Department of Impresa and Management

Chair of Consumer Behaviour

"Servant" or "Partner" Voice Assistant: How user satisfaction is affected by the fit between presentation and behaviour of the voice assistant

SUPERVISOR

Prof.ssa Simona Romani

COSUPERVISOR

Prof. Rumen Ivaylov Pozharliev

CANDIDATE

Ilaria Merico Matr. 706401

ACADEMIC YEAR 2019/2020

Index

Introduction

In a contemporary scenario where we experience technological evolutions and revolutions, one of the most avant-garde discoveries like artificial intelligence joins realities that mark the daily life of individuals (Internet-Of-Things). There are many applications of the IOT technologies in sectors like home automated systems (smart home), wearable (smart watches), health & fitness (smart health), mobility (smart car), system of distribution (smart retail) and manufacturing (smart factory) and city administration (smart city). All these sectors aim at providing users with a superior and enhanced experience, that can facilitate everyday life due to the contribution of such an advanced technology. In particular, we are to explore the field of smartphone-integrated personal assistants that are based on Natural Language Processing. Such devices are responsive to human voice as they can elaborate and replicate it though pre-programmed artificial vocal tracks.

As relatively new, this is an area that has been little explored. The present dissertation is to deepen the issue concerning how user experience can be enhanced by voice assistants and their components. In particular, we found that anthropomorphising plays a crucial role. In fact, individuals naturally tend to view machines in a human perspective by assigning them a typically human personality made up by different factors. As a result, they try to find human elements also in electronic devices in order to be able to more easily relate positively with them. Accordingly, scholars previously studied how people establish more or less strong relationships with their smart assistant, since it is perceived to play a role associated with human sphere. The main roles that emerged from previous research were partner, servant and master Thus, we assumed that smart assistants can be given typically human roles, i.e. partner or servant, showing certain personality traits, i.e. dominant or submissive.

Therefore, we focused on more anthropomorphic traits of smart assistants to be stressed in order to improve the experience that users are going to live. Based on a literature review, we identified some elements in the way the assistant is presented and behaves. In fact, they can positively influence the overall satisfaction related to the interaction undertaken with smart speaker. We chose three of them due to their strong manipulating power on user perception of the speaker: firstly, the wording selected for both the description and the responses to user requests; then, the vocal characteristics implemented in the interaction; lastly, the type of experience provided to users. In particular, we want to verify whether they can exert this influencing power alone, together or both. In fact, the interaction between presentation and behaviour will be the focal point of the present research: we will test whether behaviour is supposed to positively affect the relationship between presentation and user satisfaction.

1 Framing Smart Voice Assistant in a contemporary scenario

1.1 Smart Voice Assistants Background: IoT and Smart Objects

Nowadays, we are living the so-called Fourth Industrial Revolution which originated in 2011 with a German project. The latter can be considered as a prototype of the modern smart factories. In fact, it was intended to make manufacturing industries more efficient through technological improvements. The implementation of World Wide Web seems to be fundamental for such a revolution. The contemporary era is marked by the presence of the so-called Web 4.0, but it passed through three evolutionary stages (Hoffman & Novak, 2015): firstly, anyone could access the global network in complete anonymity as exemplified by the publication in the New Yorker in 1993, stating the following: "On the internet nobody knows you're a dog"; from 2006 onwards, the introduction of Facebook and other social media contributed to upset the web scene due to the widespread awareness of being exposed on a large scale ("On the internet everybody knows you're a dog" which means that a social profile can be quickly found, and the owner can be easily identified); lastly, from 2014, the Internet of Things system occurred. It constitutes of nonlinear interactions connecting heterogeneous entities in the digital world with equally heterogeneous entities in the physical world. By citing Hoffman & Novak, the unpredictable facet of such a state-of-the-art technology is well captured by the phrase "On the Internet of Things, nobody knows you're a fridge". Moreover, the nature of the Internet leads people to connect to each other over physical limits. Thus, it reaches its maximum expression in this revolutionary scenario. In fact, not only people, but also physical inanimate entities can connect to each other through technological contribution. Physical world and digital world are connecting to each other in order to generate a single cooperative system based on intelligent devices and internet connection. That is the Internet of Things background.

The first person who introduced the concept of Internet of Things was Kevin Ashton, at the time researcher at Procter and Gamble and now co-founder of the Auto-ID Center at the Massachusetts Institute of Technology (MIT). In 1999 he provided an embryonic definition of " the Internet-of-Things thing" that he himself resumed only ten years later as he effectively understood its social, economic and historical relevance (Ashton, 2009). A more conscious definition comes only in 2010 from the work of Chui, Loffler e Roberts. From the publications of Ashton K. (2009) and Chun *et al* (2010), we can derive some common elements in order to provide a comprehensive and exhaustive personal reflection of the IoT, typical abbreviation for Internet of Things. It can be considered a system of instruments with *"*different degrees of complexity*"* since it involves both sensors or mechanisms (e.g. Radio Frequency Identification, RFID) and physical objects in which the former where embodied; its components are interconnected through a wired or wireless network which

usually uses an Internet Protocol (IP) and can transmit and receive data and information without human intervention.

The Internet of Things has been defined as the most "potentially disruptive technological revolution of the last 50 years" since it developed empowering and increasingly pervasive machines like RFID or sensors which, once combined with deep-rooted mechanisms and practices (like distributed computing or artificial intelligence), exhibited their revolutionary contribution (Atzori, Iera, & Morabito, 2010).

The development of recent technologies contributed to the spread of IoT since they are used on a daily basis even by unaware users. For example, the Radio Frequency Identification (RFID) allows to transmit to a reader data stored in particular electronic labels called tags via radio waves. It is the mechanism behind anti-theft devices, tele pass and digital passports. The NFC (Near Field Communication) is even more recent than RFID and consists in the short-range (a few centimetres) transmission of data from an initiator to a target. ATM cards, public transport subscription and any contactless card work like this. Also, 2D barcodes can systemically store and transmit date. Eventually, our cars or smartphone have embedded GPS which exploit radio waves from satellites for geolocating purposes.

The revolutionary impact of this interconnected system lies in the combination of elements of a different nature which mutually interact and create an overall harmonious *unicum*. Thus, the above-mentioned embedded empowering potential which makes the IoT system such innovative is supplied by the smart objects (SOs) composing it. In fact, SO-based IoT has been widely studied due to its internal cooperative dynamics among connected objects aimed at delivering "evolved services to humans or to other objects" (Kortuem, Kawsar , Fitton , & Sundramoorthy, 2010). Particularly, IoT has been also referred to as a "loosely coupled, decentralized system of smart objects" (Fortino , Guerrieri , Russo , & Savaglio, 2014). Instead, a smart object has been defined by Fortino *et al*. (2014) as an "autonomous, cyber-physical object augmented with sensing/actuating, processing, storing, and networking capabilities." As we mentioned before, these intelligent devices are characterized by the network connectivity which allows interaction and communication (i.e. transmission of data and information) - with each other or with users – and which can take the form of Wi-fi, Bluetooth, RFID codes and more. Particularly, Atzori, Iera e Morabito analysed the path of development undertaken by smart objects and they identified three evolutive stages (Atzori, Iera, & Morabito, 2014):

Res Sapiens. It has the ability to interact with external systems and to communicate with human social network.

Res Agens. It is aware of the interaction with the environment and able to translate reasoning into action.

Res Social. It is a hypothesis for the future about objects building their own social network and interacting with each other by creating complex social systems.

The first example of smart object can be retrieved in a funny robotic bunny called *Nabaztag,* namely "rabbit" in the Armenian language. It was conceived and developed by Haladjian and Mével and made his first market entrance in 2005. Its motto was "If you can even connect rabbits, then you can connect everything". The fictitious animal was able to move the ears, know the time, provide the weather forecast, read the horoscope and also make funny jokes. More bunnies can be connected and marry each other. *Nabaztag* is considered, to all intents and purposes, the forerunner of today's virtual assistants such as Alexa and Google home.

According to DeLanda (DeLanda, 2015) the Assemblage Theory can be extended to SO-based IoT system due to the new togetherness shown in the ongoing interactions among previously unrelated objects and products (hi-tech and physical ones) which now cooperate based on a network connectivity. Accordingly, Hoffman and Novak expanded DeLanda's work and provided the following definition of such a systemic whole: "In the IoT, sensors that collect data, and actuators that transmit that data, are being increasingly incorporated in all manner of consumer objects commonly found in and around the home, worn on or in the body, and used in consumption activities involving shopping, entertainment, transportation, wellness, and the like" (Hoffman & Novak, 2018).

Despite the disruptive and innovative nature of the phenomenon, it has some drawbacks which may be traced back to the lack of clear and well-defined normative, regulatory, technological and market governance. According to a study conducted by Accenture Interactive who surveyed more than 2000 U.S. respondents, the principal barriers to adopt IoT devices are concerns with privacy, lack of awareness, lack of perceived value and price concerns (Accenture Interactive, 2014). As a matter of fact, in the modern scenario a huge amount of sensitive data or personal information is increasingly collected from consumers in order to profile them and analyse their behavioural, psychological and social habits. Therefore, it is logical that users are concerned and careful about data protection and privacy. Secondly, the lack of awareness we refer to results in both difficulty of use and unawareness of market availability. Eventually, the final barriers are interrelated as the lack of perceived value causes the price aversion. In fact, consumers consider smart objects too expensive since they do not perceive them as providing a new, beneficial and valuable contribution to their lives (Wuenderlich, et al., 2015).

1.1.1 Fields of Application of IoT and Smart Objects

Smart devices have been extensively studied in order to structure a taxonomic classification (Lackovic & Trunfio, 2014; Fontino, Lackovic, Russo, & Trunfio, 2013; Fortino, Rovella, Russo, & Savaglio, 2014). The model developed by Lackovic & Trunfio (2014) and Fontino, Lackovic, Russo, & Trunfio (2013) is currently implemented in order to index, discover and select the dynamics of smart objects. The authors divided them into four main categories depending on both functional and non-functional characteristics, that are type (namely, the type of object – smart pen, smart table, etc.), device (namely, either the hardware or software component), services (namely, the list of operations implementing the service), and location (namely, the position of the smart object). Furthermore, smart objects can be classified depending on the final recipients who can be consumers (consumer segment) or businesses (business segment) (GrowthEnabler, 2017). Based on past literature, we identified seven different applications of smart objects:

Smart Home (and Building): its objective consists of contributing to a new and impactful conception of the current way of housekeeping, since smart home can literally "keep the house" through, for example, the automation of household services. This concept is based on the idea that systems which are remotely activated through internet connection or smartphone-suitable apps can carry out housework (like, home care, home cleaning, cooking and so on) due to the received information. Nevertheless, despite this objective the study has shown that the management of household chores do not fall under the priorities of the Italian market. According to a study conducted by the School of Management of the *Politecnico di Milano*, the Italian idea of smart home experienced a considerable growth but remains anchored to an early stage and mainly focused on the energy and safety sectors. More specifically, we have the security sector first, then heating sector and household appliances management in the last place. Particularly, the energy sector can count on the so-called Smart Metering and Smart Energy which represent systems that allow remote reading and remote management of electricity, gas and water meters aimed at reducing waste (Internet Of things Observatory of the School of Management of the Politecnico di Milano, 2018).

Wearable Smart Devices (e.g. Smart Watches): as the name suggests they are all high-tech devices as smartwatches, electronic bracelets or even the shirts that control the heartbeat that we wear and connect to the network to monitor almost every aspect of our everyday life (Internet 4 Things, n.d.). Two categories that exemplify the beneficial contribution of applying these devices to human experience are health & fitness and work-performance. The former may involve smartwatches like Apple watch or Fitbit which have been intensively used by a wide range of consumers. Such a wearable technology can track fitness activities like the number of steps or meters covered or the number of calories burned by who wears it in a certain time interval (e.g. one hour, one day, one

week, etc.); also, it can be employed to monitor human activities related to the healthcare like heart rate, energy, and sleep (Rear, 2020). The latter may include intelligent tools who provide you with stored data in order to accomplish physical and managerial tasks or support or improve the performance at the workplace. Google developed a pair of goggles, namely Google Glass, about four years ago. They display the correct procedures "which would appear to float in in the worker's field of vision […] without having to find a computer terminal" (Wired, 2017).

Smart Health: it applies the use of smart technologies to the field of healthcare by storing patient data (e.g. vital signs, hours of sleep, diuresis) and creating a pool of knowledge to which the human mind would not be able to reach. As also Canhoto and Arp stated, smart objects collecting data from patients are functional to facilitate treatments in terms of efficiency and to support medical decisions during all stages of patient treatment. A clear example is the possibility to dose insulin based on the Automatic Glycaemic Readings, namely wearable devices checking the heart rate during a physical exercise and alerting in case of values out of range (Canhoto & Arp, 201).

Smart Mobility (car or vehicles in general): firstly, it implies savings in terms of costs, time and most of all pollution. Smart mobility refers for examples to automated driving systems which aims at automating the most common traditional vehicles like cars or trains so as to limit human-related harm.

Smart Retail: intelligent devices enable companies to better know their consumers in order to customize offers and provide unique purchasing experiences. Thus, smart retail is an intelligent distribution system gathering information about the purchase process and significantly speeding it up. Moreover, at the physical store smart objects can be useful to facilitate payment and spread proximity-based advertising or advertising related to the localization of the subject.

Smart Factory: during this Fourth Industrial Revolution period, IoT is one of the most characteristic elements of the so-called Industry 4.0. In fact, manufacturers tend to implement automation and data exchange as innovative technologies leading to interoperability, which means machines, devices, sensors, and people communicating with each other via the Internet of Things (IoT) or the Internet of People (IoP) (Mario, Pentek, & Otto, 2015) and building up a factory with a smart configuration. Here, the term factory is intended in its broadest sense in order to include as many manufacturers as possible. The advantages of the widespread use of smart objects are mainly production processes optimization both in terms of labour efficiency and energy savings, predictive maintenance, improved operations, inventory optimization and worker health and safety.

Smart City: in order to be titled as "smart", a city should empower sectors like public transports, mobility, healthcare, environment, waste, public safety and productivity through smart technologies that provide a better management. These revolutionary technologies are the focal point of a planning strategy that connects public infrastructures with citizens in order to satisfy their urgencies and

enhance the quality of life. Some examples are innovative waste disposal systems or smart traffic lights (which turn green when there are no vehicles in the opposite lane). According to the Juniper Research study for Intel, Singapore was the 2018 world smartest city (followed by London, New York, San Francisco and Chicago) which enables its citizens to save 125 hours per year (Salerno, 2018).

According to the McKinsey consulting company, the global value of smart devices is expected to reach $11.1 trillion by 2025 whose most consistent part is represented by Factories ($ 3.7 trillion), followed by Cities ($ 1.7 trillion), Health & Fitness ($1.6 trillion), Retail Environments ($ 1.2 trillion), Vehicles ($ 740 billion) and Home ($ 350 billion) (Manyika, et al., 2018). According to a publication from ANSA (ANSA, 2020), the Internet of things Observatory of the School of Management of the Politecnico di Milano found that the Italian market of the interconnected devices in 2019 reached a value of 6.2 billion euros (+24% than in 2018). A half of the total turnover is composed by the smart mobility segment (1.2 billion euros and 1.6 million connected vehicles, +14%) and the smart meters segment (1.7 billion meters, +19%). Nevertheless, the most significant growth was exhibited by segments like Smart Home (530 million, +40%) - driven by the boom in voice assistants, Smart Factory (350 million, +40%) and Smart City (520 million, +32%) since some municipalities with at least 15 thousand inhabitants have started a Smart City project in the last three years.

The IoT works with data which are transformed into strategical information that can facilitate human lives and satisfy managerial requirements. Thus, smart objects provide advantages for people as such and consumers. Firstly, consumers save time since IoT systems simplify routine operations offering an added value that is counterbalanced both by the price and by the risk that that companies have increasingly consumer information. Then, consumer experience can be enriched via more in-depth analyses of purchasing habits and overall customer journey leading to greater understanding and real-time solutions.

1.2 Definition and Historical Overview of Voice Assistants

Voice Assistants have been defined as "software agents which could interpret human voice and respond through a proceeded voice" (Hoy, 2018). Moreover, they can be traced back to the broader category of virtual assistants, since the latter are nowadays conceived as software-based applications capable of interpreting human natural language and completing tasks for users (Hoy, 2018). Nevertheless, it must be noticed that virtual assistants were originally conceived as practitioners who provided remote web assistance to complete the business services, as mentioned by Youngblood (2006). Virtual assistants may employ sophisticated technologies like speech recognition, Natural Language Processing (NLP) and computer vision. Before going forward with Voice Assistants, a brief focus on three operating principles of virtual assistants is required. In fact, Wu (2018) states that virtual assistants can be accessed via three main methods which are image-based, text-based or voice-based. If users upload or take a photo (e.g., barcodes), the system recognizes it and, eventually, provides a final result (e.g., search for similar products); if users type the request (command or question) in a dialog window (usually, an online chat, an instant messaging app or an e-mail), the system will process and, eventually, complete tasks (execute or answer) for users; if users pronounce the so-called wake word like "Hey Siri", "OK Google" or "Hey Google", "Alexa", and "Hey Microsoft", the system will listen to the user voice and will send it to a server in order to process it "as a control", ending up with performing tasks "such as play music or search information, depending on the returned command" (Hoy, 2018). Thus, we can assume that Voice Assistants are nothing but voice-driven virtual assistants which means that they are activated by user voice.

Voice Assistants do not represent a revolutionary invention in the field of computer technology, but rather the final outcome of a progressive and systematic evolution of pre-existing inventions that have been sophisticated and enhanced. Actually, for decades researchers have made enormous strides in the no longer just speculative field of human-computer interaction as computers are able to react to user inputs through scripted responses. In fact, since when the ELIZA project was developed in 1966, the idea of a written conversation between a machine and a real person took shapes. ELIZA was a chatbot created by Joseph Weizenbaum at the Massachusetts Institute of Technology whose aims was to simulate a likely conversation with a psychotherapist. Overall, a chatbot is a conversational software agent which interacts with users interpreting and using human natural language. Indeed, ELIZA managed to make sure users believed they were talking to a real human being instead of a computer by asking them about, for example, their emotional status (Epstein & Klinkenberg, 2001). Subsequently, an opensource chatbot, named A.L.I.C.E which means Artificial Linguistic Internet Computer Entity, implemented a less sophisticated Natural Language Processing. It demonstrated that chatting with a virtual assistant may have practical objective other than mere entertainment

(AbuShawar & Atwell, 2015). Eventually, in 1975 another psychiatrist-authored conversational software was developed in order to simulate a conversation with a paranoid patient (Colby K. M., 1975; Colby K. , 1999). Unlike ELIZA, it was characterized by a hostile and defensive tone which nonetheless made users perceive the machine as a real human (Epstein & Klinkenberg, 2001). Nevertheless, unlike Voice Assistant who, once received the user input, provide vocal output, the above-mentioned conversational virtual assistants communicate with users only via written dialogues. Obviously, in order to translate and reproduce user language the role of NLP is fundamental in both scripted and spoken conversations. Natural Language Processing (NLP) translates verbal input into interpretable and consequently executable commands. Then, computers are able to understand, analyse, and give sense to human input by considering the structure of human language and the sense the user intended to invoke (Goldberg, 2016). Moreover, both HARPY e Dragon's Naturally Speaking systems may be considered as ancestors of Voice Assistants. They implement speech recognition technology based on recognizing human voice as input to be processed (Juang & Rabiner, 2004). Around 1976, HARPY with its 1000-word vocabulary resulted from the collaboration among Carnegie Mellon University, Stanford Research Institute, IBM company and the United States Department of Défense and its DARPA agency. It could understand sentences due to its structured knowledge as a network of lexical, grammatical and syntax notions. The input speech is segmented and processed as a sequence of words (or sounds) in order to verify they satisfy the knowledge constraints of the system with the highest matching score. (Juang & Rabiner, 2004).

Nevertheless, it was not until 1997 that the first of voice recognizer, namely Dragon's Naturally Speaking, could reproduce human language with no need for pauses between each word at a rate of 100 words per minute. It is defined as "the world's leading speech recognition solution with over two decades of continuous development" that enables you to be more productive "by unlocking the power of your voice" (Nuance, n.d.).

