman (1997)
I am able to remember a voice the next time I hear it, even though I have only been exposed to it once	Kreiman (1997)
When listening to a voice, I tend to be quite confident in assessing whether I have heard that specific voice before	Kreiman (1997)
I tend to desire that a person I talk to, has a voice that is congruent with their identity	Kreimann & Gerratt (2018)
I am confident that when I speak, my voice carries information about my self-identity	Kreimann & Gerratt (2018)
I can sometimes react to the voice of a person sounding differently in real life than on the phone	Kreiman (1997); McGonegal et al. (1978); Rathborn et al. (1981)
I quickly determine if another person's voice is attractive	Pisanski & Feinberg (2018)
An attractive voice is important for me in order to be attracted to that someone	Pisanski & Feinberg (2018)
I'd rather listen to voice through an audiobook than reading text in a physical book	de Oliveira Neto et al. (2015)
Receiving information through voice is the preferred way	de Oliveira Neto et al. (2015)
It is easier for me to learn when listening to information through voice rather than reading text	de Oliveira Neto et al. (2015) + Lee & Nass (2004)
I am sensitive in detecting other peoples' voices around me	Belin (2018)
I tend to be able to infer how people feel based on their voice	Belin (2018)
I can perfectly fine listen to the voice of a person I cannot see without feeling a need of seeing the face of that person	Belin et al. (2004)
Sometimes I struggle in finding the right social response when using my voice in conversation	Whitehead & Armony (2018)
I can infer the emotions that a person is expressing through their voice	Whitehead & Armony (2018); Frühholz & Belin (2018)
Listening to a pleasant voice is delightful	Moriuchi (2019)
When I listen to a voice, I can easily infer the age of the speaker	Latinus & Zäske (2018)

1.2 Extracted NFV-items from Pretest.

Items	Inspiration	Factors
When I listen to a pleasant voice, it elicit positive emotions	Dörtyol (2020)	1. Voice & Hedonic Properties
Listening to a pleasant voice is delightful	Moriuchi (2019)	
I experience a sense of satisfaction when I hear certain voices	Dörtyol (2020)	
I tend to be able to assess personal characteristics of the speaker when listening to that person's voice	Kreiman (1997)	2. Inference Based on Voice
I can infer the emotions that a person is expressing through their voice	Whitehead & Armony (2018); Frühholz & Belin (2018)	
I tend to be able to infer how people feel based on their voice	Belin (2018)	
When I listen to a pleasant voice, it increases my trust in the person who speaks	Dörtyol (2020)	3. Attribution by Voice
I tend to desire that a person I talk to, has a voice that is congruent with their identity	Kreimann & Gerratt (2018)	
My attention towards a person increases when their voice appears congruent with their personality	Dörtyol (2020)	

Appendix 2: Audio Files for "Pretest 2: Personality Manipulations".

2.1. Joanna, Social personality:

https://soundcloud.com/stian-eriksen-235667574/joanna-social-pretest/s-nOMIpTaVMMG

2.2. Joanna, Intellectual personality:

https://soundcloud.com/stian-eriksen-235667574/joanna-intellectual-pretest/s-4hFJH9KbFFw

2.3. Matthew, Social personality:

https://soundcloud.com/stian-eriksen-235667574/matthew-social-pretest/s-qObV4iLqVJ1

2.4. Matthew, Intellectual personality:

https://soundcloud.com/stian-eriksen-235667574/matthew-intellectual-pretest/s-oirtZ0Al62I

Appendix 3: Audio Files and Scripts for the Main Study.

3.1. Social personality, Low involvement:

https://soundcloud.com/stian-eriksen-235667574/social-cards/s-mQ43ckMdjN3

Hey there! I'm Matthew! I am your very own digital assistant! How good to have you here today! How about I show you a deck of playing cards that you have been considering buying. I am very excited, I hope you are too! You will definitely like the beautiful scenery of nature on the background of each card. Also, the cards are created in plastic, which makes them last longer and won't get dirty as fast as cards made out of cardboard! Oh! There is one final thing I want to mention. They come in plastic packaging, so that you can store the cards securely. This deck of playing cards cost $5, which is the average compared to other playing cards in the market. Doesn't that sound good? Don't worry about delivery! Based on your profile, I will have the deck of playing cards sent to your closest post office within 3-5 working days, completely free of charge! What do you say? Should I put the deck of playing cards in your shopping cart?