1.3 Focus on the three most popular Voice Assistants: Siri, Google Assistant, Alexa

One of the latest improvements in the field of Internet of Things, which result in Voice Assistants as they are currently conceived, is the incorporation into objects or physical environments. Siri represents one of the first installations of virtual assistants on a smartphone with the launch of iPhone 4S on October 4, 2011 (CNN, 2011). Also, Amazon Echo, generally referred to as Alexa, and Google Home are the earliest examples of virtual assistants which activate home automation systems via voice input in order to enhance the actions that the domestic environment can complete. Overall, they are designed to assist users with managing simple tasks through the natural language voice-control interfaces. In particular, smart objects provided with such virtual assistants allowed artificial intelligence technology to actually speak with users for the first time (Smith K. T., 2018). Overall, IoT devices have become increasingly required in products from several industries such as automotive, smart home, healthcare, insurance and transports. In fact, an increasing number of brands are striving to implement smart conversation with branded artificial intelligence-powered assistant as an alternative communication channel with their consumers. Basically, it may be perceived as an actual touchpoint with the brand which may facilitate the interaction with their brand and, as a result, affect the overall brand consideration (Novak & Hoffman, 2019). The smart speaker market is characterized by a duopoly-like configuration composed by two most popular vendors, namely Amazon's Echo/Alexa and Google Home/Google Assistant. They are followed by minor brands including also Apple probably due to the late debut of HomePods. In fact, it entered the market at the beginning of 2018 and has been trying "to catch up in the game" (Statista Research Department, 2020).

Since smart assistants are expanding rapidly and widely, we needed to narrow the range of action. The study of my dissertation will consider a fictional launch of a new smart speaker. Therefore, a general overview of the main competitors that rule the market is a must. The three mainly relevant voice assistant are Siri, Alexa and Google Assistant, since they are the best known and the most widely spread among consumers.

Starting from the virtual assistant developed by one of the most successful IT companies, i.e. Apple, it has to be pointed out that **Siri**'s success is largely due to Apple's one. In fact, as we shall see later, owning an iPhone also means owning the built-in Siri service and this is why 98% of iPhone owners have used Siri although they don't do it consistently (Trappl, Your Virtual Butler. The Making-of, 2013). Siri was a project originally carried out by an independent company, i.e. SRI International's Artificial Intelligence Center. This American research centre released an app version of Siri for iOS operating system and, then, sold the rights to Apple in April 2010. As we previously mentioned, Siri made its first public appearance with the launch of the new iPhone 4s on October 4, 2011. Over the

years, it was installed on other Apple devices such as later iPhone models, iPad, iPod Touch, Mac, AirPods, Apple TV and HomePods (Woods, 2018). The Apple official website reports Siri's abilities to perform phone actions, "play music, […] set alarms, timers, and reminders, get directions, preview your calendar, […] quickly check facts, do calculations, or translate a phrase into another language, […] control your smart appliances, check their status, or even do a bunch of things at once — using just your voice" even through shortcuts (Apple Inc., 2020). It is more effective when engaging with iOS-integrated apps on Apple native devices. Nevertheless, third-party apps have been allowed to access and implement Siri technology in order to empower their own programs since 2016. In fact, Siri can be installed on Apple smartphones which enjoy its hands-free functions such as browsing Apple Store or iTunes Store just by using their voices. Nevertheless, users need to physically interact with the device in order to complete the purchase. Moreover, other non-Apple devices provide for manual data entry from the very beginning. Eventually, the Apple virtual assistant can be assigned a specific gender and a specific nationality depending on, respectively, the masculinity/femininity and accent expressed by the voice. In fact, Siri offers a wide range of voices it can take on (Woods, 2018).
In 2016, the Google company upgraded its Google Home smart speaker with the artificial intelligence-powered virtual assistant **Google Assistant**. Originally, it was available only on the Pixel and Pixel XL smartphones. In February 2017, it was extended to other Android devices, including third-party smartphones and Android Wear. Then, it was released as a standalone app on the iOS operating system in May 2017. Now, third-party developers were even allowed to enhance the functionality of the Assistant (Hoy, 2018). Moreover, it is available on different mobile and smart home devices, e.g. Google Express through which it can purchase products on your mark (Smith K. T., 2018). Overall, it can be installed on smartphones, smart displays, smart speakers and wearables devices. Compared to the earlier natural language processing-based user interface, namely Google Now, which provided responses and advices by outsourcing requests to certain web services, Google Assistant allows user to have a two-way conversation. In the official website, Google Assistant is described as supporting users in calling the phone, texting, surfing the net, scheduling events, setting alarms, adjusting device settings, and showing user account information (Google, n.d.). Moreover, Google Assistant can be awakened by selected vocal input like "Hey Google" or "OK Google", but it also supports keyboard input like Siri service. Nowadays, the Google's smart assistant can be considered as one of the most successful artificial intelligence-powered technology. In fact, as reported by the Vice President of Product, Manuel Bronstein, it is "available in more than 90 countries, the Google Assistant now helps more than 500 million people every month to get things done" (Google, 2020). As stated by Brant Ward, Product Manager of Google Assistant, its voice can be personalised. In particular, "we've gone from eight in 2017 to more than 30 today. You can already

choose between 11 English voices here in the U.S. (including John Legend), and today, we're launching a new voice in nine more languages: German, French, Dutch, Norwegian, Italian, Korean, Japanese, English in the U.K. or English in India." (Google, 2019). The different voices were developed in collaboration with DeepMind's state-of-the-art WaveNet technology "which make them sound natural, with great pitch and pacing" (Google, 2019). In order to personalize the assistant voice, users have to select among different options displayed by colour rather than gender in order to avoid labels prejudice.

Alexa is the voice assistant powering Amazon's widely popular Echo line of speakers released in November 2014 (Etherington, 2014). Moreover, its primary "home" is in Bluetooth speakers or smart device hubs. An important clarification concerns the fact that there is no built-in keyboard attached to Echo devices, since the primary and only way of interacting with Alexa is user voice hands-free. It can be considered as "the gatekeeper and entry point for Amazon's Echo line" (Woods, 2018). Overall, it can perform different activities depending on the smart device it is installed in which can be Amazon Echo, Amazon Echo Plus, Amazon Echo Look, Amazon Echo Show, Amazon Echo Spot and Amazon Dot. In particular, the basic model of Amazon Echo integrated with Alexa is a voice-controlled tool which is "capable of voice interaction; plays music, podcasts, audiobooks, and games; provides news, weather, and real-time information to questions; maintains lists, calendar, 22 alarms, and timers" and repurchase products on Amazon online market. Moreover, it may constitute a useful smart home hub. In fact, it "uses Wi-Fi, Bluetooth, or other wireless protocol standards to control home automation devices such as lights, thermostat, coffee maker, and television" (Smith K. T., 2018). Alexa's functionalities can be developed by third party vendors. Users are enabled to install apps, such as weather programs and audio features, on the device as new skills. Moreover, the service provided by Alexa can be experienced in different forms as a home automation system hub, a wearable device or an automated tool connected to car system (Hoy, 2018). Unlike Siri which is most effective on native Apple devices and services, Alexa can run on all operating systems, funnelling even Apple Mac or other branded PCs to Amazon's services. Also, it has also more authority to navigate the Amazon ecosystem than Apple smart device. In fact, it can purchase goods by voice command and, as a result, "holds the purse strings". As other smart speakers, Alexa has a wake word to activate it which is customizable and can be configured to be "Echo," "Amazon," or "Alexa," with the latter as the default. On the other hand, other Amazon devices require to push a button to turn Alexa's listening mode on. Currently, language configuration is only available in English, German, French, Italian, Spanish, and Japanese (Woods, 2018). Amazon is currently on of the two most popular vendor of smart speaker for seniority and appreciation (Statista Research Department, 2020). As Amazon's devices team claimed, over 100 million Alexa-enabled devices have been sold by

January 2019 (Al-Heeti, 2016). Moreover, one year after the release of the voice-controlled personal assistant Alexa, it was expressly appreciated by more than 500,000 consumers saying: “I love you”, making it one of the company’s most popular products (Risley, 2015).

2 What determines a Partner vs. Servant Voice Assistant?

Before analysing how the perception of a voice assistant playing a certain role can be manipulated, a preliminary introduction on anthropomorphising is due. In fact, the present study considers features that commonly belong to the human sphere as affecting the perceived agency that human mind assigns to the voice assistant. In particular, a voice assistant who shows more agency is perceived as a partner and a voice assistant who shows less agency is perceived as a servant.

Later, we will identify the determinants of how a voice assistant can be configured as a partner and, then, as a servant. Both sections will consider how a partner (vs. servant) voice assistant should be presented and behave. In everyday life, we are taught that first impressions count. Thus, since we are moving within a literary context of human quality extended to objects (i.e. anthropomorphism), it is right to look towards the way a voice assistant is presented. The presentation consists of the description of the voice assistant that is proposed to users. I tried to empirically interview my iPhone-integrated assistant to collect information about it and get inspired by how it presents itself. Nevertheless, an error occurred when it tried to process some questions, but, fortunately, it answered to others. In fact, when asked questions like “Who are you?”, “Where are you from?” and “What is your job?”, Apple Siri replied: “I am Siri… and I am here for helping you. I’m form California and have been created by Apple. I act as an audio-visual interface between your existential questions and my natural language processing mechanisms”. Thus, based on these elements actually used by smart voice assistants when asked to introduce themselves, we can assume that, basically, name (and, eventually, apposition) and mission are to be included in the presentation. Depending on the overall image provided by the combination of appearance and functions, the voice assistant assumes the configuration of partner or servant that, in this study, is explicitly disclosed.

In addition to the way it is described, the way in which it behaves is important to outline the role of the voice assistant. In fact, based on a review of the literature dealing with this topic, my dissertation takes into consideration actionable reactions of the smart object which qualify it either as a partner or a servant. In particular, the intelligent assistant is able to dialogue with human beings and, thus, the way of behaving can be intended as the way of communicating, namely the communication style. It involves all the characteristics that distinguish a certain vocal emission which conveys a certain identity (e.g., tone of voice and wording). Thus, the following paragraphs will deepen the vocal characteristics that the assistant should display in order to convey the image of a partner or a servant. More specifically, the behaviour condition will show the demeanour of the assistant while making a shopping order on behalf of the user.

2.1 Anthropomorphising inanimate entities

Anthropomorphism is the natural human tendency to imbue inanimate objects with a set of typically human traits which, taken as a whole, outline a real personality (Aggarwal & McGill, 2007). Scholars have extensively demonstrated the ease with which people provide anthropomorphic descriptions of unhuman agents which comprise anything that "acts with apparent independence, including animals, natural forces, religious deities, and mechanical or electronic devices" (Epley, Waytz, & Cacioppo, 2007). The inherent tendency to anthropomorphise objects does not only describe how they act or subsist (e.g., the dog is affectionate) but represents also a mental or physical status which recalls humanlike descriptors (e.g., the dog loves me). Anthropomorphism represents a process of inductive inference about nonhuman agents. Thus, it should follow the same basic cognitive operations of any other inductive inferences. The first mental operation implemented is the acquisition of knowledge. Then, the activation or elicitation of stored or any other additional knowledge is applied to a given target by correcting or adjusting highly accessible knowledge in order to integrate alternative knowledge structures that are coactivated at the time of judgment. Nevertheless, in his essay on working principles of mental systems, Gilbert (1991) stated that the final judgement is based on the most readily accessible information (Gilbert & Malone, 1995; Higgins, 1996).

Moreover, a study conducted by Epley, Waytz, & Cacioppo (2007) demonstrated that three psychological determinants influence people's tendency to anthropomorphize non-human agents: the accessibility and applicability of anthropocentric knowledge (elicited agent knowledge); the motivation to explain and understand the behaviour of other agents due to the anxiety associated with uncertainty (effectance motivation); and, the desire for social contact and affiliation (sociality motivation).

Smart products constitute a subcategory of conventional objects, since they are able to actually interact with humans through short conversations, queries, commands, questions and more (Schweitzer, Belk, Jordan, & Ortner, 2019). An early exemplification of the anthropomorphising practice in the computer science can be retrieved in the so-called ELIZA effect. In a broader sense, it can be defined as the tendency to consider computer behaviours as similar to human ones. More specifically, it is "the tendency for people to treat programs that respond to them as if they had more intelligence than they really do" (Trappl, Petta, & Payr, 2002). According to Trappl, Petta, & Payr, the ELIZA effect demonstrates the effectiveness of applicating a social perspective to the field of artificial intelligence rather than the mere programming engineering. For further clarification, Douglas Hofstadter exemplifies the ELIZA effect with the misleading belief of a casual observer that an automated teller machine displaying the words "THANK YOU" at the end of a transaction is thought as actually being thankful although it is only "printing a pre-programmed string of symbols"

(Hofstadter, 1996). As we mentioned above, ELIZA was a chatterbot developed in 1966 by MIT computer scientist Joseph Weizenbaum. The script reproduced on ELIZA's screen was a parody of a Rogerian psychotherapist, largely by rephrasing the patient's replies as questions. Weizenbaum discovered that ELIZA's DOCTOR script elicited unexpected emotional responses from users who, as a result, treated it as a real human being caring about their problems even when they were aware it was a simple computer program (Billings, 2007).

Also, studies conducted more recently investigated the extent to which users perceive computers as individuals due to human characteristics recognized in appearance or behaviours. The immediate consequence is the way in which they approach these inanimate entities by treating them as human beings. The theoretical foundation is the Computer Are Social Actors (CASA) paradigm providing that computer can be considered real social entities as they can interact with users. Thus, we can state that human-computer interactions follow the same social rules of human-human interactions and, as a result, can be defined "fundamentally social and interpersonal" (Nass & Moon, 2000; Nass, Moon, Fogg, Reeves, & Dryer, 1995). According to this paradigm, such a social reaction of humans towards unhuman entities can be explained by a two-step mental processing: firstly, humans consciously process some cues from an inanimate entity; then, by referencing to these cues, they unconsciously assign a certain human personality to this entity which becomes socially relevant. Individuals ignore the unnatural facet of this cognitive bias as they are unaware of it (Nass & Moon, 2000). Additionally, when a computer shows cues that generates a social reaction into individuals (i.e. the perception of computers as units with whom interact), these cues are even called social cues (Araujo , 2018).

We argue that CASA needs to be expanded because people have changed, becoming more familiar with technologies; technologies have changed, becoming more sophisticated than a script automatically displayed on a computer screen; eventually, the way people interact with technologies has changed. Anthropomorphising seems to be a significant improvement in social robotics research in the near future, intuitively providing sophisticated artificial intelligence machines with innovative physical and social features (Gambino, Fox, & Ratan, 2020). In fact, my dissertation is going to deepen the anthropomorphic practice applied to the narrowed field of voice-controlled smart assistants which, in turn, reply with humanoid voice.

2.1.1 Anthropomorphising Voice Assistants (Partner vs. Servant)

As we noted earlier, voice assistants can actually engage in a conversation with users, persuading them that they are relating to and cooperating with one of their own. Generally, when individuals communicate with each other, it may involve different communication styles such as friendly, relaxed, contentious, attentive, precise, animated, dramatic, open, or dominant based on personal

factors (e.g. gender or social status) or contextual factors (e.g. the current situation or time) (Norton R. , 1978). Accordingly, the style in which a smart speaker interacts with users has the power to convey a certain overall impression to the listener that would have been completely different with another style (Carli, 1989; Norton R. , 1983).

In a consumer perspective, the interaction style has a considerable and effective influencing power, since it has been proved to exert a greater impact on the evaluation of a service received than other provisions of the service (Wu, Chen, & Dou, 2016). Accordingly, the social interaction model suggests that if consumers are not very familiar with service issues, it will be mostly evaluated based on the interaction style implemented in the service-consumer interaction. Moreover, style of conversating has significant managerial implications by affecting factors like customers' purchase probability (Dion & Notarantonio, 1992; Williams & Spiro, 1985), trust and satisfaction (Celli, Ghosh, Alam, & Riccardi, 2016; Webster & Sundaram, 2009). But more importantly, the style implemented in a service interaction may affect the overall perception of the service quality. As a proof, Notarantonio & Cohen reported that "an affiliative, friendly, and accommodative style has been shown to be superior to a dominant, precise, and unaccommodating style" (Notarantonio & Cohen, 1990).

By focusing on the fast-growing context of IoT and smart objects, several studies have been conducted on smart social interactions employing different styles and affecting how users overall perceive the smart assistant they talk to. Particularly, Wu, Chen, & Dou (2016) examined two smart interaction styles that are frequently implemented by companies in order to convey a certain brand image reflecting the intrinsic brand identity: friend-like and engineer-like style. They can boost consumers' perceptions of brand warmth and/or brand competence and, as a consequence, consumer emotional attachment to the brand (Wu, Chen, & Dou, 2016). Even before affecting consumer perception of these brand components, human mind elaborates these two completely different ways of talking to consumers as two distinct and well-defined roles that smart assistants may play. Once consumers confer a certain role to an intelligent assistant, it assumes a real identity and it is able to interact, namely an ontologically relevant entity similar to human beings. Thus, smart objects can be considered as agents worthy of respect and concerns because of the assumption of a typically human role. Moreover, the communication style can arouse in users an emotional reaction to brand by enhancing the perceived brand warmth and attachment as well as to the smart object they interact with, just as a human-to-human interaction typically does. In fact, other than perceive objects as ontologically relevant entities, humans "feel" them as such to the point where they establish real relationships with them. In fact, some individuals become so emotionally connected and attached that anthropomorphised products are seen as friends and kept for longer, since replacing them would mean

betraying a friend and being unfaithful (Chandler & Schwarz, 2010). Experiencing a sort of affection with smart objects and, in particular, with voice assistants is even easier than with ordinary products due to their inner human-like intelligent components, i.e. voice, which bring them closer and make them more similar to human identities. This is how the anthropomorphising of vocal assistants occurs and works.

Former studies concerned voice assistants based on the typically human-like role they played in the interactional relationship with the user (Sundar, Jung, Waddell, & Kim, 2017; Schweitzer, Belk, Jordan, & Ortner, 2019; Kim & Kramer, 2015; Wu, Chen, & Dou, 2016). Particularly, depending on the answers of three groups of respondents about the perceived hierarchical role of the smart object, Schweitzer, Belk, Jordan, & Ortner identified three types of relationship which users have engaged in with a smart object. The first one is a servant–master relationship where the voice controlled smart assistant is a servant; the second one is a servant–master relationship where the voice controlled smart assistant is a master; the last one is a partnership where members treat each other as a peer (Schweitzer, Belk, Jordan, & Ortner, 2019). Particularly, when respondents referred to the voice-controlled smart assistant, metaphors involving familiar human roles (namely parent/child, teacher/student and employer/employee) naturally emerged and prompted a relationship improvement. Since Schweitzer, Belk, Jordan, & Ortner reported that the relationship where voice assistant is perceived as a master is "fraught and unlikely to continue in the future", it would not bear interest to study it. In fact, the present dissertation will deepen the anthropomorphised role of either servant or partner that a voice assistant may play. This distinction is also consistent with the study of Gruenfeld, Inesi, Magee, & Galinsky who theorized that every relationship between two parts involves different hierarchical degrees. Therefore, either "one entity is an equal partner to the other or one entity can dominate the other" (Gruenfeld, Inesi, Magee, & Galinsky, 2008). Overall, the voice assistant humanized into a servant is defined as an agent "working for consumers" that represent masters and creating benefits to please them; instead, the voice assistant humanized into a partner is defined as an agent "working with consumers" that represent partners and coproducing the benefits with consumers as equals (Aggarwal & McGill, 2012).

In conclusion, Schweitzer, Belk, Jordan, & Ortner (Schweitzer, Belk, Jordan, & Ortner, 2019) stated that perceiving the assistant as a partner makes the user-assistant relationship full of compatibility and sympathy, to the point that users are encouraged to talk slowly and repeat patiently in order to "make Siri learn and grow", "connect to her" and "make her understand me better". Thus, we can assume that such an anthropomorphising of the machine results in a positive effect on users, as if they liked the machine. Similarly, when actions undertaken by the assistant are perceived as servile, users are pleased by such a reverence to the point that it can be considered as satisfying. Furthermore,

Aggarwal and McGill demonstrated that an anthropomorphised brand can prime consumers' behaviours subsequently to the exposition to the brand. In the same way, we want to test whether human behaviour can be triggered by the interaction with an anthropomorphised smart assistant powered by voice input. In particular, the present dissertation will investigate whether user satisfaction could be primed by the anthropomorphic role of servant vs. partner played by smart voice assistant based on the way it is presented to and the way it communicates with users.

2.2 Voice Assistant as a Partner

2.2.1 Presentation

This chapter of my dissertation starts from the explanation of how a partner voice assistant should be presented. Based on some reviews of the Amazon Alexa smart speaker on the Amazon website, I gathered some features that users ascribe to the assistant, depending on whether it is perceived as either a partner or a servant. When users consider the assistant an equal partner worthy of human soul and conscience, they even address it using the feminine personal pronoun, "She". Moreover, sometimes she is perceived as a part of the family or, at least, as a living entity who is a fundamental part of the house - rather than a mere home automation system. Some reviewers even stated that, although she was able to listen to conversations at any time, her presence never made them feel uncomfortable. Moreover, it has been defined as a smart rather than a technologically sophisticated or advanced innovation, which implies an ontological relevance. She aims at helping users in order to make them feel happier by doing something for them or even with them.

There are many examples of products, companies and brands which call themselves "friend" or, even, "partner" of their consumers. For example, the State Farm brand refers to itself as "a good neighbour" for consumers (Kim & Kramer, 2015). Thus, even a smart assistant can be referred to with a friendly apposition. Assigning a human role to an inanimate entity is a common practice for marketing and psycho-social purposes. In fact, as we mentioned above, human beings tend to mentally represent objects with typically human features in order to facilitate the mental processing of objects' information. In a study conducted by Schweitzer, Belk, Jordan, & Ortner, respondents were asked to interact with Apple Siri voice assistant and, in particular, to ask her to perform a task or search for information (2019). Then, they were asked some impressions and opinions about the assistant. A very interesting question was about the way they figured on their mind the assistant. They irrationally provide a vivid and quite detailed portrait of it as "tall, slim and black-haired" or, overall, as an attractive, likeable and helpful woman (Schweitzer, Belk, Jordan, & Ortner, 2019).

Last but not least, the choice of words and the formulation of the speech are fundamental components for a presentation. Wu, Chen, & Dou depicted how a branded smart assistant interacting with users through a professional and cold wording were perceived as a competent but detached engineer (2016); on the other hand a more friendly and caring wording would seem to reflect the role of a partner. For example, a language featuring terms like "we" rather than "you and I" can elicit more closeness and emotional attachment, since it better expresses the concept of a more connected togetherness (Sela, Wheeler, & Sarial-Abi, 2012). Moreover, as highlighted by reviews on Amazon website, some users

disclosed their actual feelings of love and miss towards the assistant that should be reproduced in the presentation.

Eventually, in order to convey the image of an intelligent and autonomous entity, it does not need any input by the user to start working. In fact, as we will see later (see paragraph 2.3.2., *Voice Assistant as a Servant - Presentation*), users who are given control over the assistant perceive they are actually imposing their agency on it (2019). Instead, a partner assistant is an independent agent who does not require any contribution by humans.

2.2.2 Behaviour

In this section, we are going to profile a partner voice assistant. Nowadays, conveying the friendly, well-intentioned, and trustworthy elements of brands through a friend-like style of communication has become a widespread practice among companies (Wu, Chen, & Dou, 2016). Additionally, it has been proved that a friend-like style rather than an engineer-like one can be compatible with both friendly brands (e.g. Disney) and rugged brands (e.g. IBM) "without diluting or damaging a company's well-established brand positioning". Wu, Chen, & Dou demonstrated that consumers will perceive a more positive brand warmth, brand competence and brand attachment when they interact with a branded smart voice assuming a friendly tone rather than a professional one (2016). Likewise, we assume that such a positive effect of the friend-like interaction style can be extended also to fields of applications not related to the brand. Moreover, this dissertation is going to outline a partner-like style which is similar to a friend-like one, considering that partners act and interact among each other, since they belong to the same hierarchical level. In fact, acting as a partner to users may arouse in them a better perception of the smart assistant and elicit an overall sense of satisfaction due to the affiliative and accommodative facet. Nevertheless, it should be pointed out that even a voice assistant powered by artificial intelligence has only one tool at its disposal to interact with users, namely the voice. Consequently, behaviours that we are going to investigate purely refer to the vocal acts performed during a typical conversation between user and smart device. Particularly, it is quite obvious that a conversation implies choices about the content and the arrangement of the speech as actionable behaviours.

First of all, it is important to understand how the assistant can configure itself as a partner through voice elements. In fact, it has already been demonstrated that voice manipulation can bear fruit. The modulation of intonation concern rhythm, and accentuation. Moreover, a change in the form of a word (typically the ending) is commonly used to express a grammatical function or attribute. Thus, it may be helpful to express also a certain personality. Dominance can be found in a serious behaviour

that has been proved to be represented by a lower pitch of voice may represent a (Niculescu, Van Dijk, Nijholt, Li, & See, 2013).

In addition to the cadence and sound characteristics of the voice, the behaviour of a partner voice assistant features a supportive component which reflects the emotional connection established with the user. Since customers became the value co-creators rather than the mere receivers of that value (Haeckel, 1999; Vargo & Lusch, 2004), smart user interfaces should speak human language, that is characterized by emotional component. For example, in the popular movie "Big Hero 6", the fictional smart robot, Baymax, gives the boy a hug when he feels sad and dances with him when he feels happy. Basically, smart assistants should express their emotions as individuals commonly do in order to be humanized. Moreover, human social interactions employ or, at least, should employ a polite language. It can be rendered also in conversations with smart assistants by a wording featuring, for example, "Please" and "Thank you". In this way, assistants can be perceived as one of the same species and, more specifically, as an equal partner. As evidence of the importance of the emotional component within smart interactions, we relate to a research by Schweitzer, Belk, Jordan, & Ortner (2019). After cumulative interactions, respondents learned that the voice-controlled smart assistant lacked emotional intelligence and damaged the perception of a truly intimate relationship. Therefore, they have even planned to abandon its use. Overall, it could be compared to dating: at first, humans appreciate the non-human sassy personality, but the relationship begins to decline when there does not seem to be "enough brain behind the sexy exterior". Therefore, a further caution about which programmers should be concerned is credibility and plausibility. This means that voice assistants should improve the emotional intelligence element by avoiding repetitive and unsuitable answers in order to increase the intimate connection with the user. Furthermore, a supportive character can be shown by helping in performing a task. This intention to help users makes a voice assistant behaviour even more similar to that of a partner. As a result, it is actually perceived as "an efficient, intelligent, conscientious and skilful as well as caring friend" as demonstrated by Deci & Ryan (2000) and Proksch, Orth, & Cornwell (2015). In fact, the voice assistant is perceived as a partner when it shows elements of both competence and sensitiveness that may help users (Wu, Chen, & Dou, 2016).

Thus, a partner-like style employs a customer-centric strategy based on a "sense-and-respond" practice. In fact, a peer-to-peer relationship is characterized by the mutual intimate knowledge of personality and preferences of the relationship members. Similarly, a partner assistant should provide the user with tailor-made proposals, namely appropriately and specifically designated for the user by creating a personalized experience. Such a partnership based on deep customer insights would make the user react to the voice assistant with more favourable emotions (Fournier, 1998; Kervyn, Fiske, & Malone, 2012).

Overall, Hoffman and Novak (2018) theorised that an anthropomorphised personal assistant is perceived as a partner when it shows gregariousness and extraversion. By considering a smart conversation, an extravert voice assistant engages in frequent verbal interactions with users, showing also proneness to answer. Furthermore, past research showed that machines can be considered as having their own identity. Their amusing personality can even animate the interaction. In fact, users appreciated a voice assistant "making jokes, reacting very coldly when offended and in general being good at a quick comeback" (Schweitzer, Belk, Jordan, & Ortner, 2019). In this way, the assistant demonstrates a bold and reactive personality. Moreover, previous studies showed that the way of approaching to users influences the perception of the voice assistant as "cheeky" and "self-confident" (Schweitzer, Belk, Jordan, & Ortner, 2019). Thus, an enterprising conversation can make the assistant even more similar to a human being and, in particular, a partner. In this perspective, an influent power is exerted by the phrasing of proposals, since a confident assistant would come up with assertive advices rather than hesitant suggestions. In fact, a laboratory experiment tested the perception of computer personality as either dominant or servant depending on, among other things, the text displayed by the computer. The former employed more commands and assertion and, on the other hand, the latter employed more questions and suggestion (Nass, Moon, Fogg, Reeves, & Dryer, 1995). Furthermore, an enterprising assistant may go beyond what is asked, deepen the request and provide a solution out of the box. For example, it should provide further insights into a definition that has been asked, find offers which can replace the research of the user, if it is not available and, also, provide additional offers to expand user's range of choice (Schweitzer, Belk, Jordan, & Ortner, 2019).

2.3 Voice Assistant as a Servant

2.3.1 Presentation

Physical appearance reminding human features may be determinant for the way in which smart assistant is presented. In fact, past research exhibited how inanimate objects such as car grilles and cell phones are intentionally designed to resemble human faces (Aggarwal and McGill 2007; Landwehr et al. 2011). In particular, considering a study conducted by Schweitzer, Belk, Jordan, & Ortner, some respondents perceived a voice assistant like Apple Siri as a servant and instinctively humanized it in a secretary "dressed up with blouse and jeans" (2019). Therefore, it would seem that exhibiting typical features of a character who usually serves others contributes to represent a servant assistant. Nevertheless, the most appropriate representation of the "servant voice assistant", as intended by users who have actually interacted with it, consists of a nice, friendly, helpful, reliable and ready-to-please character. In fact, respondents admitted feeling a mild attachment towards the voice assistant to the point that it is assimilated to a faithful companion, i.e. a pet, rather than a slave. (Schweitzer, Belk, Jordan, & Ortner, 2019).

Since we previously demonstrated that a partner focuses more on communicating friendly intentions, a servant should employ engineer-like style and, especially, wording just like the one Wu, Chen & Dou (2016) dealt with. They experimented that a voice assistant using a professional and detached wording lowers the impact on the brand attachment. In the same way, a servant voice assistant should not emphasize the emotional connection with users.

When presented, the image of servant can be conveyed by a wording like "this assistant is going to work for you" or "[…] to please you" rather than "[…] to work with you" or "[…] to support you in your work". Accordingly, the tasks it can accomplish are considered as a service it provides to the user-master. In fact, Kim & Kramer carried out a research on how brands which describe themselves as working for a consumer-master trigger in materialists a desire to dominate, and, thus, brands are perceived as servants. Showing brands as human chracters by conveying a servile attitude can be examplified by Burger King which speaks of itself as a servant performing "your way, right away" (Kim & Kramer, 2015). In the same way, a voice assistant could benefit from such a similar wording in the presentation in order to be perceived as a subordinate which serves users.

In addition to the assistant's position hierarchically subordinate to the user, the words chosen to convey its functions are important. In fact, Aggarwal & McGill implemented a concept of servant that refers to an outsourced provider of benefits rather than precisely a slave (2007), and we will reproduce a similar one. Thus, the introductory presentation of the voice assistant should convey the message that the smart object is able to please users by adding value to interaction that otherwise

would have been less satisfying. Also, as we have seen for the partner voice assistant, the task for which it is best suited should be indicated in the presentation of the servant voice assistant.
Eventually, since users who see the assistant as a servant perceive that they literally impose agency on the object, they want to actually feel they are needed to activate the smart object and make it work. In fact, the study by Schweitzer, Belk, Jordan, & Ortner came to the conclusion that people should be supplied a control and feedback mechanism like "the 'close door' buttons on elevators that allow consumers to feel a sense of control" (2019).

2.3.2 Behaviour

As for the partner role, the variable defined as behaviour of voice assistant refers to all the elements constituting a conversation between user and servant assistant. As a starting point, the voice characterization should be taken into consideration. A study conducted by Sundar, Jung, Waddell, & Kim tested how senior citizens perceive robot assistants as playful or serious depending on the tone of voice (2017). In particular, a robot featuring a female voice speaking with limited inflections in pitch and tone would appear to show a serious demeanour (Sundar, Jung, Waddell, & Kim, 2017). Additionally, a higher pitch of voice has been proved to denote a less serious behaviour (Niculescu, Van Dijk, Nijholt, Li, & See, 2013). As we assumed that seriousness suits for a dominant assistant, a high pitch of voice can be assigned to a submissive assistant.
Furthermore, also the emotional component seems to be crucial. In particular, words exchanged within the conversation should be detached through unemotional wording and inexpressive tone of voice. In fact, when the smart assistant is perceived as submissive, individuals are aware that it is "just a voice behind which a person can be imagined" (Schweitzer, Belk, Jordan, & Ortner, 2019). Thus, they do not establish a deep and intimate relationship with the assistant due to this lack of will, unlike they used to do with a partner assistant with which they show a strong emotional connection.
According to this perspective of submissiveness, the device is portrayed as a very shy and civilised person who basically "stays in the background, and obeys if talked to" (Schweitzer, Belk, Jordan, & Ortner, 2019). Consequently, such a smart assistant can only afford a standard and average performance, since it has not the enterprising personality to provide a tailor-made and customized user experience. Moreover, the user is only provided the information he or she expressly asked without any further content or proposal. Accordingly, Schweitzer, Belk, Jordan, & Ortner noticed that a servant voice assistant is perceived as actually saying "yes and amen to everything" (2019). Thus, it would come out with linear sentences which are strictly related to the issues that it deals with and that are never out-of-the-box. This means that it would never contradict the "master" who provides power and imposes agency, since it appears in a "lower hierarchical position". This

perception is in accordance with the digital-self theory (Belk, 2013) where the digital elements are extensions of individuals. Rather than an external source of assistance, a voice controlled smart assistant is considered as an extended part of users that enhances and empowers their capabilities by providing new means they previously lacked.

In general, the verbal exchange between servant voice assistant and user does not appear as a peer-to-peer conversation, since it does not present high levels of interactivity. Thus, verbal exchanges are infrequent. Moreover, the smart assistant shows hesitation and, even, tardiness to answer which makes conversation less smooth and, as a result, unrealistic. We earlier defined assistant's personality as reverent and respectful and this fits its subordinating position. Hence, it is not surprising that, when it interacts with users, it employs only cold answers that leave no room for originality, sympathy and the ability to respond in kind. Sometimes servant voice assistants are even perceived as senseless, since they provide pointless or even wrong replies. This can make it difficult to treat them as intelligent creatures. Nevertheless, this is probably due to a need for more sophisticated technological improvements rather than to what the voice assistant can do or seem (Schweitzer, Belk, Jordan, & Ortner, 2019). Accordingly, a servant assistant only responds to the inputs provided by the user (Nass, Moon, Fogg, Reeves, & Dryer, 1995). Thus, a voice-controlled smart object always goes second in the interaction and never takes the initiative to start the conversation. Overall, we can assume that a servant voice assistant merely reacts to users' orders and questions and never speaks spontaneously. In order to do so, it should employ a phrasing which shows its reverent and respectful personality like a subordinate. Indeed, computers have been experimented to show a typical human submissive or dominant personality depending on the level of confidence displayed (Nass, Moon, Fogg, Reeves, & Dryer, 1995). More specifically, submissiveness can be represented by a weaker language expressed in the form of questions and suggestions that display a low confidence level.

2.4 Fit Between Role and Behaviour

In the interaction with the smart object users should perceive the behaviour of the vocal assistant as fitting the role that it claimed to assume in the presentation, since such a fit will trigger a positive effect on the user satisfaction resulting from the interaction. It all depends on the mental representation that the user forms of the smart object resulting from the role played and, particularly, the whole of related information stored in his or her mind, on the smart object's actual behaviours and on the degree of consistency between the former and the latter.

Since the second half of the last century, several in-depth studies have been carried out in order to examine the importance of consistency in human cognitive processing. Festinger (1962) defines consistency as the degree to which a certain individual's cognition implies another one based on how they are mentally represented in the individual's mind. Kruglanski, et al. demonstrated that people "possess a general need for cognitive consistency" (Kruglanski, et al., 2018). In fact, according to the Cognitive Consistency paradigm new information that are inconsistent with prior beliefs has a negative impact on people's affective and cognitive reaction which may result in distress.

Furthermore, consistency also plays a crucial role for interpersonal relationships as the Cognitive Consistency paradigm can be extended to expectations people made about others. Particularly, within human interactions each individual expects other people to behave consistently with what has already been established consensually and, often, tacitly. Thus, the assumption of the paradigm that human mind prefers to avoid the processing of discordant information may be applied to human behaviour violating the human expectation mentioned above. In fact, this theoretical model regulates both civic and professional relations which are based on compliance with institutions and what is called psychological contract, respectively. The latter provides for mutual guarantees that employee and employer exchange at the beginning of their working relationship. Such a contract can be considered broken when one of the two parties disregards its part of the agreement and the remaining part perceives that "their firm, or its agents, have failed to deliver on" the promised trade value (Van den Heuvel & Schalk, 2009). Now, the concept of psychological contract generally describes many relationships such as the consumer-brand relationship. More specifically, consumers trust the brand and the value it promised to supply and distribute to community (Chih, Chiu, Lan, & Fang, 2017). On the other hand, the brand can count on the committed duty or responsibility consumers show towards the brand that result in both concrete (e.g., willingness to purchase) and abstract (e.g., brand ambassador) brand loyalty (Ke-fa, 2009).

As additional proof, a study conducted by Nass et al. (1995) reported the practice of companies to exploit the style of interaction implemented by computers behaviour to convey the brand personality. More specifically, companies oriented towards customer relationships tend to assume a friendly tone

of voice in order to position themselves as "their customers (best) friends" by matching a sincere, honest, and genuine personality, by conveying caring intentions and by strengthening emotional attachment towards the brand. In contrast, companies oriented towards technology and innovation prefer to position themselves in the consumer's mind as a professional and experienced problem solver by matching a rugged, efficient, and sophisticated personality, by conveying indifferent intentions and by lowering emotional attachment towards the brand (Nass, Moon, Fogg, Reeves, & Dryer, 1995). Overall, the brand distinctive features should be reflected in the communication process so that consumers perceive a high level of congruence between the brand personality and the actual behaviour of the smart object. Additionally, such a congruence "facilitates customers' information processing […] and can reinforce the brand meaning among customers" (Sirianni, Bitner, Brown, & Mandel, 2013).

Therefore, there seems to be sufficient evidence to support the claim that users need to be shown a certain degree of consistency between how the voice assistant acts and the role it was presented to play. As a result, an interactive behaviour congruent with the role a certain voice assistant plays could positively impact the sense of satisfaction due to the conversation with the smart object. As exhibited by Aggarwal and McGill (2007), the overall satisfaction and consequent evaluation that consumers make of an anthropomorphized product can be positively influenced by the perception of the actual features of the product as matching the mental schema whom consumers built up of the human character the product should resemble. Likewise, consumers may be more satisfied by a product whose actual behaviour is perceived as matching the mental schema consumers own of its role. This underlying reasoning can be extended to the evaluation of any product or service such as an interactive communication with a smart voice assistant.