3.2. Social personality, High involvement:

https://soundcloud.com/stian-eriksen-235667574/social-laptop/s-meXFTwtkcJ0

Hey there! I'm Matthew! I am your very own digital assistant! How good to have you here today! How about I show you a laptop that you have been considering buying? I am very excited, I hope you are too! The laptop has a black sleek design with a 13-inch screen and 12 hours of battery life, which are really great features that I think you will appreciate! These will let you watch full-HD videos without having to connect the charger for a really long time! Also, with 8 gigabytes of RAM, a 256 GB SSD hard drive and an i5-intel processor, you will have great storage and be able to do all kinds of multitasking. Oh! There are a few final things I want to mention. These are the secure Face-ID, the comfortable keyboard, and the ports. This laptop has lightning bolt, HDMI, two times USB-C and a USB-A, which will have you connected at all times! The price is $700, which is the average price compared to other laptops in the market. I hope that sounds good! Don't worry about delivery! Based on your profile, I will have

the laptop sent to your closest post office within 3-5 working days, completely free of charge! What do you say? Should I put the laptop in your shopping cart?

3.3. Intellectual personality, Low involvement: **https://soundcloud.com/stian-eriksen-235667574/intellectual-cards/s-HXNOBBu4Lm6**

Hi, my name is Matthew, your digital assistant. You will now be introduced to a deck of playing cards that you have been considering purchasing. First, the background of each card is filled with a scenery of nature. Second, the cards are manufactured in plastic, making them more manageable than cards of cardboard and mitigating the probability of the cards becoming dirty. Lastly, the cards come in a plastic packaging for secure storage. This deck of playing cards costs $5, which is the average compared to other playing cards in the market. Based on the information that is stored in your profile, this product can be sent to your closest post office within approximately 3-5 working days. There will be no shipping fee. Please, take your time, and consider if you want me to add the deck of playing cards to your shopping cart.

3.4. Intellectual personality, High involvement: **https://soundcloud.com/stian-eriksen-235667574/intellectual-laptop/s-PfkMR4F0Q8g**

Hi, my name is Matthew, your digital assistant. You will now be introduced to a laptop that you have been considering purchasing. First, the laptop has a black, sleek design with a 13-inch screen and 12 hours of battery life. This will ensure that you can watch full-HD videos without the necessity of always bringing a charger. Second, the laptop provides 8 Gigabytes of RAM, a 256 GB SSD hard drive and an i5-intel processor, which enables satisfactory storage capabilities and gives you the opportunity to multitask with several programs simultaneously. Lastly, it comes with advanced and secure authentication through face-ID, a comfortable keyboard, and several ports. These ports include lightning bolt, HDMI, 2 times USB-C and one USB-A. This laptop costs $700, which is the average price compared to other laptops in the market. Based on the information that is stored in your profile, this product can be sent to your closest post office within approximately 3-5 working days. There will be no shipping fee.

Please, take your time, and consider if you want me to add the laptop to your shopping cart.

Appendix 4: Pictures of High and Low Involvement Products.

High involvement: Laptop

Low involvement: Deck of Cards

Appendix 5: Video Files per Treatment Condition in the Main Study.

5.1. Voice, Social personality, Low involvement:

https://youtu.be/SWLnd_kCP9I

5.2. Voice, Social personality, High involvement:

https://youtu.be/X6HSaTrj1hM

5.3. Voice, Intellectual personality, Low involvement:

https://youtu.be/cyrc105nYuA

5.4. Voice, Intellectual personality, High involvement:

https://youtu.be/hUPRkKvAsnA

5.5. Voice and Visual, Social Personality, Low involvement:

https://youtu.be/NYXyU70vI9g

5.6. Voice and Visual, Social Personality, High involvement:

https://youtu.be/60EIHOrvQdw

5.7. Voice and Visual, Intellectual Personality, Low involvement:

https://youtu.be/B8mPKWTOCn0

5.8. Voice and Visual, Intellectual Personality, High involvement:

https://youtu.be/ARX0Z5sv5p0

Appendix 6: Main Study Questionnaire.