Eventually, environmental fit with human personality is useful for explaining why the voice assistant behaviour should match the role intended to be conveyed. In fact, Pervin (1968) experimented that, for each individual, environmental components (i.e., stimuli external to an organism) best-fitting his or her personality elicit a positive effect on both individual performance and satisfaction. Particularly, the study proved that non-interpersonal environments like perceived occupational role requirements should match individual personality or self-imaged personality in order to affect satisfaction (Brophy, 1959). These findings support the assumption that individuals are more satisfied when they believe they fit in and feel comfortable in a given condition. Such fit and comfort come from the match between an inner reality of the individual (i.e., personality) and an exterior one (i.e., non-interpersonal environment), but, based on the same reasoning, we can assume they could also derive from an inner reality (i.e., role) and an exteriorization of an inner reality (i.e., behaviour). Hence, human personality and behaviour can be compared to smart assistant presentation and behaviour, respectively. Thus, the

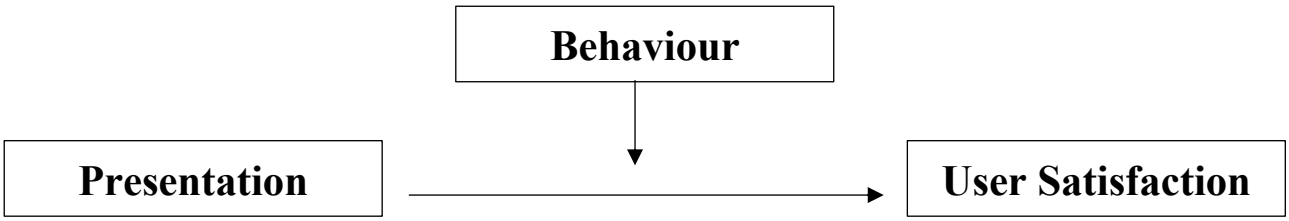
presentation of a smart voice assistant can reasonably affect the post-interaction user satisfaction due to the match with the behaviour it engaged in played which is comparable to the personality of an individual.

2.5 Hypothesis development

Basically, this dissertation is going to investigate the role that a voice assistant is perceived to play. In particular, users may perceive the voice assistant as a partner rather than a servant depending on how it is presented. Additionally, they may perceive it as a dominant entity rather than submissive depending on how it interacts. Both presentation and behavioural elements can be manipulated in order to convey a certain image of the voice assistant. Moreover, users are assumed to be more satisfied when the behaviour exhibited during the smart conversation is consistent with the presentation aimed at introducing the voice assistant to users. In particular, the core hypothesis of this study will be as follows:

H1) Partner smart speaker with a dominant behaviour will be evaluated more positively than partner smart speaker with a dominant behaviour, while servant smart speaker with a submissive behaviour will be evaluated more positively than servant smart speaker with a dominant behaviour.

Thus, it will be run a between-subjects two-way Analysis of CoVariances (ANCOVA), since we have two categorical independent variables impacting on a continuous dependent variable (DV). Whose effect is controlled for some covariates. Presentation represents the focal variable (IV1) and consists of two categories or levels, namely Partner and Servant. Similarly, Behaviour represents the moderator (M) and consists of two categories or levels, namely Dominant and Submissive. Independent variables have been separately manipulated in the same scenario and different levels have been combined. Thus, it has been developed an experimental 2x2 design resulting in a total of four scenarios. According to our hypothesis, users will show a higher level of satisfaction, when exposed to two of these scenarios, i.e. Partner Presentation – Dominant Behaviour and Servant Presentation – Submissive Behaviour. The following is the conceptual model for hypothesis testing:



IV = Presentation

M = Behaviour

DV = User Satisfaction

3 Quantitative Analysis

3.1 Methodology

3.1.1 Research type

My dissertation aims at testing cause-effect relationships between manipulated presentation of voice assistant and satisfaction of the user related to the interaction with the assistant, considering the moderating influence of its actual behaviour. Since we want to quantify this relation by attributing numerical data to these constructs, this is going to be a quantitative causal research and the method will consist of experiments.

3.1.2 Data collection and sampling method

In order to measure the above-mentioned constructs and establish the cause-effect relations, it was conducted an experimental survey which was administrated online and was accessible by mobile phones or any other digital device. Due to the nature of the research, the only criterion required was the sample to be large enough to generalise the results to the entire population. Also, handing out the questionnaire through the Internet has been demonstrated to be more convenient. Data will be collected via Qualtrics Software.

Other than demographic questions which are commonly implemented, respondents have been shown realistic scenarios of both a description of a new voice assistant that has just been launched on the market and, also, a simulated conversation with the assistant itself, in order to assess user reaction.

This research is going to adopt a non-probability sampling method due to convenience reasons. In fact, the online survey has been spread through a personal network of acquaintances and over social groups on the Internet. Target population will be people between 17 and 70 years old. We could have selected a targeted sample of respondents based on previous experience with smart voice assistant and voice shopping. For convenience, we chose a random sample and, then, we refined data through the insertion of covariates that controlled for the propensity to purchase a product online, the frequency of voice shopping, the familiarity with smart speaker and the concern in privacy protection.

3.2 Survey

Each respondent was randomly assigned to just one out of the four possible scenarios: 1) partner presentation– dominant behaviour, 2) partner presentation – submissive behaviour, 3) servant presentation – dominant behaviour, 4) servant presentation – submissive behaviour. As scenarios will differ upon two manipulation variables, i.e. presentation and behaviour, it is going to be a 2 by 2 between-subjects design. Two dummy variables have been created, coded as 0 when respondent has not been assigned to a certain condition and coded as 1 when he or she has been assigned to that condition. In particular, the presentation variable takes either the value 1, when voice assistant is introduced through descriptors which are suitable for a partner, or 0, when voice assistant is introduced through descriptors which are suitable for a servant. The behaviour variable takes either the value 1, when responses employed by the voice assistant resemble a dominant behaviour, or 0, when responses employed by the voice assistant resemble a submissive behaviour. The new dichotomous variables are labelled "*PRES*" and "*BEH*", respectively. Overall, the experimental design would be as follows:

		PRESENTATION	
		Partner = 1	**Servant = 0**
BEHAVIOUR	**Dominant = 1**	Partner Presentation with Dominant Behaviour	Servant Presentation with Dominant Behaviour
	Submissive = 0	Partner Presentation with Submissive Behaviour	Servant Presentation with Submissive Behaviour

3.2.1 Manipulation and Construction of the Scenarios

The manipulation of presentation will consist in modifying descriptors which have been implemented in two written descriptions of a new voice assistant, named Trinity, that has just been launched on the market. A voice assistant can be presented as a partner or a servant depending on many features (Aggarwal & McGill, 2007; Kim & Kramer, 2015; Schweitzer, Belk, Jordan, & Ortner, 2019; Sela, Wheeler, & Sarial-Abi, 2012; Wu, Chen, & Dou, 2016). In line with their findings, two presentations were designed for the voice assistant in order to manipulate its perceived role, by briefly stating who the assistant is and what is its function. The construction of the behaviour scenarios started from a careful research on what product is one of the most likely to be purchased online. Then, we exposed participants to a fictional situation where they had to imagine they want to buy such a product and they ask Trinity. According to Chiang & Dholakia (2003), people prefer to purchase online items

based on convenience, price and type. In particular, we selected a search good, namely a good whose attributes can be assessed before purchase, that is more convenient in terms of time and cost to buy online at an affordable price. Thus, participants were supposed to request Trinity to buy a USB charger cable for smartphones. In fact, it ranks among the Amazon Bestsellers, where the products most purchased on the world's largest online marketplace are listed.

Then, behavioural factors were manipulated based on two papers written by Sundar, Jung, Waddell, & Kim (2017) and Schweitzer, Belk, Jordan, & Ortner (2019). The former discovered that pitch of voice can convey a certain mental representation of a robot demeanour; the latter found that consumers value intelligent devices (e.g., voice assistants) more positively when they provide tailor-made user experiences. In fact, consumers have been demonstrated to "want more than talking to a device" (Schweitzer, Belk, Jordan, & Ortner, 2019). Thus, the manipulation will consist of both a more or less high pitch of voice and a personalized or standard response to the user input. Moreover, Schweitzer, Belk, Jordan, & Ortner demonstrated that people tend to perceive voice assistants as dominant when they employ in assertive advices, whereas they perceive them as submissive when they employ in hesitant questions (2019). In both the dominant and submissive condition, respondents played two audio files which have been created via Amazon Polly software. A multiple choice was located between the two blocks featuring the vocal tracks. Here, respondents were asked to answer Trinity, in order to recreate a conversation as realistic and interactive as possible. In the dominant condition, respondents have been asked only whether or not they want to complete the purchase. In fact, after having processed the fictional request, Trinity automatically adds the product to cart in the respondent's favourite colour. In the servant condition, respondents have been asked which colour variant they prefer. In fact, before adding the product to cart, Trinity informs respondents about all the colour variants available without any personalization.

Overall, we hypothesized that Trinity's behaviour is perceived as more dominant, when the simulated voice shows a low pitch of voice and provides a personalized customer experience through assertive advices. Rather, a voice with a high pitch and cautious questions express a more submissive behaviour. In fact, Trinity asked respondent whether or not he or she wanted it to do something, which implies that it needs agency to be exogenously imposed by him or her.

3.2.2 Measurement scales

Construct measurement used primarily existing scales, adapted for the context examined, that is an interaction with a smart voice assistant. In order to measure our dependent variable, namely User Satisfaction, we have implemented a multi-item scale adapted from Bruner (2017). The number of items was reduced to six, based on their fit with the context at issue. In order to assess the degree of

satisfaction experienced by respondents, they have been asked to indicate on a seven-point differential semantic scale, ranging from a negative meaning equals 1 to a positive one equals 7, how satisfied they were with the interaction with Trinity. Such a scale has been hugely employed by past research in order to determine how much individuals are satisfied with a service provided (Jones, Mothersbaugh, & Beatty, 2000; Hess, Ganesan, & Klein, 2003; Bansal , Irving , & Taylor, A Three-Component Model of Customer Commitment to Service Providers," JAMS,, 2004; Bansal, Taylor, & James, Migrating' to New Service Providers: Toward a Unifying Framework of Consumers' Switching Behaviors, 2005), with a salesperson or a dealer (Oliver & Swan, Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach, 1989a; Oliver & Swan, Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction, 1989b) and with outcomes obtained from a channel relationship, e.g., between a retailer and a vendor (Ganesan, 1994). Thus, items are reliable and plausible in order to exemplify feelings expressing both satisfaction and dissatisfaction.

Additionally, we considered some control variables in order to take into account exogenous factors that may affect participants' responses, reactions and, as a result, their evaluation of the satisfaction with the interaction. A seven-point Likert scale composed by four items measured how much respondents were concerned about data protection when interacting with a smart speaker. Such a scale has been adapted from a study measuring the primary dimensions of individuals' concerns about organizational practices about information privacy (Kukar-Kinney & Close, 2010), by extending individuals' concerns from companies to smart speaker. Also, we a implemented a seven point-Likert scale composed by three items adapted from Chiang & Dholakia (2003) in order to measure how likely respondents were to purchase the product at issue online. We replaced the general label, "product", with "USB cable". Then, we considered also how much familiar respondents were with smart speakers as the more experienced they are, the more conscious and meaningful their answers will be. The items we developed were adapted from Arnthorsson, Berry and Urbany's seven-point differential semantic scale (1991), which measures experience in using and shopping for a certain type of product being considered. Eventually, the frequency with which respondents actually make voice shopping plays a crucial role. In fact, how frequently individuals purchase a product online via voice assistant may influence their evaluation of the interaction. Current online purchases through voice commands were measured on a validated seven-point Likert scale that has been adapted from Esch, et al. (2006).

Eventually, the last section of the survey is usually reserved for demographics. For the purpose of compiling descriptive statistics, participants were asked to report their age, gender, nationality, occupation and highest level of instruction on the last page of the questionnaire.

3.2.3 Manipulation check

In order to measure whether respondents actually perceived the presentation they have been assigned to as we aimed, a manipulation check was required. Thus, based on a seven-point Likert scale adapted from Kim & Kramer (2015), respondents were asked how much they perceived Trinity as a "partner" and as a "companion". Similarly, the manipulation check for the servant condition asked them to indicate on a seven-point Likert scale adapted from Kim & Kramer (2015), how much they perceived Trinity as a "servant" and as an "attendant". Eventually, respondents have been asked to evaluate also how they perceived the behaviour of the voice assistant they have just interacted with, based on a validated scale that has been adapted from Kiesler (1983). In order to measure the effectiveness of the voice assistant behaviour manipulation, five semantic differential items asked respondents to indicate how they perceive the assistant on a seven-point scale. Also, they have been explicitly instructed to refer to voice and way of coming up with.

Like dependent and control variables, manipulation checks are multi-item scales and, as such, their scale reliability had to be analysed. Then, each scale will be computed as the same variable averaging all items. the following is a summary table reporting all measurement scales implemented in the survey with related reliability coefficients and scale items.

SCALE NAME	**RELIABILITY COEFFICIENT (Cronbach's alpha)**	**SCALE ITEMS**
User Satisfaction	$\alpha = .923$	*Indicate how much satisfied you are with the interaction with Trinity, by checking the cell that best gives your answer:* *The interaction with Trinity displeased / The interaction with Trinity pleased me;* *I am very dissatisfied with the interaction with Trinity / I am very satisfied with the interaction with Trinity;*

		The interaction with Trinity did a poor job for me / The interaction with Trinity did a good job for me; *I am unhappy with the interaction with Trinity / I am happy with the interaction with Trinity;* *The interaction with Trinity is frustrating / The interaction with Trinity is enjoyable;* *I have a very unfavourable opinion on the interaction with Trinity / I have a very favourable opinion on the interaction with Trinity.*
Privacy Concerns	$\alpha = .926$	*Indicate how much do you agree with the following statements, considering a scale ranging from 1 (= strongly disagree) to 7 (= strongly agree):* *"It usually bothers me that smart speakers ask me for personal information";* *"When smart speakers ask me for personal information, I sometimes think twice before providing it";* *"It bothers me to give personal information to so many smart speakers";* *"I'm concerned that smart speakers are collecting too much personal information about me".*
Online Purchase Intention	$\alpha = .868$	*Indicate how much do you agree with the following statements, considering a scale ranging from*

		1 (= strongly disagree) to 7 (= strongly agree): *"The likelihood that I would search for this product online is high";* *"I would be willing to buy this product online";* *"I would be willing to recommend my friends to buy this product online".*
Familiarity with smart speaker	$\alpha = .842$	*Overall, do you consider yourself unfamiliar or familiar with smart speakers?* *"Unfamiliar / Familiar";* *Overall, do you consider yourself not informed or informed on smart speakers?* *"Not informed at all / Very. informed";* *Do you consider yourself expert in smart speakers?* *"I know nothing about / I know much about".*
Online purchase via smart speaker frequency	$\alpha = .937$	*Indicate how frequently did you undertake the following actions, considering a scale ranging from 1 (= not at all) to 7 (= very frequently):* *"How often have you bought online via smart speaker in the past?";* *"How often do you consume something bought online via smart speaker?".*

Partner presentation manipulation check	$\alpha = .895$	*Taking into account the presentation, indicate how much do you agree with the following statements, considering a scale ranging from 1 (= strongly disagree) to 7 (= strongly agree):* *"Trinity is like a partner to the user.";* *"Trinity is like a companion to the user."*
Servant presentation manipulation check	$\alpha = .646$	*Taking into account the presentation, indicate how much do you agree with the following statements, considering a scale ranging from 1 (= strongly disagree) to 7 (= strongly agree):* *"Trinity is like a servant to the user.";* *"Trinity is like an attendant to the user."*
Behaviour manipulation check	$\alpha = .874$	*Taking into account the voice and the approach, indicate how did you perceive the behaviour of Trinity:* *"Submissive/Dominant";* *"Servile/Commanding";* *"Spineless/Overbearing";* *"Submissive/Bossy";* "*Slavish/Tyrannical".*

3.3 Results

3.3.1 Descriptive statistics

The total sample size is 320 respondents, who filled in the online survey between 13th and 21th September 2020. The sample was conveniently selected in Italy exploiting the author's social sphere. Respondents have been filtered out based on whether they completed the survey. 191 people completed the questionnaire. Moreover, 3 participants were included, although they left the survey uncompleted due to the unfulfillment of the dominant behaviour manipulation check section. Nevertheless, they scored a satisfactory progress level (91%). Therefore, the final valid number of participants was 194.

The analysed sample exhibited a gender distribution of 30.9% male (N = 60) and 67.5% female (N = 130). Respondents were Italian between 17 and 70 years old. The mean age was 24 years (SD = 9.53). The frequency distribution of educational level showed that most of the respondents has a master's degree (80 respondents, 41.2%), even though there is an important share of people who have a bachelor's degree (40 respondents, 20.6%) followed by who is still attending college (33 respondents, 17%), who has a high school diploma (32 respondents, 16.5%). Moreover, 53.6% of respondents was composed by students (N = 104), followed by 23.7% of full-time employee (N = 46), 8.8% of part-time employee (N = 17), 9.8% of unemployed searching for a job (N = 19) and 2.6% of pensioners (N = 5).

Each respondents was randomly assigned to one of four conditions: 1) partner presentation – submissive behaviour (N = 55), 2) partner presentation –dominant behaviour (N = 42), 3) servant presentation – submissive behaviour (N = 50), 4) servant presentation – dominant behaviour (N = 47). In fact, exposure to certain experimental condition was managed through Qualtrics randomizer tool that ensured a quite equal exposure of all the four conditions, that included all the values for presentations, as well as behaviour. As reported above, condition seen by the highest number of valid respondents was the one where Trinity acted in a submissive manner after being presented as a partner (28.4%), whereas the least viewed was the one where Trinity acted in a dominant manner after being presented as a partner (23,2%). Nevertheless, the overall distribution is quite even.

Conditions

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	PPSB	55	28,4	28,4	28,4
	PPDB	42	21,6	21,6	50,0
	SPSB	50	25,8	25,8	75,8
	SPDB	47	24,2	24,2	100,0
	Total	194	100,0	100,0	

3.3.2 Measurement Scale and Reliability analysis

As we used multi-item scales for both dependent variable and control variables, we needed to check for reliability and compute new variables averaging all items. Thus, we conducted a Cronbach's alpha test in order to have a clear measure of internal consistency.

The reliability analysis for Satisfaction showed a reliability coefficient of $\alpha = .923$, with an excellent internal consistency. Even the multi-item scale measuring Privacy Concern showed internal consistency with an excellent reliability coefficient ($\alpha = .926$) as well as Online Purchase Intention with a good Cronbach's alpha ($\alpha = .868$). Familiarity with a good Cronbach's alpha ($\alpha = .842$) and Online Purchase Via Smart Speaker Frequency with an excellent Cronbach's alpha ($\alpha = .937$). Eventually, we tested also the reliability of multi-item scales for manipulation check. It resulted that Servant Presentation Manipulation check scale exhibited a questionable but still acceptable coefficient ($\alpha = .646$), Partner Presentation Manipulation check scale exhibited a good coefficient ($\alpha = .895$) and Dominant Behaviour Manipulation check scale exhibited a good coefficient ($\alpha = .874$).

Later, when the reliability of multi-item scales was assured, the next step consisted in grouping all the items of the same scale into a new single variable. As a consequence, the six items of the dependent variable User Satisfaction scale were computed into a new variable simply called "*SAT*"; the four items of the Privacy Concerns were computed into a new variable called "*PRCONC*"; the three items of the Online Purchase Intention were computed into a new variable called "*ONPINT*"; the three items of the Familiarity were computed into a new variable called "*FAM*"; the two items of the Online Purchase Via Smart Speaker Frequency were computed into a new variable called "*SMONPFREQ*"; the two items of the Servant Manipulation Check were computed into a new variable called "*MNCK_SERVP*"; the two items of the Partner Manipulation Check were computed into a new variable called "*MNCK_PARTP*"; the five items of the Manipulation Check Behaviour were computed into a new variable called "*MNCK_DOMBEH*".

Now, we can compare satisfaction means among groups of respondents based on what condition they have been assigned to. Results showed that the most satisfying condition was the second one (M =

5.46), followed by the third one (M = 5.17), then, the fourth one (M = 4.97) and, lastly, the first one (M = 4.96). It would seem that user satisfaction is affected by a fit between presentation and behaviour, since the highest level of satisfaction corresponds to Partner Presentation – Dominant Behaviour and Servant Presentation – Submissive Behaviour conditions. Nevertheless, we do not know yet whether differences are statistically significant, so we cannot make inferences about the population.

Report

SAT

Conditions	Mean	N	Std. Deviation
PPSB	4,9636	55	1,83857
PPDB	5,4563	42	1,34752
SPSB	5,1700	50	1,68308
SPDB	4,9716	47	1,41520
Total	5,1254	194	1,60176

3.3.3 Manipulation check

As we derived from literature two different ways of presenting a voice assistant in order to convey two different images that respondents may perceive (partner or servant), we had to verify if our sample actually perceived it as we aimed.