Block: Introduction (3 Questions)
Standard: Intro to introductory questions (1 Question)
Standard: Endogenous variables (5 Questions)
Standard: Intro to susceptibility to voice (1 Question)
Standard: Susceptibility to voice (10 Questions)
Standard: Intro to treatment (2 Questions)
Standard: Treatments (8 Questions)
Standard: Purchase Intent (1 Question)
Standard: Intro to trust and risk (1 Question)
Standard: Trust (2 Questions)
Standard: Demographics (2 Questions)
Standard: Closing statement (1 Question)

Start of Block: Introduction

Q1.
Welcome to our study!

In this study, by MSc. students Stian Eriksen and Henrik Christensen at BI Norwegian Business School, we are interested in examining people's interaction and behavior with digital assistants/voice assistants (such as Google Home, Amazon Alexa, Microsoft Cortana or Apple Homepod). If you decide to take part, you are to (1) answer some general questions (2) listen to audio files, and (3) rate your level of agreement with statements following the audio files. There are no right or wrong responses, please respond according to what feels right to you. We hope you enjoy it. Your participation is very important to us.

NB! This study requires you to turn up the volume on your device.

Q2.
Before starting, please read the below consent statement and indicate whether you consent.If you have any questions, concerns, or wish to make a complaint, please contact us at: stian.eriksen5@student.bi.no / henrik.christensen@student.bi.no

Statement of Informed Consent:1. I have read and understand the information about this experiment and understand its general purpose.2. I understand that I can withdraw from the questionnaire at any time, for any reason, and without penalty, and that doing so will delete my data. (NB – Please be aware that data collected is anonymous – if you change your mind after completing the experiment, we will be unable to trace your specific data to remove it).3. I understand how to raise a concern or make a complaint (detailed above).4. I understand that my responses are anonymous.

Q3.
I agree with the above information and wish to partake in this online experiment.

- o Yes (1)
- o No (2)

Skip To: End of Survey If Consent = 2

End of Block: Introduction

Start of Block: Intro to introductory questions

Q4.
The first questions will pertain to your experience and familiarity with digital assistants (Google Home, Amazon Alexa, Microsoft Cortana, Apple Homepod).

End of Block: Intro to introductory questions

Start of Block: Endogenous variables

Ownership: Do you, or anyone in your household, own a digital assistant?

- o Yes (1)
- o No (0)

Interaction 1.0.: Have you ever interacted with a digital assistant?

- o Yes (1)

o No (0)

Display This Question:

If Interacting with a DA = 1

Interaction 2.0: How often do you interact with your digital assistant?

o Several times a day (8)

o Once a day (7)

o Several times a week (6)

o Once a week (5)

o Once every other week (4)

o Once a month (3)

o Less than once a month (2)

o I do not interact with it anymore (1)

Display This Question:

If Interacting with a DA = 1

Purchase 1.0: Have you ever made a purchase through a digital assistant?

o Yes (1)

o No (0)

Display This Question:

If Purchases through DA = 1

Purchase 1.1: How often do you purchase through a digital assistant?

o Several times a day (8)

o Once a day (7)

- Several times a week (6)
- Once a week (5)
- Once every other week (4)
- Once a month (3)
- Less than once a month (2)
- I do not purchase through it anymore (1)

End of Block: Endogenous variables

Start of Block: Intro to susceptibility to voice

Q10.

The following questions pertain to your perception of voice.

Please answer what feels right to you.