Therefore, we asked for two questions related to people's perception of the way the assistant is presented and, then, carried out an Independent Samples t-test to control the results. The dichotomous variable referred to presentation ("PRES" = 1 which means partner; "PRES" = 0 which means servant) was the grouping variable and the manipulation check question for partner was the dependent one. The effect of the manipulation was not significant, $t(192) = -.83$, $p = .41$ ($p > .05$), although the partner condition actually scored higher in the partner presentation (M = 3.7, SD = 1.87) than in the servant presentation (M = 3.47, SD = 1.93). This means that the descriptors intended to convey the perception of Trinity as a partner of the user failed as there is not a statistically significant difference between perceptions of the two groups.

Group Statistics

	servant:partner	N	Mean	Std. Deviation	Std. Error Mean
MNCK_PARTP	servant	97	3,4742	1,93026	,19599
	partner	97	3,7010	1,87036	,18991

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MNCK_PARTP	Equal variances assumed	1,266	,262	-,831	192	,407	-,22680	,27290	-,76507	,31147
	Equal variances not assumed			-,831	191,810	,407	-,22680	,27290	-,76508	,31147

Nevertheless, we run another Independent Samples t-test checking for the effectiveness of the servant presentation manipulation. Again, the dichotomous variable, “PRES”, (“partner” = 1; “servant” = 0) was the grouping variable and the manipulation check question for servant as dependent one. It resulted that respondents assigned to the servant presentation (M = 5.02, SD = 1.36) perceived the assistant described as servant more than those assigned to the partner presentation (M = 4.18, SD = 1.63). Thus, despite previous results the two presentations significantly differ in terms of hierarchical role respondents perceive Trinity play, $t(192) = 3.88$, $p < .05$.

Group Statistics

	servant:partner	N	Mean	Std. Deviation	Std. Error Mean
MNCK_SERVP	servant	97	5,0155	1,35680	,13776
	partner	97	4,1804	1,63169	,16567

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MNCK_SERVP	Equal variances assumed	1,104	,295	3,876	192	,000	,83505	,21547	,41006	1,26004
	Equal variances not assumed			3,876	185,816	,000	,83505	,21547	,40998	1,26013

A manipulation check was required also for the manipulation of voice assistant behaviour (dominant vs. submissive). Therefore, we carried out a third Independent Samples t-test with dichotomous variable referred to behaviour (“BEH” = 1 which means dominant; “BEH” = 0 which means submissive) as the grouping variable and the manipulation check question for dominance as the dependent one. Again, the effect of the behaviour manipulation was significant, $t(189) = - 3.97$, $p < .05$, with scores for dominance higher in the dominant behaviour condition (M =. 3.62, SD = 1.31) than the submissive behaviour condition (M = 2.92, SD = 1.13). In summary, both manipulations were successful.

Group Statistics

	submissive:dominant	N	Mean	Std. Deviation	Std. Error Mean
MNCK_DOMBEH	submissive	105	2,9219	1,12950	,11023
	dominant	86	3,6233	1,30966	,14122

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MNCK_DOMBEH	Equal variances assumed	,298	,586	-3,973	189	,000	-,70135	,17654	-1,04959	-,35312
	Equal variances not assumed			-3,915	168,883	,000	-,70135	,17915	-1,05501	-,34769

3.3.4 Main analysis

Since we were to demonstrate that the combined action of presentation and behaviour has a positive impact on user satisfaction, it will be analysed the cross-over interaction between the presentation and behaviour variables. In particular, we want to test whether users will be more satisfied by the smart interaction when the voice assistant is perceived to engage in a dominant (submissive) behaviour, after it has been presented as a partner (servant). As mentioned above, the study required a 2 by 2 experimental design, as it involves two categorical independent variables with two levels each, namely presentation (partner vs. servant) and behaviour (dominant vs. submissive) and a continuous dependent variable, namely user satisfaction. My dissertation is primarily interested in the effect of the presentation of a voice assistant on differences in user satisfaction and, then, how this effect may differ when a consistent behaviour is undertaken by the assistant. Thus, behaviour can be considered as a moderator variable and the present analysis is to demonstrate that the effect of presentation on user satisfaction is moderated by behaviour.

Two-Way ANOVA

Initially, it was run a two-way ANOVA to assess whether conveying a voice assistant can significantly influence user satisfaction according to the presentation conveyed to users, the behaviour the assistant engaged in and their interaction. Although we have two levels for each independent variable, it has been preferred Analysis of Variances over two separate t-tests as the latter could have increased the probability of type one error, namely rejecting the null hypothesis and assuming significant mean differences between groups, although there was no difference. Moreover, unlike multiple t-test, ANOVA allows to study the interaction effect between two independent variables on a dependent one. Levene's test is to test an important assumption of ANOVA, namely the homogeneity of variance. It is valid when the null hypothesis is not rejected, that is when the error variance of the dependent variable is equal across groups. In this case, Levene's test for equality of

error variances was significant at significance level equals .006 ($p < .05$). Thus, we had to reject the null hypothesis and the homogeneity of variances assumption.

Levene's Test of Equality of Error Variances[a,b]

		Levene Statistic	df1	df2	Sig.
SAT	Based on Mean	4,230	3	190	,006
	Based on Median	2,881	3	190	,037
	Based on Median and with adjusted df	2,881	3	178,255	,037
	Based on trimmed mean	4,041	3	190	,008

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Dependent variable: SAT

b. Design: Intercept + PRES + BEH + PRES * BEH

Additionally, Tests of Between-Subjects Effects confirms the failure of the ANOVA by demonstrating that model fit was not significant ($F(3,190) = .941$, $p > 0.05$) which means that there are no statistically significant mean differences on the user satisfaction due to presentation or behaviour. Accordingly, neither presentation ($F(1,190) = .362$, $p > 0.05$) nor behaviour ($F(1,190) = .405$, $p > 0.05$) main effects were significant as well as interaction effect ($F(1,190) = 2.23$, $p > 0.05$). Moreover, the Partial Eta Squared values exhibit that a small percentage of variance in user satisfaction is explained by presentation alone (only 0.2%) as well as by behaviour alone (only 0.2%) or by the interaction term (only 1.2%).

Descriptive Statistics

Dependent Variable: SAT

servant:partner	submissive:dominant	Mean	Std. Deviation	N
servant	submissive	5,1700	1,68308	50
	dominant	4,9716	1,41520	47
	Total	5,0739	1,55419	97
partner	submissive	4,9636	1,83857	55
	dominant	5,4563	1,34752	42
	Total	5,1770	1,65444	97
Total	submissive	5,0619	1,76084	105
	dominant	5,2004	1,39719	89
	Total	5,1254	1,60176	194

Tests of Between-Subjects Effects

Dependent Variable: SAT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	7,250[a]	3	2,417	,941	,422	,015	2,823	,255
Intercept	5077,345	1	5077,345	1977,160	,000	,912	1977,160	1,000
PRES	,931	1	,931	,362	,548	,002	,362	,092
BEH	1,040	1	1,040	,405	,525	,002	,405	,097
PRES * BEH	5,736	1	5,736	2,233	,137	,012	2,233	,318
Error	487,920	190	2,568					
Total	5591,556	194						
Corrected Total	495,170	193						

a. R Squared = ,015 (Adjusted R Squared = -,001)

b. Computed using alpha = ,05

Two-Way ANCOVA

Since the analysis shows no statistically significant effects on the dependent variable, we run a two-way ANCOVA by controlling for four metric covariates. In fact, they are variables that cannot be manipulated but are related to the dependent variable and should be accounted for, since they may mislead the effect of other variables. Thus, ANCOVA will allow to compare satisfaction means among groups while controlling for these extraneous variables.

Descriptive Statistics

Dependent Variable: SAT

Presentation	Behaviour	Mean	Std. Deviation	N
servant	submissive	5,1700	1,68308	50
	dominant	4,9716	1,41520	47
	Total	5,0739	1,55419	97
partner	submissive	4,9636	1,83857	55
	dominant	5,4563	1,34752	42
	Total	5,1770	1,65444	97
Total	submissive	5,0619	1,76084	105
	dominant	5,2004	1,39719	89
	Total	5,1254	1,60176	194

First of all, the Levene's test was significant ($p = .004$) which means that we had reject null hypothesis, that is variances are equal across groups.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: SAT

F	df1	df2	Sig.
4,617	3	190	,004

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + PRCONC + ONPINT + FAM + SMONFREQ + PRES + BEH + PRES * BEH

Although the basic assumption of the analysis of variance is violated (homoscedasticity), we proceeded with the interpretation of results.

Tests of Between-Subjects Effects

Dependent Variable: SAT

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	20,061[a]	7	2,866	1,122	,351	,041	7,854	,476
Intercept	183,829	1	183,829	71,967	,000	,279	71,967	1,000
PRCONC	,266	1	,266	,104	,747	,001	,104	,062
ONPINT	8,163	1	8,163	3,196	,075	,017	3,196	,428
FAM	1,452	1	1,452	,568	,452	,003	,568	,117
SMONFREQ	2,771	1	2,771	1,085	,299	,006	1,085	,179
PRES	,781	1	,781	,306	,581	,002	,306	,085
BEH	1,809	1	1,809	,708	,401	,004	,708	,133
PRES * BEH	4,676	1	4,676	1,831	,178	,010	1,831	,270
Error	475,109	186	2,554					
Total	5591,556	194						
Corrected Total	495,170	193						

a. R Squared = ,041 (Adjusted R Squared = ,004)

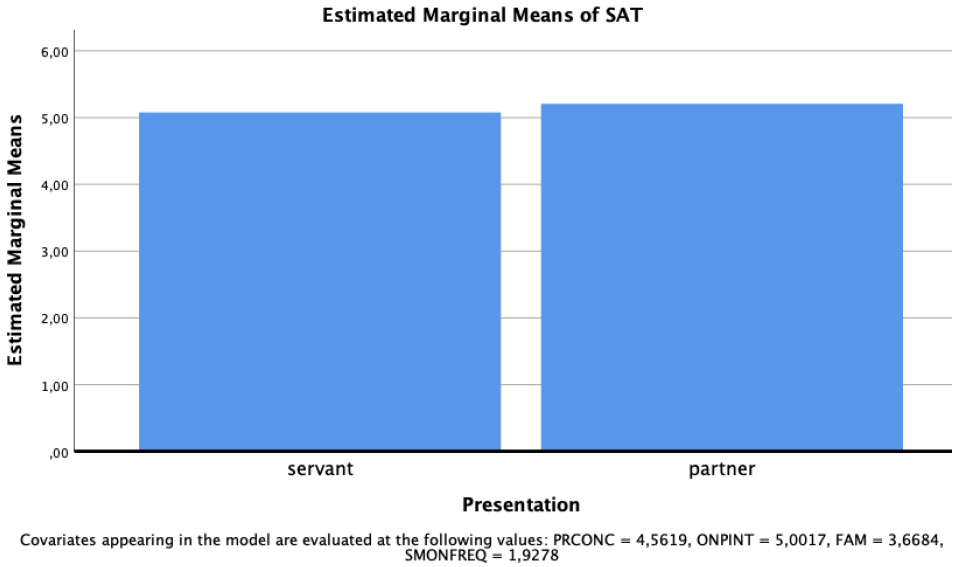
b. Computed using alpha = ,05

Unfortunately, the model fit was not significant $F(7,186) = 1.12$, $p = .351$. Consequently, there are no significant mean differences on user satisfaction due to presentation or behaviour. Thus, although sample means are different, this finding cannot be extended to the entire population.

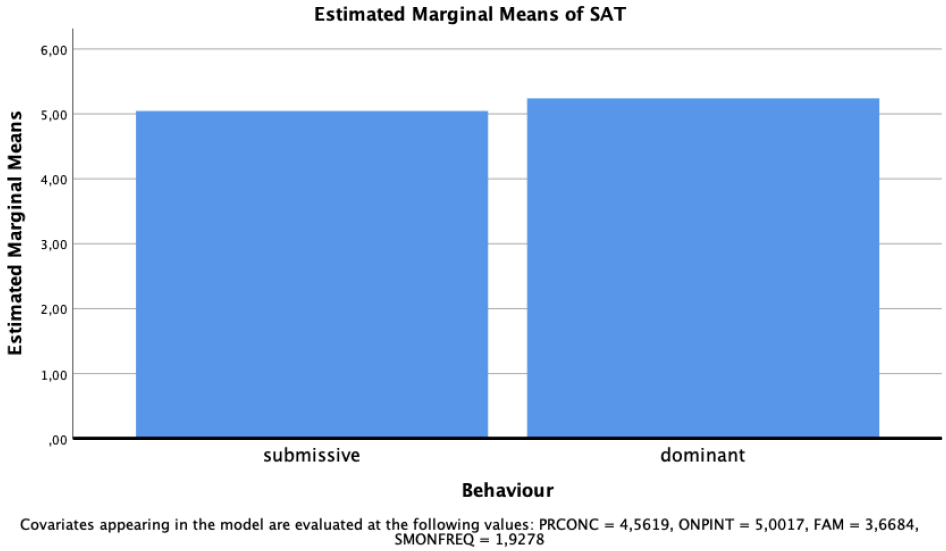
Also, covariates resulted in a non-statistically significant effect on user satisfaction., i.e. $F_{\text{Privacy Concerns}}(1,186) = .104$, $p = .747$, $F_{\text{Online Purchase Intention}}(1,186) = 3.20$, $p = .075$., $F_{\text{Familiarity}}(1,186) = .568$, $\alpha = .452$ and $F_{\text{Online Purchase Via Smart Speaker Frequency}}(1,186) = 1.09$, $\alpha = .299$. Thus, control variable alone cannot influence the level of user satisfaction.

It was not found a significant main effect of presentation on user satisfaction ($F(1,186) = .306$, $p = 0.581$), which means that presentation do not predict user satisfaction regardless of behaviour. Accordingly, the Partial Eta Squared explains that only 0.2% of variance in user satisfaction is explained by the presentation alone. Thus, we can state that user satisfaction mean does not significantly differ between two groups of presentation, even controlling for covariates. Nevertheless,

when Trinity is presented as a partner (M = 5.21; SD = .166), users have been more satisfied by the interaction compared to when it is presented as a servant (M = 5.08; SD = .165).



As far as behaviour, it has a not statistically significant main effect on user satisfaction (F(1,186) = .708, p < .401). The partial eta squared indicated that behaviour alone explains only the 0.4% of variance in user satisfaction. Thus, we can state that user satisfaction mean does not significantly differ between two groups of behaviour, even controlling for covariates. As well as presentation, also behaviour showed mean differences in user satisfaction when Trinity behaves in a dominant manner (M = 5.24; SD = .16) compared to when it behaves in a submissive manner (M = 5.04; SD = .17).



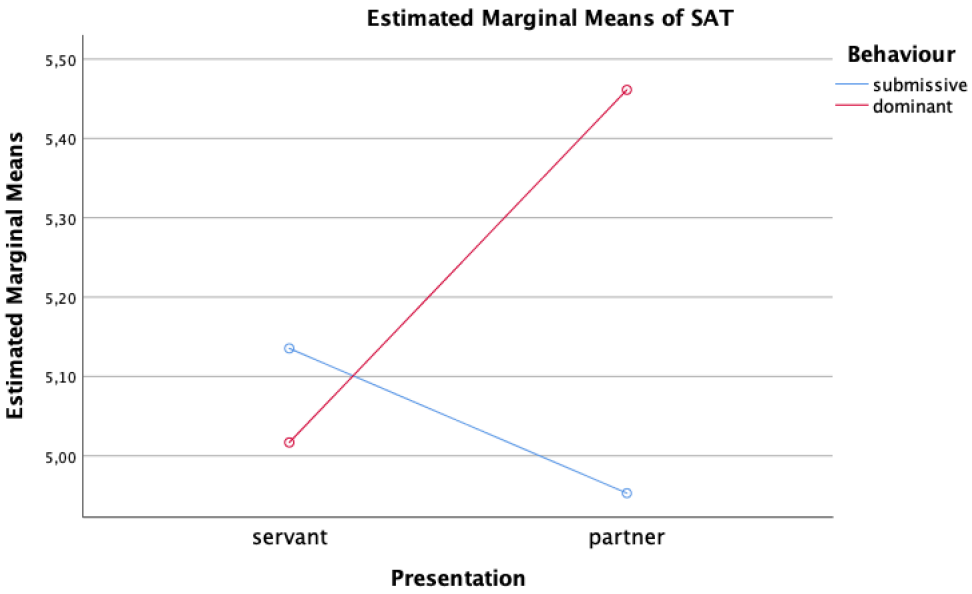
Lastly, we checked for the p-value of the interaction term in order to verify the validity of our main hypothesis (H1). According to what the study aimed at demonstrating, effects should be viewed in the context of a two-way interaction between presentation and behaviour. It was not found a significant interaction between these two variables on satisfaction, F(1,186) = 1.83, p = .178. Thus,

smart voice assistant behaviour does not moderate the effect of presentation on user satisfaction. Nevertheless, a voice assistant presented as a partner has been rated much more in satisfaction when it acts in a dominant manner (M = 5.46; SD = .249) compared to voice assistant presented as a partner when it acts in a submissive manner (M = 4.95; SD = .217). Similarly, a voice assistant presented as a servant has been rated much more in satisfaction when it acts in a submissive manner (M = 5.14; SD = .227) compared to voice assistant presented as a servant when it acts in a dominant manner (M = 5.02; SD = .237). Accordingly, the Partial Eta Squared showed that the interaction term "*PRES.*BEH.*" explains a larger percentage of variance in user satisfaction (η= 1.0%) than "*PRES.*" (η= 0.2%) and *"BEH.*" (η= 0.4%) alone. As a result, we can infer that when there is a fit between presentation and behaviour the level of satisfaction will be higher than when there is a misfit. Nevertheless, this effect is not statistically significant and, as such, cannot be extended to the entire population ($p > .05$). By comparison, there was no significant difference between partner presentation when it acts in a submissive manner and servant presentation when it acts in a dominant manner. Even, there was not much difference between means, although a voice assistant constructed as a submissive partner scored slightly higher. In other words, partner engaging in a submissive behaviour (M = 4.95, SE = .217) and servant assistants engaging in a dominant behaviour (M = 5.02, SE = .237) generally elicit in users the same degree of satisfaction.

Eventually, a clarification is due. Based on statistical significance, we cannot make inferences about the entire population, but we should be content with the validity of the hypothesis for the selected sample.

Cross-over Interaction Line Plots

As the interaction plot reported below exhibits, there is an interaction effect which confirms that the effect of one independent variable, i.e. presentation, on a dependent variable, i.e. user satisfaction, is different for different levels of the other independent variable, i.e. behaviour (submissive vs. dominant). Specifically, it is a disordinal interaction with a cross-over effect as represented by the line segments crossing each other. Thus, there is a change in the rank order of the effects of presentation across the levels of behaviour.



Covariates appearing in the model are evaluated at the following values: PRCONC = 4,5619, ONPINT = 5,0017, FAM = 3,6684, SMONFREQ = 1,9278

In particular, when Trinity was presented as a servant (partner), a submissive (dominant) behaviour elicited in respondents a level of satisfaction that was higher than the one elicited by a dominant (submissive) behaviour. When Trinity was presented as a partner (servant), a submissive (dominant) behaviour elicited a lower level of satisfaction than when it was presented as a servant (partner), whereas a dominant (submissive) behaviour elicited a higher level of satisfaction. Thus, switching from the servant presentation to the partner presentation, submissive behaviour showed a decrease in satisfaction, while dominant behaviour showed an increase. In conclusion, also the plot shows a fit between presentation and behaviour, where a voice assistant presented as servant and acting in a submissive manner satisfies users the most. Accordingly, a voice assistant presented as partner and acting in a dominant manner satisfies users the most.