End of Block: Intro to susceptibility to voice

Start of Block: Susceptibility to voice

NFV - 1: When listening to a pleasant voice, it elicits positive emotions

- Strongly agree (7)
- Agree (6)
- Somewhat agree (5)
- Neither agree nor disagree (4)
- Somewhat disagree (3)
- Disagree (2)
- Strongly disagree (1)

NFV - 2: Listening to a pleasant voice is delightful

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 3: I experience a sense of satisfaction when I hear certain voices

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

Attention check: This is an attention check. Please answer "Agree"

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 4: I tend to be able to assess personal characteristics of the speaker when listening to that person's voice

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 5: I can infer the emotions that a person is expressing through their voice

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 6: I tend to be able to infer how people feel based on their voice

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)

- o Strongly disagree (1)

NFV - 7: I tend to desire that a person I talk to has a voice that is congruent with their identity

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 8: My attention towards a person increases when their voice appears congruent with their personality

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)
- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

NFV - 9: When I listen to a pleasant voice, it increases my trust in the person who speaks

- o Strongly agree (7)
- o Agree (6)
- o Somewhat agree (5)

- o Neither agree nor disagree (4)
- o Somewhat disagree (3)
- o Disagree (2)
- o Strongly disagree (1)

End of Block: Susceptibility to voice

Start of Block: Intro to treatment

SOUNDCHECK

Below is a soundcheck to ensure that your sound is functioning.

Please listen to the audio file and fill in the numbers you hear.
https://soundcloud.com/stian-eriksen-235667574/soundcheck-joanna/s-lmAqtGbP59I

PageBreak

Story
You are momentarily about to listen to an audio file which is attached to a short video. You can play the file as many times as you would like.

Please read the following information before you start:

Imagine that you have been, in the last week, showing interest in a certain product (checking websites, searching for it online etc.). Imagine further that you now sit at home in your living room with your digital assistant in front of you, which already has your personal information stored (address, phone, payment). The product that you have shown interest in, has become a product that you desire to purchase. You now start a dialogue with your digital assistant.

- ☐ I have read the information (1)

End of Block: Intro to treatment

Start of Block: Treatments

Please play the video and pay attention. If you are watching this on a smartphone, click the "full screen" button in the lower right corner and turn the phone horizontally.

- o I have listened to/watched the video

See Appendix 6 for links to all the eight video files.

End of Block: Treatments

Start of Block: Purchase Intention

Purchase intention
What do you think are the chances that you will purchase this product through the digital assistant?

- o Absolutely certain to buy (10)
- o Almost certain to buy (9)
- o Much better than even chance (8)
- o Somewhat better than even chance (7)
- o Slightly better than even chance (6)
- o About even chance (50-50) (5)
- o Slightly less than even chance (4)
- o Somewhat less than even chance (3)
- o Much less than even chance (2)
- o Almost no chance (1)
- o Absolutely no chance (0)

End of Block: Purchase Intent

Start of Block: Intro to trust and risk

Q31.

The next part contains statements pertaining to your experience with the digital assistant you were just exposed to.

End of Block: Intro to trust and risk

Start of Block: Trust

Trust: Please rate to which extent you agree with the following statements:

	Strongly agree (7)	Agree (6)	Somewhat agree (5)	Neither agree nor disagree (4)	Somewhat disagree (3)	Disagree (2)	Strongly disagree (1)
I feel that the digital assistant made truthful claims (1)	o	o	o	o	o	o	o
I feel that the digital assistant was trustworthy (2)	o	o	o	o	o	o	o
I believe what the digital assistant told me (3)	o	o	o	o	o	o	o

Risk: Please rate to which extent you agree with the following statements:

	Strongly agree (7)	Agree (6)	Somewhat agree (5)	Neither agree nor disagree (4)	Somewhat disagree (3)	Disagree (2)	Strongly disagree (1)
I feel safe making purchases through the digital assistant using my credit card (1)	o	o	o	o	o	o	o
I feel safe allowing my personal information to be kept by the digital assistant for purchases (2)	o	o	o	o	o	o	o

Compared with other ways of making purchases, I think that using the digital assistant is more risky (3)	o	o	o	o	o	o	o
There is too much uncertainty associated with using the digital assistant to make purchases (4)	o	o	o	o	o	o	o

End of Block: Trust

Start of Block: Demographics

Gender: What is your gender?