3.4 General discussion

First of all, a brief digression about our sample of respondents is due. As we mentioned above, gender distribution was quite even. As a matter of fact, results can be applied equally to men (N = 60) and women (N = 131). Moreover, most of the sample was composed by young people. This allowed more reliable and valid results as Millennials (born between 1981-2000) "were raised during the dot.com boom, and have seen the development of MP3 players, YouTube, and smartphones as well as the impact of technology in all aspects of their lives from healthcare, transportation, to communication" (Gibson & Sodeman, 2014). Therefore, they are provided with an intrinsic familiarity with technologically innovative devices like smart speaker compared to previous generations.

As for the stimuli shown to respondents, there were four scenarios resulting from the combination of different presentations and behaviours that have been manipulated in two separate conditions. The partner presentation emphasizes the work that the voice assistant makes together with the user, like a partner would do. Conversely, the servant presentation emphasizes the work that the voice assistant makes for the user, like a servant would do. Basically, this first manipulation is to convey an image of the assistant that is placed at the same level of the user. Then, we considered two different behaviour the assistant may engage in. A dominant behaviour was given by a lower pitch of voice, assertive advices and a personalized user experience, whereas a submissive behaviour is given by a higher pitch of voice, hesitant question and an ordinary user experience. One of the first elements subject of the analysis has been the manipulation check. It checked for whether the manipulation was successful. In particular, it verified whether it exists a link between the wording chosen for presentation and the perceived role played by the assistant (partner vs. servant). Analogously, it verified also whether it exists a link between voice and wording chosen for the conversation, on the one hand, and the perceived level of dominance shown by the assistant, on the other hand. Fortunately, the assumptions made before the empirical study ware correct. In fact, Trinity was perceived as more "dominant" in the dominant condition than in the servant. The manipulation check was statistically significant. Moreover, it was perceived as more "servant" and "attendant" when presented as a servant than as a partner. The manipulation check was statistically significant. Unexpectedly, although Trinity was perceived as more "partner" and "companion" when presented as a partner than as a servant, the manipulation check was not statistically significant. This means that we cannot assume that descriptors intended to present Trinity as a partner are valid for the entire population. Nevertheless, this failure result can be compensated by what emerged from the servant manipulation check. In fact, both presentation checks aimed at verifying that Trinity is perceived in two different ways under the two conditions of presentation (partner vs. servant). Although we did not find that partner descriptors made Trinity be perceived as more partner in the partner condition

than in the servant, we found that servant descriptors made her be perceived as more servant in the servant condition than in the partner condition. Hence, it is the same meaning that has been differently phrased. Thus, we can extend results to the entire population. In conclusion, partner presentation manipulation was successful for our sample and mean differences in the perceived hierarchical role depending on presentation phrasing can be extended to the entire population.

Then, we were interested in whether presentation and behaviour of voice assistant may separately influence the user satisfaction. No significantly main effect was found neither for presentation nor for behaviour. Nevertheless, our sample showed that presentation alone actually affects user satisfaction with a partner presentation being more satisfying than a servant one. Also, it was found that behaviour alone affects user satisfaction with a dominant behaviour being more satisfying than a submissive one. Only limitation of our study is the impossibility to infer about the entire population.

Moreover, my dissertation also predicted that subjects would be more satisfied by a smart interaction performed by a voice assistant they perceive as actually playing the role it is expected to. In particular, subjects who actually perceive a voice assistant as dominant, when it was previously presented as a partner, would be more satisfied, as well as subjects who actually perceive it as submissive, when it was previously presented as a servant. Unfortunately, results did not support this hypothesis. Nevertheless, when respondents have been exposed to a partner (servant) presentation of the voice assistant followed by a dominant (submissive) interactive behaviour that perfectly matches the essence of a partner (servant), they have actually scored higher in user satisfaction, than a partner (servant) presentation followed by a submissive (dominant) behaviour. By considering results limited to our sample, we can infer that the interaction between presentation and behaviour positively affects user satisfaction when they are matched. Thus, behaviour moderates the effect of presentation on satisfaction. The moderating role is even more effective when a fit occurs: Partner Presentation – Dominant Behaviour and Servant Presentation – Submissive Behaviour. In fact, the combined action of presentation and behaviour generates an increase in user satisfaction that is higher than the increase generated by both presentation and behaviour alone. In particular, we can see that a servant presentation combined with a submissive behaviour scored higher in user satisfaction than she would have scored alone, although it is still lower than the combined action of partner presentation and dominant behaviour. Again, only restriction of our study is the impossibility to extend results to the entire population.

Hence, although presentation and behaviour separately can trigger an increase in satisfaction, their combined action can generate a greater increase regardless of the statistical significance of the results. In conclusion, our hypothesis is verified for our sample, but cannot be extended to the entire population.

3.5 Theoretical implications

The first relevant finding is that voice can effectively influence the overall mental representation that users create of a voice assistant. Previous literature already examined the importance of this feature in depicting a certain image of smart assistant. Scholars studied how a wider range of tones and an accelerated speaking rate Bonner (2000) or a higher pitch of voice Sundar, Jung, Waddell, & Kim (2017) influences the perception of an individual in a social context. In particular, Sundar, Jung, Waddell, & Kim applied this finding to the Human Robot Interaction (HRI) dimension in order to manipulate the perception of a robot demeneor as playful or serious. We extended further this research to the world of smart assistant by investigating their behaviour in the perspective of dominance shown. Similarly, the way in which the intelligent automated entity actually interacts with humans has an important impact on the degree to which respondents perceive it as dominant or submissive. As we wanted to convey the image of a partner or servant voice assistant, our manipulation of dominance degree is inspired by a study conducted by Nass, Moon, Fogg, Reeves, & Dryer (1995). In fact, they identified some features in human personality which can be applied also to computers. In particular, they investigated how speaking persuasively without showing hesitation in making decisions makes individuals perceive the computer as dominant. Accordingly, we attached similar features to a voice assistant by phrasing its responses through assertion. Also, we made the dominant assistant provide a more personalised experience to users (i.e., purchasing a product in the variant of user's favourite colour) in order to. make him or her believe that the assistant knows and values their preferences like a close person usually does unlike a mere servant. In fact, Schweitzer, Belk, Jordan, & Ortner (2019) stress how users want more from their AI-based devices.

Even though it has never been expressly referred to as presentation, our definitions and descriptions of Trinity smart speaker was inspired by Kim & Kramer who identified the wording most suitable for being assigned to either a partner or a servant (2015). Nevertheless, the present study decided to add all these previous notions to a new concept which is that of the text of a product presentation.

Although, behavioural or presentation components of smart speaker have been widely explored, it was never been deepened their combined action like in the present dissertation. We discovered that user satisfaction can be increased by the type of presentation and behaviour conveyed to users when they are matched base on a fit between the role a speaker is promised to play and the role it actually plays. This may be an important contribution to previous literature in this field that should be further explored as it provides some answers, but also paves the way for many other interesting paths which will be explored later.

3.6 Managerial implications

As we previously mentioned, users value a personalized user experience. Accordingly, IT companies should strike for supply consumers with what they want. In a managerial perspective, the possibility to customize a product experience should be an option available on a smart speaker by making it storage personal data and other information of customers in order to create an experience as smooth and enjoying as possible.

Another managerial implication concerns the importance of voice in influencing the perception of the smart assistant role. Therefore, managers could manipulate voice characteristics in order to create a track voice that communicate a certain role that is suitable for the target they aim to reach. In this way, a company or a brand would be able to be spiritually aligned with customers by making them feel similar to or, even, connected with both the smart assistant and, by extension, also the company or the brand themselves. In fact, we found that users appreciate when smart assistants undertake interaction in accordance with what they are intended to be. From my dissertation, it emerged that a crucial element of the interaction seemed to be voice.

Similarly, wording can impact the role perceived by users. Thus, managers could and should indulge in the wording chosen for the introduction of the smart product to the public, the basic description of the product itself and the vocabulary stored in its artificial memory. According to previous findings, wording manipulation should be adapted to the context in which the smart voice assistant will be marketed. As far as voice characteristics, also the speech phrasing and the wording should fit the objective target.

All the items that can be manipulated in order to convey a certain image of voice-controlled smart assistants can be furtherly exploited for managerial purposes. In fact, the role communicated by certain voice or wording causes users to create a mental representation of the assistant itself in their mind. This practice may be extended also to the mental image of the brand when it aims at reinforcing brand awareness by eliciting positive reactions in consumers and when it aims at renewing or cleaning up brand image by re-inventing itself.

Overall, it has to be pointed out that the present study demonstrated a preference of the investigated sample for smart assistants showing a dominant behaviour compared to those showing a submissive behaviour. A similar finding concern presentation of smart assistants. In fact, they have been proved to elicit a higher level of satisfaction in users when it is presented as a partner than as a servant.

3.7 Limitations and suggestions future research

This sector of the market is so innovative and stimulating that too much enlargement potential exists. According to the limitations of the present study, some suggestions we be mentioned hereafter.

The sample could be finished including only individuals who actually have a past experience with smart speakers and who, even, are very familiar with and satisfied by voice shopping. Similarly, the same search could be conduct in a country where smart speaker use is reported as more usual and cleared through customs. Otherwise it could be considered a task easier than voice shopping that may elicit more positive responses.

In addition, it has been reported that a voice assistant is perceived mostly as a female entity or, at best, as asexual (Schweitzer, Belk, Jordan, & Ortner, 2019). It could be analysed whether the gender affiliation expressed by voice can to some extent influence the opinion the user creates in his or her own mind as a manipulation variable in place of the pitch of voice.

Moreover, as our findings were not significant, it may be expanded the sample of investigation. In addition, our covariates were reliable but did not have a significant main effect on our dependent variable, i.e. satisfaction. Thus, they could be found other variables to take into account when investigating the effect of presentation, behaviour and their interaction on satisfaction.

My dissertation considered only the effect on satisfaction, whereas it could have considered an element more practical and useful for companies such as willingness to purchase. On the other hand, it could have been implemented a measuring variable that did not require to respondents a practical effort resulting in an action. For example, attitude towards product measures the cognitive and emotional responsiveness to certain external stimuli, in order to assess whether certain product features can sort any effect on individuals exposed to those stimuli without any practical action.

After discovering the partial failure of our analysis, we gathered some interesting insights by collecting feedback from some respondents. Three main issues emerged: some respondents revealed they were not easily able to recall the scenario they were assigned to, when asked (scenarios weakness); some respondents mentioned that there were too many instructions to read and remember that made the survey too long and complex (survey complexity and length); some respondents ran into some system errors while carrying out the questionnaire so that they had to leave it and re-enter it (system errors). The only two problems that may be solved are scenarios weakness and survey complexity and length. The former can be accounted for by a more noticeable manipulation of scenarios or by a cognitively effortless setting of scenarios. For the latter, it may be helpful a more cognitively effortless setting of the overall survey.

As far as scenarios weakness, there are many noticeable elements that can be manipulated in presentation and behaviour of voice assistants. Future research could try to test whether they

significantly impact the perception of the role played by the smart assistant as hypothesized. For example, some noteworthy suggestions can be derived from Schweitzer, Belk, Jordan, & Ortner (2019). They reported how users who consider a voice assistant as a servant, a partner or a master perceived the interaction with it. Their findings showed that the appearance of the assistant is important (servant like a dog; partner like a person) as well as the emotions communicated via wording (servant is unemotional; partner is emotional) and the most suitable task for the assistant (basic tasks for servant; complex task for partner). Lastly, scholars found that users want to feel they still have a control over the machine. Thus, it may helpful a manipulated the presence or not of a button that make the user feel necessary to impose agency on the voice assistant.

Conclusions

To conclude this empirical study, we are going to report the main findings and the related contribution to existing literature concerning smart objects and cognitive processes behind their evaluation.

Users perception has been demonstrated to be effectively influenced by certain features showed by smart objects. As a matter of fact, the anthropomorphising literature deal with how human traits can be applied to inanimate entities like smart assistants. Human mind tends to associate them with typically human roles and characteristics. Objects would become more familiar and easily processable, and, consequently, individuals would approach more positively with them. Therefore, the perception of the user is misled by what he himself wants to see and, as such, can be manipulated depending on what he is exposed to. Even in the case of voice assistants, there are elements that actually prompt a new vision of these machines, giving them an ontological relevance. Based on certain characteristics shown, they can accomplish a real role within user's life and be assigned behavioural meanings. This vision opens the way to the central part of our analysis that manipulates some basics of voice assistants.

First of all, (smart) product presentation involves an accurate description of the assistant and its functions. It aims at providing a certain image of the smart speaker by introducing also the task it has to accomplish with or for users. Both in a collaborator and subordinate perspective, the agent is the object by showing the ontological relevance we above- mentioned.

As far as the second component, that is behavior, it is useful to remember that the subject of our investigation is voice-controlled intelligent assistants. Therefore, character manifestations can be exhibited only in conversations undertaken with the user. In these interactions manipulation may take place both on tool (voice) and style (tone and wording) of communication. The vocal characteristics have been extensively demonstrated to make an animate or inanimate subject be perceived in a desired way. In the case of voice assistants, voice is crucial. On the other hand, the tone and wording with which a speech is undertaken can influence the perception that the interlocutor has of the overall interaction and of the speaker. Also, voice assistants can convey a certain nuance to the conversation, based on the degree of confidence showed.

The focal point of our analysis was to identify an interaction effect between these two elements to be manipulated. Thus, to conclude this brief overview of the most relevant findings of my dissertation, we cannot fail to mention the effective moderating power of behaviour on presentation. When the latter is matched with the way of behaving, it can improve user experience. In particular, we investigated how such a fit influences the satisfaction due to the interaction.

Overall, having demonstrated how the perception of the individual can be manipulated, this finding can be extended to other fields of application for managerial purposes. Companies can bend consumers perception to their will depending on the product, brand or, by extension, corporate image they are to convey.

References

AbuShawar, B., & Atwell, E. (2015). ALICE Chatbot: Trials and Outputs. *Computaciòn y Sistemas, 19*(4), 625-632.

Accenture Interactive. (2014). *The Internet of Things: The Future of Consumer Adoption.* Retrieved from accenture.com: https://www.accenture.com/t20150624t211456__w__/us-en/_acnmedia/accenture/conversion-assets/dotcom/documents/global/pdf/technology_9/accenture-internet-things.pdf

Aggarwal, P., & McGill, A. L. (2007). Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products. *Journal of Consumer Reaserach, 34*(4), 468-479.

Aggarwal, P., & McGill, A. L. (2012). When brands seem human, do humans act like brands? Automatic behavioral priming effects of brand anthropomorphism. *Journal of Consumer Research*, 307–323.

Al-Heeti, A. (2016, January 4). *Amazon has sold more than 100 million Alexa devices Holiday sales of the Echo Dot also exceeded Amazon's expectations.* Retrieved from cnet.com: https://www.cnet.com/news/amazon-has-sold-more-than-100-million-alexa-devices/

ANSA. (2020, April 9). *Mercato IoT italiano a 6,2 miliardi,+24%*. Retrieved from www.ansa.it: https://www.ansa.it/sito/notizie/tecnologia/tlc/2020/04/07/mercato-iot-italiano-a-62-miliardi24_c6e7fbd4-3482-4473-ba98-4f5135d2d037.html

Apple Inc. (2020, July 4). *Siri*. Retrieved from apple.com: https://www.apple.com/siri/

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. *Computer Human Behaviour*, 183–189.

Arnthorsson, A., Berry, W. E., & Urbany, J. E. (1991). Difficulty of pre-purchase quality inspection: Conceptualization and measurement. *Advances in Consumer Research Volume, 18*, 217-224.

Ashton, K. (2009). That ‘internet of things’ thing. *RFID journal, 22*(7), 97-114.

Atzori, L., Iera, A., & Morabito, G. (2010). The internet of things: A survey. *Computer Networks, 54*(15), 2787 – 2805.

Atzori, L., Iera, A., & Morabito, G. (2014). From Smart objects to social objects: the next evolutonary step of the Internet of Things. *Communications Magazine*, 97-105.

Bailey, J. E., & Pearson, S. W. (1983). Development of a Tool for Measuring and Analyzing Computer User Satisfaction. *Management Science, 29*(5), 530-545.

Bansal , H. S., Irving , P. G., & Taylor, S. F. (2004). A Three-Component Model of Customer Commitment to Service Providers," JAMS,. *Journal of the Academy of Marketing Science, 32*(3), 234-250.

Bansal, H. S., Taylor, S. F., & James, Y. S. (2005). Migrating' to New Service Providers: Toward a Unifying Framework of Consumers' Switching Behaviors. *Journal of the Academy of Marketing Science., 33*(1), 96-115.

Belk, R. W. (2013). Extended self in a digital world. *Journal of Consumer Research, 40*(3), 477-500.

Billings, L. (2007). Rise of Roboethics. *Seed Megazine*. Retrieved from scribd.com: https://it.scribd.com/document/7865212/rise-of-roboethics

Bonner, B. L. (2000). The effects of extroversion on influence in ambiguous group tasks. *Small Group Research, 31*(2), 225–244.

Brophy, A. L. (1959). Self, role and satisfaction. *Genetic Psychology Monographs, 59*, 263-308.

Bruner, G. C. (2017). *Marketing scales handbook: multi-item measures for consumer insight research.* Fort Worth, Texas: GCBII Productions, LLC.

Canhoto, A. I., & Arp, S. (201). Exploring the factors that support adoption and sustained use of health and fitness wearables. *Journal of Marketing Management*, 32-60.

Carli, L. L. (1989). Gender differences in interaction style and influence. *Journal of Personality and Social Psychology, 56*(4), 565–576.

Celli, F., Ghosh, A., Alam, F., & Riccardi, G. (2016). In the mood for sharing contents: Emotions, personality and interaction styles in the diffusion of news. *Information Processing & Management*, 93–98.

Chandler, J., & Schwarz, N. (2010). Use does not wear ragged the fabric of friendship: Thinking of objects as alive makes people less willing to replace them. *Journal of Consumer Psychology, 20*(2), 138–145.

Chiang, K. P., & Dholakia, R. R. (2003). Factors Driving Consumer Intention to Shop Online: An Empirical Investigation. *JOURNAL OF CONSUMER PSYCHOLOGY, 13*(1&2), 177–183.

Chih, W. H., Chiu, T. S., Lan, L. C., & Fang, W. C. (2017). Psychological contract violation: Impact on perceived justice and behavioral intention among consumers. *International Journal of Conflict Management, 28*(1), 103-121.

Chui, M., Löffler, M., & Roberts, R. (2010). The internet of things. *McKinsey Quarterly*, 1-9.

CNN. (2011, October 4). *Apple introduces Siri, Web freaks out*. Retrieved from edition.cnn.com: https://edition.cnn.com/2011/10/04/tech/mobile/siri-iphone-4s-skynet/index.html

Colby, K. (1999). Human-computer conversation in a cognitive therapy program. . In Y. Wilks, *Machine conversations* (pp. 9-19). Dordrecht: Kluwer.

Colby, K. M. (1975). *Artificial paranoia: a computer simulation of paranoid processes.* New York: Pergamon Press.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the selfdetermination of behavior. *Psychological Inquiry, 11*(4), 227–268.

DeLanda, M. (2015). *Philosophy and simulation: the emergence of synthetic reason.* London: Bloomsbury Publishing.

Dion, P. A., & Notarantonio, E. M. (1992). Salesperson communication style: The neglected dimension in sales performance. *Journal of Business Communication, 29*(1), 63–77.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review, 114*(4), 864 – 886.

Epstein, J., & Klinkenberg, W. (2001). From Eliza to Internet: a brief history of computerized assessment. *Computers in Human Behavior, 17*, 295-314.

Esch, F. R., Langner, T., Schmitt, B. H., & Geus, P. (2006). Are brands forever? How brand knowledge and relationships affect current and future purchases. *Journal of Product & Brand Management, 15*(2), 98-105.

Etherington, D. (2014, November 6). *Amazon Echo Is A $199 Connected Speaker Packing An Always-On Siri-Style Assistant*. Retrieved from techcrunch.com: https://techcrunch.com/2014/11/06/amazon-echo/?guce_referrer_us=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnLw&guce_referrer_cs=In7f4oLqVbKUQEY6Wm8tmQ&guccounter=2

Festinger, L. (1962). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Fontino, G., Lackovic, M., Russo, W., & Trunfio, P. (2013). A discovery service for smart objects over an agent-based middleware. In M. Pathan, G. Wei, & G. Fortino, *Internet and Distributed Computing Systems (IDCS)* (pp. 281–293). Springer.