- o Male (1)
- o Female (2)
- o Non-binary / third gender (3)
- o Prefer not to say (4)

Age: How old are you?

- o 18-24 (1)
- o 25-29 (2)
- o 30-35 (3)
- o 36-40 (4)
- o 41-44 (5)
- o Prefer not to say (6)

End of Block: Demographics

Start of Block: Closing statement

Q36.
Thank you for completing this online experiment.

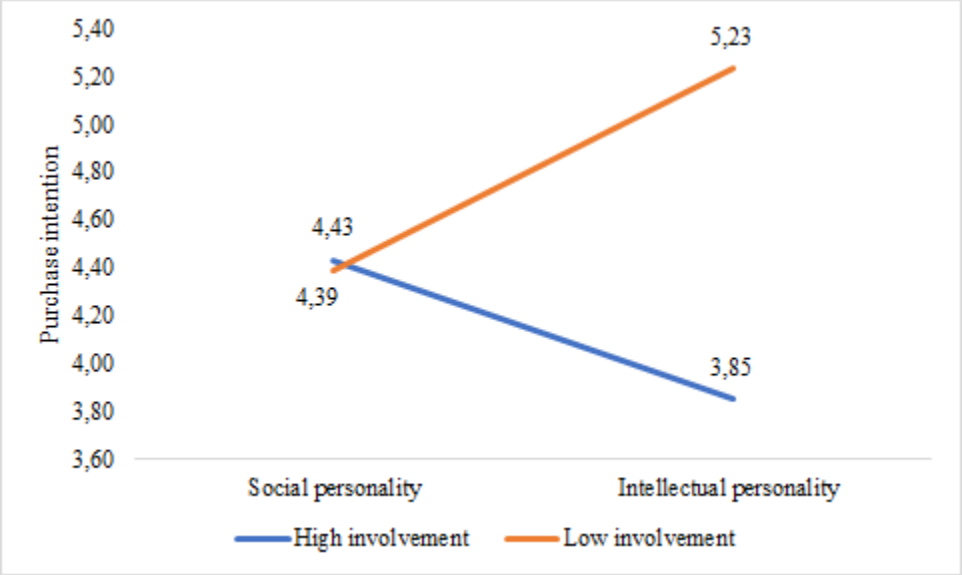
To remind you, if you have any questions or wish to make a complaint, please contact us at stian.eriksen5@student.bi.no / henrik.christensen@student.bi.no.

By pressing continue in the bottom right corner you will automatically be redirected back to Prolific to register your participation.

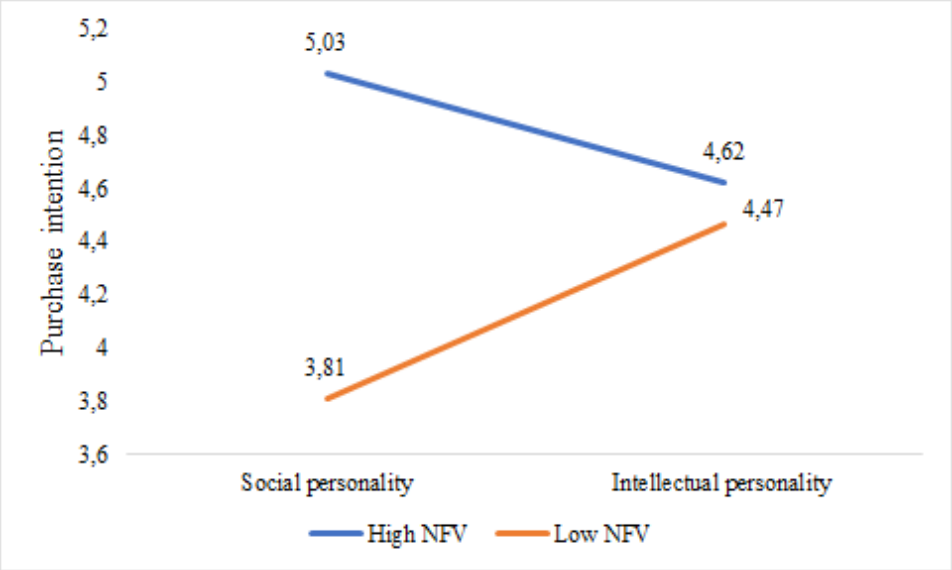
End of Block: Closing statement

Appendix 7: Interaction Plots.

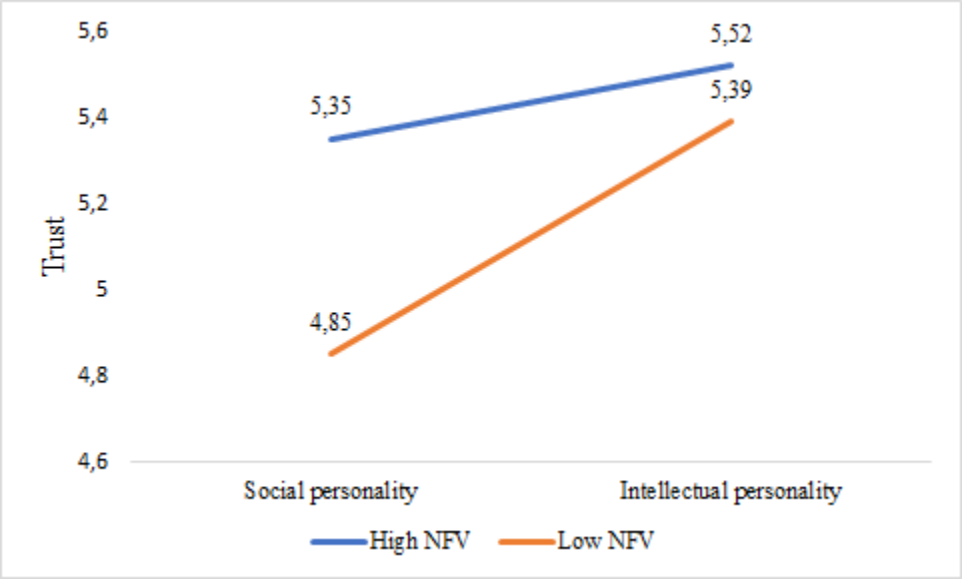
7.1 : Two-way Interaction: Personality and Involvement on Purchase Intention.



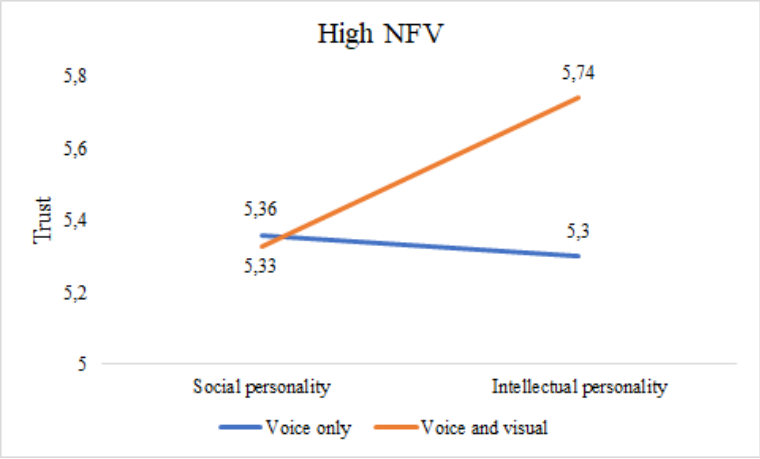
7.2 : Two-way Interaction: Personality and NFV on Purchase Intention.