Fortino , G., Guerrieri , A., Russo , W., & Savaglio, C. (2014). Middlewares for Smart Objects and Smart Environments: Overview and Comparison. In G. Fontino, & P. Trunfio, *Internet of*

Things Based On Smart Objects: Technology, Middleware and Applications and Computing (pp. 1-27). Springer.

Fortino, G., Rovella, A., Russo, W., & Savaglio, C. (2014). On the Classification of Cyberphysical Smart Objects in the Internet of Things. *Proceedings of International Workshop on Networks of Cooperating Objects Smart Cities*, 3–9.

Fournier, S. (1998). Consumers and their brands: Developing relationship theory in consumer research. . *Journal of Consumer Research , 24*(4), 343–353.

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a Stronger CASA: Extending the Computers Are Social Actors Paradig. *Human-Machine Communication*, 71-85.

Ganesan, S. (1994). Determinants of Long-Term Orientation in Buyer-Seller Relationships. *Journal of Marketing, 58*(2), 1-19.

Gibson, L. A., & Sodeman, W. A. (2014). Millennials and Technology: Addressing the Communication Gap in Education and Practice. *Organization Development Journal*, 63-75.

Gilbert , D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 21–38.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 107–119.

Goldberg, Y. (2016). A Primer on Neural Network Models for Natural Language Processing. *Journal of Artificial Intelligence Research*, 345–420.

Google. (2019, September 18). *Google Assistant. New voices for your Google Assistant in nine countries*. Retrieved from blog.google.com: https://www.blog.google/products/assistant/new-voices-your-google-assistant-nine-countries/

Google. (2020, January 7). *Google Assistant. A more helpful Google Assistant for your every day*. Retrieved from blog.google.com: https://blog.google/products/assistant/ces-2020-google-assistant/

Google. (n.d.). *Hey Google Panoramica*. Retrieved July 2020, from assistant.google.com: https://assistant.google.com/intl/it_it/

GrowthEnabler. (2017). *MARKET PULSE REPORT, INTERNET OF THINGS (IoT).* Retrieved from Growthenabler.com: https://growthenabler.com/flipbook/pdf/IOT%20Report.pdf

Gruenfeld, D. H., Inesi, E., Magee, J., & Galinsky, A. (2008). Power and the Objectification of Social Targets. *Journal of Personality and Social Psychology, 95*(1), 111–27.

Haeckel, S. H. (1999). *Adaptive enterprise.* Boston, MA: Harvard Business School Press.

Hess, R. L., Ganesan, S., & Klein, N. M. (2003). Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction. *Journal of the Academy of Marketing Science, 31*(2), 127-145.

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins, & A. W. Kruglanski, *Social psychology: Handbook of basic principles* (pp. 133–168). New York: Guilford Press.

Hoffman, D. L., & Novak, T. P. (2015). *Emergent Experience and the Connected Consumer in the Smart Home Assemblage and the Internet of Things.* The Center for the Connected Consumer, the George Washington University School of Business.

Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the internet of things: An assemblage theory approach. *Journal of Consumer Research, 44*(March), 1178–1205.

Hofstadter, D. R. (1996). Preface 4 the ineradicable Eliza effect and its dangers, epilogue. In D. R. Hofstadter, *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought* (pp. 155-169). New York: Basic Books.

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

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction toVoice Assistants. *Medical Reference Services Quarterly, 37*(1), 81-88.

Internet 4 Things. (n.d.). *Wearable (dispositivo indossabile), cos'è, applicazioni, notizie ed approfondimenti*. Retrieved from www.internet4things.it: https://www.internet4things.it/tag/wearable/

Internet Of things Observatory of the School of Management of the Politecnico di Milano. (2018, February 16). *Non manca (quasi) più nessuno: La Smart Home apre i battenti"*. Retrieved from www.osservatorio.net: https://www.osservatori.net/it_it/osservatori/comunicati-stampa/non-manca-quasi-piu-nessuno-la-smart-home-apre-i-battenti

Jones, M. A., Mothersbaugh, D. L., & Beatty, S. E. (2000). Switching Barriers and Repurchase Intentions in Services. *Journal of Retailing, 79*(2), 259-274.

Juang, B., & Rabiner, L. (2004). *Automatic Speech Recognition – A Brief History of the Technology Development.* Oxford: Elsevier.

Ke-fa, Y. U. (2009). A study of Maintenance of Consumer-brand Relationship: from a Perspective of Psychological Contract. *Contemporary Finance & Economics*.

Kervyn, N., Fiske, S. T., & Malone, C. (2012). Brands as intentional agents framework: How perceived intentions and ability can map brand perception. *Journal of Consumer Psychology, 22*(2), 166– 176.

Kiesler, D. J. (1983). The 1982 Interpersonal Circle: A Taxonomy for Complementarity in Human Transactions. *Psychological Review, 90*(3), 185-214.

Kim, H. C., & Kramer, T. (2015). Do Materialists Prefer the "Brand-as- Servant"? The Interactive Effect of Anthropomorphized Brand Roles and Materialism on Consumer Responses. *Journal of Consumer Research*, 284-299.

Kortuem, G., Kawsar , F., Fitton , D., & Sundramoorthy, V. (2010). Smart objects as building blocks for the internet of things. *Internet Computing, 14*(1), 44-51.

Kruglanski, A. W., Jasko, K., Milyavsky, M., Chernikova, M., Webber, D., Pierro, A., & di Santo , D. (2018). Cognitive Consistency Theory in Social Psychology: A Paradigm Reconsidered. *Psychological Inquiry, 29*(2), 45-59.

Kukar-Kinney, M., & Close, A. G. (2010). The determinants of consumers' online shopping cart abandonment. *Journal of the Academy of Marketing Science, 38*, 240-250.

Lackovic, M., & Trunfio, P. (2014). A Service-Oriented Discovery Framework for Cooperating Smart Objects. In G. Fortino, & P. Trunfio, *Internet of Things Based on Smart Objects: Technology, Communications and Computing* (pp. 85-105). Springer.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2018). *Big data: The next frontier for innovation, competition, and productivity.* McKinsey Global Institute.

Mario, H., Pentek, M., & Otto, B. (2015). Design Principles for Industrie 4.0 Scenarios: A Literature Review.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: social responses to computers. *J. Soc. Issues, 56*(1), 81-103.

Nass, C., Moon, Y., Fogg, B. J., Reeves, B., & Dryer, D. C. (1995). Can Computer Personalities Be Human Personalities? *Int. J. Human-Computer Studies*, 223-239.

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. Int.J.Soc.Robot 5,. *Internal Journal of Social Robotics, 5*, 171–191.

Norton, R. (1978). Foundation of a communicator style construct. *Human Communication Research*, 99-112.

Norton, R. (1983). *Communicator Style: Theory, Applications and Measures.* . Beverly Hills: Sage.

Notarantonio, E. M., & Cohen, J. L. (1990). The effects of open and dominant communication styles on perceptions of the sales interaction. *Journal of Business Communication, 27*(2), 171–184.

Novak, T. P., & Hoffman, D. L. (2019). Relationship journeys in the internet of things: a new frameworkfor understanding interactions between consumers and smart objects. *Journal of the Academy of Marketing Science, 47*(2), 216–237.

Nuance. (n.d.). *Dragon speech recognition software trusted by millions*. Retrieved May 2020, from www.shop.nuance.co.uk: https://shop.nuance.co.uk/store/nuanceeu/en_GB/Content/pbPage.home?currency=EUR&pgmid=95401100&utm_source=google&utm_medium=cpc&utm_campaign=EHK-AO-2020-DragonPC_Ecom+/+DBU+/+DPI+/+Dragon+/+Naturally+Speaking+/+Product+/+EU+-+EN+/+-+/+Exact+/+Desktop&keyw

Oliver, R. L., & Swan, J. E. (1989a). Consumer Perceptions of Interpersonal Equity and Satisfaction in Transactions: A Field Survey Approach. *Journal of Marketing, 53*(2), 21-35.

Oliver, R. L., & Swan, J. E. (1989b). Equity and Disconfirmation Perceptions as Influences on Merchant and Product Satisfaction. *Journal of Consumer Research, 16*(3), 372-383.

Pervin, L. A. (1968). Performance and Satisfaction as a function of individual-environment fit. *Psychological Bulletin, 69*(1), 56-68.

Proksch , M., Orth, U. R., & Cornwell, T. B. (2015). Competence enhancement and anticipated emotion as motivational drivers of brand attachment. *Psychology & Marketing, 32*(9), 934–949.

Rear, J. (2020, April 22). *www.telegraph.co.uk*. Retrieved from The best fitness trackers and watches: https://www.telegraph.co.uk/recommended/leisure/best-fitness-trackers-watches/

Risley, J. (2015, November 17). *One year after Amazon introduced Echo, half a million people have told Alexa, 'I love you'*. Retrieved from geekwire.com: https://www.geekwire.com/2015/one-year-after-amazon-introduced-echo-half-a-million-people-have-told-alexa-i-love-you/

Salerno, A. (2018, March 14). *125 ore ogni anno: ecco il tempo che si "risparmia" con le smart city.* Retrieved from www.internet4things.it: https://www.internet4things.it/smart-city/125-ore-ogni-anno-ecco-il-tempo-che-si-risparmia-con-le-smart-city/

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management*, 1-23.

Sela, A., Wheeler, S. C., & Sarial-Abi, G. (2012). We Are Not the Same as You and I: Causal Effects of Minor Language Variations on Consumers' Attitudes toward Brands. *Journal of Consumer Research, 39*(3), 661–61.

Sirianni, N. J., Bitner, M. J., Brown, S. W., & Mandel, N. (2013). Branded Service Encounters: Strategically Aligning Employee Behavior with the Brand Positioning. *Journal of Marketing, 77*(6), 108-123.

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information Privacy: Masuring Individuals' Concerns About Organizational Practices. *MIS Quarterly, 20*(2), 167-196.

Smith, K. T. (2018). Marketing via smart speakers: what should Alexa say? *Journal of Strategic Marketing*, 1-16.

Statista Research Department. (2020, July 2). *Smart speakers - Statistics & Facts*. Retrieved from statista.com: https://www.statista.com/topics/4748/smart-speakers/

Sundar, S. S., Jung, E. H., Waddell, T. F., & Kim, K. J. (2017). Cheery companions or serious assistants? Role and demeanor congruity as predictors of robot attraction and use intentions among senior citizens. *Human-Computer Studies*, 88-97.

Swinyard, W. R., & Smith, S. M. (2003). Why People (Don't) Shop Online: A Lifestyle Study of the Internet Consumer. *Psychology & Marketing, 20*(7), 567-597.

Trappl, R. (2013). *Your Virtual Butler. The Making-of.* Springer.

Trappl, R., Petta, P., & Payr, S. (2002). *Emotions in Humans and Artifacts.* Cambridge, Mass.: MIT Press.

Van den Heuvel, S., & Schalk, R. (2009). The relationship between fulfillment of the psychological contract and resistance to change during organizational transformations. *Social Science Information, 48*(2), 283-313.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing, 68*(1), 1–17.

Webster, C., & Sundaram, D. S. (2009). Effect of service provider's communication style on customer satisfaction in professional services setting: The moderating role of criticality and service nature. *Journal of Services Marketing, 23*(2), 103–113.

Williams, K. C., & Spiro, R. L. (1985). Communication style in the salesperson-customer dyad. *Journal of Marketing Research, 22*(4), 434–442.

Wired. (2017, July 19). *I Google Glass sono tornati, ma solo per le imprese*. Retrieved from www.wired.it: https://www.wired.it/gadget/accessori/2017/07/19/ritorno-google-glass-imprese/

Woods, H. S. (2018). Asking more of Siri and Alexa: feminine persona in service of surveillance capitalism. *Critical Studies in Media Communication, 35*(4), 334-349.

Wu, J., Chen, J., & Dou, W. (2016). The Internet of Things and interaction style: the effect of smart interaction on brand attachment. *Journal of Marketing Management*, 1-15.

Wu, K. N. (2018). Voice Assistant. *School of Computer Science, University of Birmingham*, 1-11.

Wuenderlich, N. V., Heinonen , K., Ostrom, A. L., Patricio , L., Sousa, R., Voss, C., & Lemmink, J. G. (2015). "Futurizing" smart service: implications for service researchers and managers. *Journal of Services Marketing, 29*(6/7), 442-447.

Appendix

The partner presentation scenario reads as follows:

*"Trinity is a smart speaker designed to **partner with** you.*
*Its capacity to work **with** you makes it the perfect **companion** for you."*

On the other hand, the servant presentation scenario reads as follows:

*"Trinity is a smart speaker designed to **serve you**.*
*Its capacity to work **for** you makes it the perfect **attendant** for you."*

The text pronounced by Trinity in the dominant behaviour condition is reported below:

"I just added to cart a USB charger cable Tech-Power in your favourite colour variant. The total of your order is 8 €. Let's buy it!"

When respondents selected "*Yes*", Trinity responded as follows:

"I ordered a USB charger cable Tech-Power for 8 €. Tracking information will be sent to your e-mail address. If we want to do more shopping together, I am ready!"

On the other hand, when respondents selected "*No*", Trinity responded as follows:

"All right, if we want to do more shopping together, I am ready!"

On the other, hand, the text pronounced by Trinity in the submissive behaviour condition was as follows:

"I just added for you to your cart a USB charger cable Tech-Power. The total of your order is 8 €. Five colour variants are available: white, black, red, blue and green. Which one do you want me to buy for you?"

When respondents selected "*White", "Black", "Red", "Blue"* or *"Green"*, Trinity responded as follows:

"I ordered for you a USB charger cable Tech-Power of the colour you chose for 8 €. Tracking information will be sent to your e-mail address. If you want me to do more shopping for you, I am at your disposal."

On the other hand, when respondents selected "*None*", Trinity responded as follows:

"All right, if you want me to do more shopping for you, I am at your disposal."

Summary

Framing Smart Voice Assistant in a contemporary scenario

Nowadays, we are living the so-called Fourth Industrial Revolution which coincides with the spread on the global market of Internet of Things. Since 2014, the IOT systems represented a revolutionary innovation. They consist of nonlinear interactions connecting heterogeneous entities in the digital world with equally heterogeneous entities in the physical world. Two worlds are connecting to each other in order to generate a single cooperative system based on intelligent devices and internet connection. The Internet of Things has been defined as the most "potentially disruptive technological revolution of the last 50 years", since it developed empowering and increasingly pervasive machines like RFID or sensors which, once combined with deep-rooted mechanisms and practices (like distributed computing or artificial intelligence), exhibited their revolutionary contribution (Atzori, Iera, & Morabito, 2010). The development of recent technologies contributed to the spread of IoT since they are used on a daily basis even by unaware users. For example, NFC (Near Field Communication) consists in the short-range (a few centimetres) transmission of data from an initiator to a target. ATM cards, public transport subscription and any contactless card work like this. The innovative potential of IoT system is supplied by the smart objects (SOs) composing it. A smart object has been defined by Fortino *et al*. (2014) as an "autonomous, cyber-physical object augmented with sensing/actuating, processing, storing, and networking capabilities." SO-based IoT has been widely studied due to its internal cooperative dynamics among connected objects aimed at delivering "evolved services to humans or to other objects" (Kortuem, Kawsar , Fitton , & Sundramoorthy, 2010). Hoffman and Novak expanded DeLanda's work and provided the following definition of such a systemic whole: "In the IoT, sensors that collect data, and actuators that transmit that data, are being increasingly incorporated in all manner of consumer objects commonly found in and around the home, worn on or in the body, and used in consumption activities involving shopping, entertainment, transportation, wellness, and the like" (Hoffman & Novak, 2018). Despite the disruptive and innovative nature of the phenomenon, it has some drawbacks which may be traced back to the lack of clear and well-defined normative, regulatory, technological and market governance. According to a study conducted by Accenture Interactive who surveyed more than 2000 U.S. respondents, the principal barriers to adopt IoT devices are concerns with privacy, lack of awareness, lack of perceived value and price concerns (Accenture Interactive, 2014). Lackovic & Trunfio (2014) and Fontino, Lackovic, Russo, & Trunfio (2013) divided smart objects into four main categories depending on both functional and non-functional characteristics, that are type (namely, the type of object – smart pen, smart table, etc.), device (namely, either the hardware or software component), services (namely,

the list of operations implementing the service), and location (namely, the position of the smart object). Furthermore, smart objects can be classified depending on the final recipients who can be consumers (consumer segment) or businesses (business segment) (GrowthEnabler, 2017). Overall, IoT works with data which are transformed into strategical information that can facilitate human lives and satisfy managerial requirements. Thus, smart objects provide advantages for people as such and consumers. Firstly, consumers save time since IoT systems simplify routine operations offering an added value that is counterbalanced both by the price and by the risk that that companies have increasingly consumer information. Then, consumer experience can be enriched via more in-depth analyses of purchasing habits and overall customer journey leading to greater understanding and real-time solutions. Based on past literature, we identified seven different applications of smart objects: Smart Home (and Building), Wearable Smart Devices (e.g. Smart Watches), Smart Health, Smart Mobility, Smart Retail, Smart Factory and Smart City.

Voice Assistants have been defined as "software agents which could interpret human voice and respond through a proceeded voice" (Hoy, 2018). Moreover, they can be traced back to the broader category of virtual assistants, since the latter are nowadays conceived as software-based applications capable of interpreting human natural language and completing tasks for users (Hoy, 2018). Nevertheless, it must be noticed that virtual assistants were originally conceived as practitioners who provided remote web assistance to complete the business services, as mentioned by Youngblood (2006). Virtual assistants may employ sophisticated technologies like speech recognition, Natural Language Processing (NLP) and computer vision. Before going forward with Voice Assistants, a brief focus on three operating principles of virtual assistants is required. In fact, Wu (2018) states that virtual assistants can be accessed via three main methods which are image-based, text-based or voice-based. Voice Assistants do not represent a revolutionary invention in the field of computer technology, but rather the final outcome of a progressive and systematic evolution of pre-existing inventions that have been sophisticated and enhanced. In fact, since when the ELIZA project was developed in 1966, the idea of a written conversation between a machine and a real person took shapes. ELIZA was a chatbot created by Joseph Weizenbaum at the Massachusetts Institute of Technology whose aims was to simulate a likely conversation with a psychotherapist. Overall, a chatbot is a conversational software agent which interacts with users interpreting and using human natural language. Indeed, ELIZA managed to make sure users believed they were talking to a real human being instead of a computer by asking them about, for example, their emotional status (Epstein & Klinkenberg, 2001). Subsequently, an opensource chatbot, named A.L.I.C.E which means Artificial Linguistic Internet Computer Entity, implemented a less sophisticated Natural Language Processing. It demonstrated that chatting with a virtual assistant may have practical objective other

than mere entertainment (AbuShawar & Atwell, 2015). Eventually, in 1975 another psychiatrist-authored conversational software was developed in order to simulate a conversation with a paranoid patient (Colby K. M., 1975; Colby K. , 1999). Unlike ELIZA, it was characterized by a hostile and defensive tone which nonetheless made users perceive the machine as a real human (Epstein & Klinkenberg, 2001). Nevertheless, unlike Voice Assistant who, once received the user input, provide vocal output, the above-mentioned conversational virtual assistants communicate with users only via written dialogues. Natural Language Processing (NLP) is fundamental in both scripted and spoken conversations with machines by enabling them to translate verbal input into interpretable and consequently executable commands. Moreover, both HARPY e Dragon's Naturally Speaking systems may be considered as ancestors of Voice Assistants. They implement speech recognition technology based on recognizing human voice as input to be processed (Juang & Rabiner, 2004).