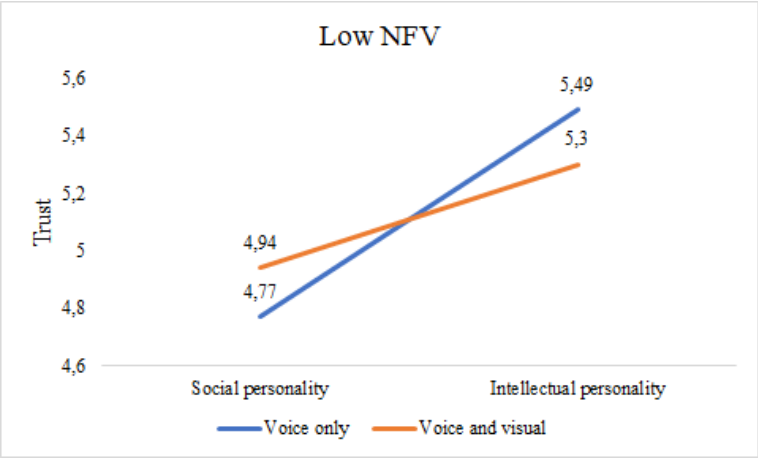


7.3. Two-way Interaction: Personality and NFV on Trust.

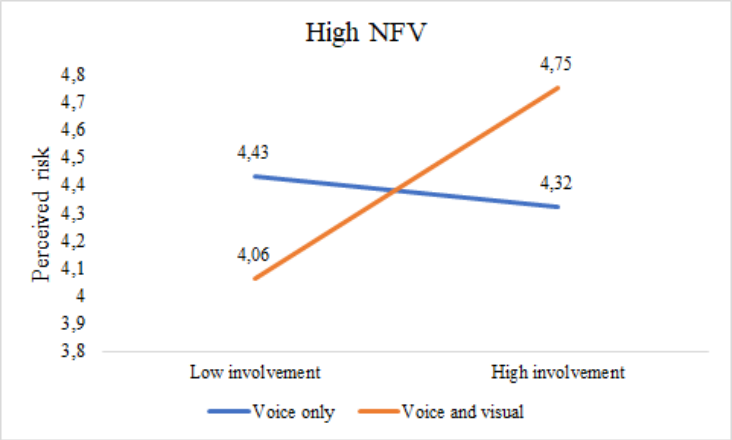


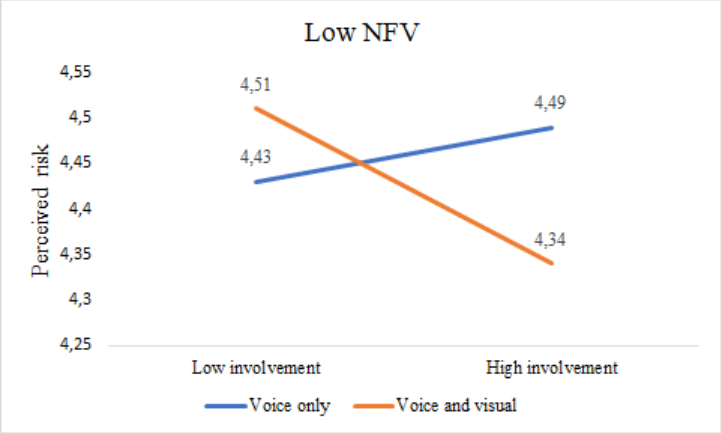
7.4. Three-way Interaction: NFV, Personality, and Modality on Trust.





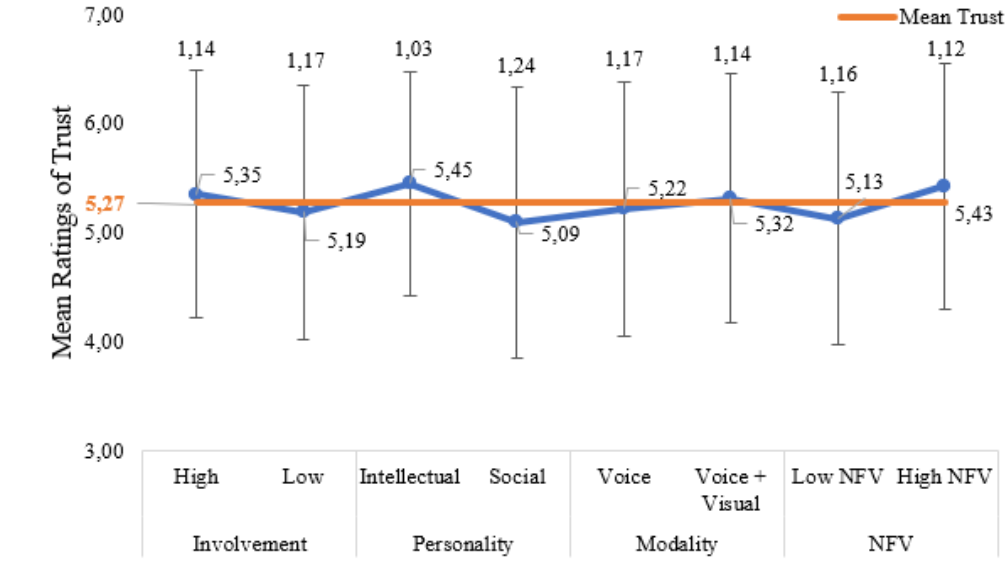
7.5. Three-way Interaction: NFV, Involvement, and Modality on Perceived Risk.



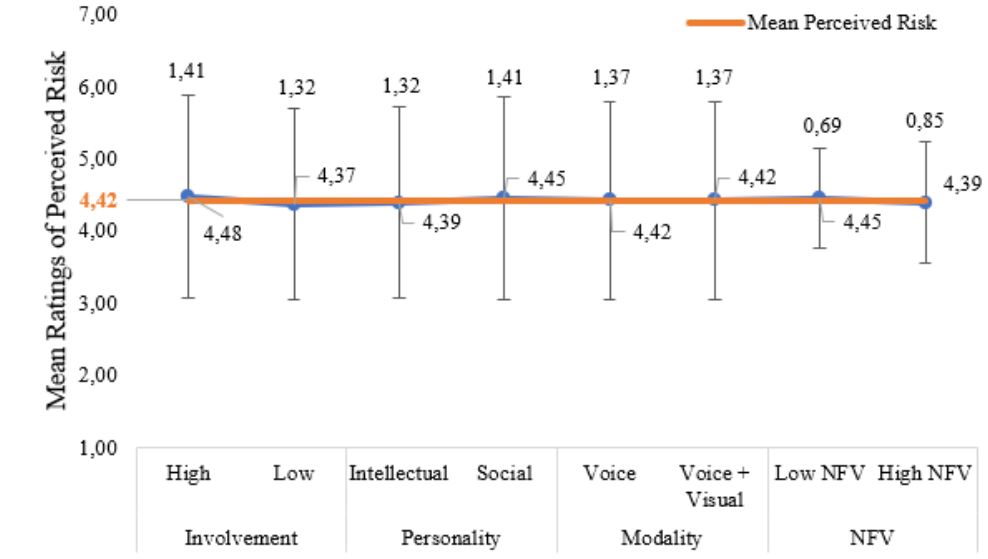


Appendix 8: Summary Figures of Independent Variables on Trust and Perceived Risk.

8.1. Summary Figure: Trust



8.1. Summary Figure: Perceived Risk



Appendix 9: Structure Matrix EFA - Four-factor Solution.

Items	Factors			
	1	2	3	4
When I listen to a pleasant voice, it elicits positive emotions	.858			
Listening to a pleasant voice is delightful	.868			
I experience a sense of satisfaction when I hear certain voices	.682			
I tend to be able to assess personal characteristics of the speaker when listening to that person's voice		.673		
I can infer the emotions that a person is expressing through their voice		.731		
I tend to be able to infer how people feel based on their voice		.891		
When I listen to a pleasant voice, it increases my trust in the person who speaks			.591	
I tend to desire that a person I talk to has a voice that is congruent with their identity			.749	
My attention towards a person increases when their voice appears congruent with their personality			.767	
Using my voice in interaction with others is easy for me				.808
Sometimes I struggle in finding the right social response when using my voice in interaction				.692
I am confident that when I speak, my voice carries information about my self-identity				.631

Extraction Method: Maximum Likelihood.
Rotation Method: Promax with Kaiser Normalization.