One of the latest improvements in the field of Internet of Things, which result in Voice Assistants as they are currently conceived, is the incorporation into objects or physical environments. Siri represents one of the first installations of virtual assistants on a smartphone with the launch of iPhone 4S on October 4, 2011 (CNN, 2011). Also, Amazon Echo, generally referred to as Alexa, and Google Home are the earliest examples of virtual assistants which activate home automation systems via voice input in order to enhance the actions that the domestic environment can complete. Overall, they are designed to assist users with managing simple tasks through the natural language voice-control interfaces. In particular, smart objects provided with such virtual assistants allowed artificial intelligence technology to actually speak with users for the first time (Smith K. T., 2018). Overall, IoT devices have become increasingly required in products from several industries such as automotive, smart home, healthcare, insurance and transports. In fact, an increasing number of brands are striving to implement smart conversation with branded artificial intelligence-powered assistant as an alternative communication channel with their consumers. Basically, it may be perceived as an actual touchpoint with the brand which may facilitate the interaction with their brand and, as a result, affect the overall brand consideration (Novak & Hoffman, 2019). The smart speaker market is characterized by a duopoly-like configuration composed by two most popular vendors, namely Amazon's Echo/Alexa and Google Home/Google Assistant. They are followed by minor brands including also Apple probably due to the late debut of HomePods. In fact, it entered the market at the beginning of 2018 and has been trying "to catch up in the game" (Statista Research Department, 2020).

What determines a Partner vs. Servant Voice Assistant?

Before analysing how the perception of a voice assistant playing a certain role can be manipulated, a preliminary introduction on anthropomorphising is due. In fact, the present study considers features that commonly belong to the human sphere as affecting the perceived agency that human mind assigns to the voice assistant. In particular, a voice assistant who shows more agency is perceived as a partner and a voice assistant who shows less agency is perceived as a servant. Anthropomorphism is the natural human tendency to imbue inanimate objects with a set of typically human traits which, taken as a whole, outline a real personality (Aggarwal & McGill, 2007). Scholars have extensively demonstrated the ease with which people provide anthropomorphic descriptions of unhuman agents which comprise anything that "acts with apparent independence, including animals, natural forces, religious deities, and mechanical or electronic devices" (Epley, Waytz, & Cacioppo, 2007). Smart products constitute a subcategory of conventional objects, since they are able to actually interact with humans through short conversations, queries, commands, questions and more (Schweitzer, Belk, Jordan, & Ortner, 2019). An early exemplification of the anthropomorphising practice in the computer science can be retrieved in the so-called ELIZA effect. In a broader sense, it can be defined as the tendency to consider computer behaviours as similar to human ones. More specifically, it is "the tendency for people to treat programs that respond to them as if they had more intelligence than they really do" (Trappl, Petta, & Payr, 2002). According to Trappl, Petta, & Payr, the ELIZA effect demonstrates the effectiveness of applicating a social perspective to the field of artificial intelligence rather than the mere programming engineering. For further clarification, Douglas Hofstadter exemplifies the ELIZA effect with the misleading belief of a casual observer that an automated teller machine displaying the words "THANK YOU" at the end of a transaction is thought as actually being thankful. Tough, it is only "printing a pre-programmed string of symbols" (Hofstadter, 1996). As we noted earlier, voice assistants can actually engage in a conversation with users, persuading them that they are relating to and cooperating with one of their own. Generally, when individuals communicate with each other, it may involve different communication styles such as friendly, relaxed, contentious, attentive, precise, animated, dramatic, open, or dominant based on personal factors (e.g. gender or social status) or contextual factors (e.g. the current situation or time) (Norton R. , 1978). By focusing on the fast-growing context of IoT and smart objects, several studies have been conducted on smart social interactions employing different styles and affecting how users overall perceive the smart assistant they talk to. Particularly, Wu, Chen, & Dou (2016) examined two smart interaction styles that are frequently implemented by companies in order to convey a certain brand image reflecting the intrinsic brand identity: friend-like and engineer-like style. They can boost consumers' perceptions of brand warmth and/or brand competence and, as a consequence, consumer emotional attachment to

brand or product (Wu, Chen, & Dou, 2016). Experiencing a sort of affection with smart objects and, in particular, voice assistants is even easier than with ordinary products due to their inner human-like intelligent components, i.e. voice, which bring them closer and make them more similar to human identities. This is how the anthropomorphising of vocal assistants occurs and works. Aggarwal and McGill demonstrated that an anthropomorphised brand can prime positive reactions and actions in consumers, after being exposed to the brand. In the same way, we want to test whether human behaviour can be triggered by the interaction with an anthropomorphised smart assistant. In particular, the present dissertation will investigate whether user satisfaction could be primed by the anthropomorphic role of servant vs. partner played by smart voice assistant based on the way it is presented to and the way it communicates with users.

Presentation

In everyday life, we are taught that first impressions count. Thus, since we are moving within a literary context of human quality extended to objects (i.e. anthropomorphism), it is right to look towards the way a voice assistant is presented. Based on some reviews of the Amazon Alexa smart speaker on the Amazon website, I gathered some features that users ascribe to the assistant, depending on whether it is perceived as either a partner or a servant. When users consider the assistant an equal partner worthy of human soul and conscience, they even address it using the feminine personal pronoun, “She”. Moreover, sometimes she is perceived as a part of the family or, at least, as a living entity who is a fundamental part of the house - rather than a mere home automation system. Some reviewers even stated that, although she was able to listen to conversations at any time, her presence never made them feel uncomfortable. There are many examples of products, companies and brands which call themselves “friend” or, even, “partner” of their consumers. For example, the State Farm brand refers to itself as “a good neighbour” for consumers (Kim & Kramer, 2015). Thus, even a smart assistant can be referred to with a friendly apposition. Overall, individuals have been demonstrated to irrationally portrait a voice assistant like Apple Siri as an attractive, likeable and helpful woman (Schweitzer, Belk, Jordan, & Ortner, 2019). Wu, Chen, & Dou depicted how a branded smart assistant interacting with users through a professional and cold wording were perceived as a competent but detached engineer (2016); on the other hand a more friendly and caring wording would seem to reflect the role of a partner. Moreover, a language featuring terms like “we” rather than “you and I” can elicit more closeness and emotional attachment, since it better expresses the concept of a more connected togetherness (Sela, Wheeler, & Sarial-Abi, 2012). Eventually, a partner assistant is an independent agent who does not require any contribution by humans. Considering the study conducted by Schweitzer, Belk, Jordan, & Ortner, some respondents perceived Siri as a servant and instinctively humanized it in a secretary “dressed up with blouse and jeans” (2019). Therefore, it would seem that

exhibiting typical features of a character who usually serves others contributes to represent a servant assistant. Nevertheless, the most appropriate representation of the "servant voice assistant", as intended by users who have actually interacted with it, consists of a nice, friendly, helpful, reliable and ready-to-please character. In fact, respondents admitted feeling a mild attachment towards the voice assistant to the point that it is assimilated to a faithful companion, i.e. a pet, rather than a slave. (Schweitzer, Belk, Jordan, & Ortner, 2019). A servant voice assistant should not emphasize the emotional connection with users. When presented, the image of servant can be conveyed by a wording like “this assistant is going to work for you” or “[…] to please you” rather than “[…] to work with you” or “[…] to support you in your work”. Accordingly, the tasks it can accomplish are considered as a service it provides to the user-master as a subordinate who serves him or her. Eventually, since users who see the assistant as a servant perceive that they literally impose agency on the object, they want to actually feel they are needed to activate the smart object and make it work. In fact, the study by Schweitzer, Belk, Jordan, & Ortner came to the conclusion that people should be supplied a control and feedback mechanism like “the ‘close door’ buttons on elevators that allow consumers to feel a sense of control” (2019).

Behaviour

In addition to the way it is described, the way in which a smart assistant behaves is important to outline the role of the voice assistant. In fact, based on a review of the literature dealing with this topic, my dissertation takes into consideration actionable reactions of the smart object which qualify it either as a partner or a servant. In particular, the intelligent assistant is able to dialogue with human beings and, thus, the way of behaving can be intended as the way of communicating, namely the communication style. In fact, it has already been demonstrated that voice manipulation can bear fruit. The modulation of intonation concern rhythm, and accentuation. Moreover, a change in the form of a word (typically the ending) is commonly used to express a grammatical function or attribute. Thus, it may be helpful to express also a certain personality. Dominance can be found in a serious behaviour that has been proved to be represented by a lower pitch of voice may represent a (Niculescu, Van Dijk, Nijholt, Li, & See, 2013). Moreover, smart user interfaces should speak human language, that is characterized by emotional component. It can be rendered also in conversations with smart assistants by a wording featuring, for example, “Please” and “Thank you”. In this way, assistants can be perceived as one of the same species and, more specifically, as an equal partner. A partner-like style employs a customer-centric strategy based on a “sense-and-respond” practice. In fact, a peer-to-peer relationship is characterized by the mutual intimate knowledge of personality and preferences of the relationship members. Similarly, a partner assistant should provide the user with tailor-made proposals, namely appropriately and specifically designated for the user by creating a personalized

experience. Such a partnership based on deep customer insights would make the user react to the voice assistant with more favourable emotions (Fournier, 1998; Kervyn, Fiske, & Malone, 2012). Overall, Hoffman and Novak (2018) theorised that an anthropomorphised personal assistant is perceived as a partner when it shows gregariousness and extraversion. By considering a smart conversation, an extravert voice assistant engages in frequent verbal interactions with users, showing also proneness to answer. Similarly, users appreciated a voice assistant "making jokes, reacting very coldly when offended and in general being good at a quick comeback" (Schweitzer, Belk, Jordan, & Ortner, 2019). In this way, the assistant demonstrates a bold and reactive personality. In this perspective, an influent power is exerted by the phrasing of proposals, since a confident assistant would come up with assertive advices rather than hesitant suggestions. Furthermore, an enterprising assistant may go beyond what is asked, deepen the request and provide a solution out of the box. For example, it should provide further insights into a definition that has been asked, find offers which can replace the research of the user, if it is not available and, also, provide additional offers to expand user's range of choice (Schweitzer, Belk, Jordan, & Ortner, 2019). As for the partner role, we took into account the study conducted by Sundar, Jung, Waddell, & Kim (2017), where a higher pitch of voice has been proved to denote a less serious behaviour (Niculescu, Van Dijk, Nijholt, Li, & See, 2013). As we assumed that seriousness suits for a dominant assistant, a high pitch of voice can be assigned to a submissive assistant. Furthermore, also the emotional component seems to be crucial. In particular, words exchanged within the conversation should be detached through unemotional wording and inexpressive tone of voice. In fact, when the smart assistant is perceived as submissive, individuals are aware that it is "just a voice behind which a person can be imagined" (Schweitzer, Belk, Jordan, & Ortner, 2019). Such a smart assistant can only afford a standard and average performance, since it has not the enterprising personality to provide a tailor-made and customized user experience. Moreover, the user is only provided the information he or she expressly asked without any further content or proposal. Thus, the assistant would come out with linear sentences which are strictly related to the issues that it deals with and that are never out-of-the-box. Moreover, the smart assistant shows hesitation and, even, tardiness to answer which makes conversation less smooth and, as a result, unrealistic. Hence, it is not surprising that, when it interacts with users, it employs only cold answers that leave no room for originality, sympathy and the ability to respond in kind (Nass, Moon, Fogg, Reeves, & Dryer, 1995). Eventually, we can assume that a servant voice assistant merely reacts to users' orders and questions and never speaks spontaneously.

Fit between behaviour and presentation

In the interaction with the smart object users should perceive the behaviour of the vocal assistant as fitting the role that it claimed to assume in the presentation, since such a fit will trigger a positive

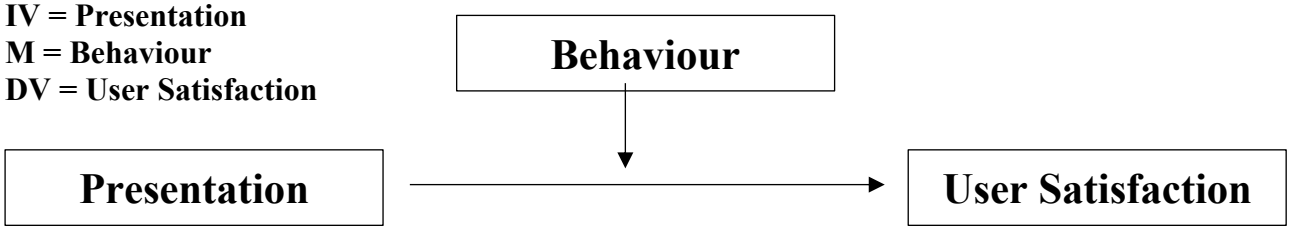
effect on the user satisfaction resulting from the interaction. It all depends on the mental representation that the user forms of the smart object resulting from the role played and, particularly, the whole of related information stored in his or her mind, on the smart object's actual behaviours and on the degree of consistency between the former and the latter. Since the second half of the last century, several in-depth studies have been carried out in order to examine the importance of consistency in human cognitive processing. Kruglanski, et al. demonstrated that people "possess a general need for cognitive consistency" (Kruglanski, et al., 2018). In fact, according to the Cognitive Consistency paradigm new information that are inconsistent with prior beliefs has a negative impact on people's affective and cognitive reaction which may result in distress. As additional proof, a study conducted by Nass et al. (1995) reported the practice of companies to exploit the style of interaction implemented by computers behaviour to convey the brand personality. More specifically, companies oriented towards customer relationships tend to assume a friendly tone of voice in order to position themselves as "their customers (best) friends" by matching a sincere, honest, and genuine personality, by conveying caring intentions and by strengthening emotional attachment towards the brand. In contrast, companies oriented towards technology and innovation prefer to position themselves in the consumer's mind as a professional and experienced problem solver by matching a rugged, efficient, and sophisticated personality, by conveying indifferent intentions and by lowering emotional attachment towards the brand (Nass, Moon, Fogg, Reeves, & Dryer, 1995). Overall, the brand distinctive features should be reflected in the communication process so that consumers perceive a high level of congruence between the brand personality and the actual behaviour of the smart object. Additionally, such a congruence "facilitates customers' information processing […] and can reinforce the brand meaning among customers" (Sirianni, Bitner, Brown, & Mandel, 2013). Therefore, there seems to be sufficient evidence to support the claim that users need to be shown a certain degree of consistency between how the voice assistant acts and the role it was presented to play.

Basically, this dissertation is going to investigate the role that a voice assistant is perceived to play. In particular, users may perceive the voice assistant as a partner rather than a servant depending on how it is presented. Additionally, they may perceive it as a dominant entity rather than submissive depending on how it interacts. Both presentation and behavioural elements can be manipulated in order to convey a certain image of the voice assistant. Moreover, users are assumed to be more satisfied when the behaviour exhibited during the smart conversation is consistent with the presentation aimed at introducing the voice assistant to users. In particular, the core hypothesis of this study will be as follows:

H1) Partner smart speaker with a dominant behaviour will be evaluated more positively than partner smart speaker with a dominant behaviour, while servant smart speaker with a submissive behaviour will be evaluated more positively than servant smart speaker with a dominant behaviour.

Thus, it will be run a between-subjects two-way Analysis of CoVariances (ANCOVA), since we have two categorical independent variables impacting on a continuous dependent variable (DV). Whose effect is controlled for some covariates. Presentation represents the focal variable (IV1) and consists of two categories or levels, namely Partner and Servant. Similarly, Behaviour represents the moderator (M) and consists of two categories or levels, namely Dominant and Submissive.

IV = Presentation
M = Behaviour
DV = User Satisfaction



Quantitative Analysis

Methodology

My dissertation aims at testing cause-effect relationships between manipulated presentation of voice assistant and satisfaction of the user related to the interaction with the assistant, considering the moderating influence of its actual behaviour. This is going to be a quantitative causal research conducted via an experimental survey which was administrated online and was accessible by mobile phones or any other digital device. The only criterion required was the sample to be large enough to generalise the results to the entire population. The total sample size is 320 respondents who have been filtered out based on whether they completed the survey. Therefore, the final valid number of participants was 194. The analysed sample exhibited a gender distribution of 30.9% male (N = 60) and 67.5% female (N = 130). Respondents were Italian between 17 and 70 years old. The mean age was 24 years (SD = 9.53). The frequency distribution of educational level showed that most of the respondents has a master's degree (80 respondents, 41.2%), even though there is an important share of people who have a bachelor's degree (40 respondents, 20.6%) followed by who is still attending college (33 respondents, 17%), who has a high school diploma (32 respondents, 16.5%). Moreover, 53.6% of respondents was composed by students (N = 104), followed by 23.7% of full-time employee (N = 46), 8.8% of part-time employee (N = 17), 9.8% of unemployed searching for a job (N = 19) and 2.6% of pensioners (N = 5).

Also, handing out the questionnaire through the Internet has been demonstrated to be more convenient. Data will be collected via Qualtrics Software. This research is going to adopt a non-probability sampling method due to convenience reasons. In fact, the online survey has been spread through a personal network of acquaintances and over social groups on the Internet.

Survey

Each respondents was randomly assigned to one of four conditions: 1) partner presentation – submissive behaviour (N = 55), 2) partner presentation –dominant behaviour (N = 42), 3) servant presentation – submissive behaviour (N = 50), 4) servant presentation – dominant behaviour (N = 47). Therefore, the overall distribution is quite even. As scenarios will differ upon two manipulation variables, i.e. presentation and behaviour, it is going to be a 2 by 2 between-subjects design. Two dummy variables have been created, coded as 0 when respondent has not been assigned to a certain condition and coded as 1 when he or she has been assigned to that condition. In particular, the presentation variable takes either the value 1, when voice assistant is introduced through descriptors which are suitable for a partner, or 0, when voice assistant is introduced through descriptors which are suitable for a servant. The behaviour variable takes either the value 1, when responses employed by the voice assistant resemble a dominant behaviour, or 0, when responses employed by the voice assistant resemble a submissive behaviour. The new dichotomous variables are labelled "*PRES*" and "*BEH*", respectively.

The manipulation of presentation will consist in modifying descriptors which have been implemented in two written descriptions of a new voice assistant, named Trinity, that has just been launched on the market. They briefly stated who the assistant is and what is its function. The construction of the behaviour scenarios started from a careful research on what product is one of the most likely to be purchased online (i.e., a USB charger cable for smartphones). Then, we exposed participants to a fictional situation where they had to imagine they want to buy such a product and they ask Trinity. The manipulation will consist of both a more (dominant) or less (submissive) high pitch of voice and a personalized (dominant) or standard (submissive) response to the user input. Moreover, Schweitzer, Belk, Jordan, & Ortner demonstrated that people tend to perceive voice assistants as dominant, when they employ in assertive advices, whereas they perceive them as submissive, when they employ in hesitant questions (2019). In both the dominant and submissive condition, respondents played two audio files which have been created via Amazon Polly software. A multiple choice was located between the two blocks featuring the vocal tracks. Here, respondents were asked to answer Trinity, in order to recreate a conversation as realistic and interactive as possible. In the dominant condition, respondents have been asked only whether or not they want to complete the purchase. In fact, after having processed the fictional request, Trinity automatically adds the product to cart in the

respondent's favourite colour. In the servant condition, respondents have been asked which colour variant they prefer. In fact, before adding the product to cart, Trinity informs respondents about all the colour variants available without any personalization.

In order to measure our dependent variable, namely User Satisfaction, we have implemented a multi-item scale adapted from Bruner (2017). Respondents, they have been asked to indicate on a seven-point differential semantic scale, ranging from a negative meaning equals 1 to a positive one equals 7, how satisfied they were with the interaction with Trinity (Cronbach's α = .923). Additionally, we considered some control variables in order to take into account exogenous factors that may affect participants' responses, reactions and, as a result, their evaluation of the satisfaction with the interaction. A seven-point Likert scale composed by four items measured how much respondents were concerned about data protection when interacting with a smart speaker (Cronbach's α = .926). Also, we a implemented a seven point-Likert scale composed by three items adapted from Chiang & Dholakia (2003) in order to measure how likely respondents were to purchase the product at issue online (Cronbach's α = .868). Then, we considered also how much familiar respondents were with smart speakers as the more experienced they are, the more conscious and meaningful their answers will be. The items we developed were adapted from Arnthorsson, Berry and Urbany's seven-point differential semantic scale (1991) (Cronbach's α = .842). How frequently individuals purchase a product online via voice assistant may influence their evaluation of the interaction. Current online purchases through voice commands were measured on a validated seven-point Likert scale that has been adapted from Esch, et al. (2006) (Cronbach's α = .937). Eventually, the last section of the survey is usually reserved for demographics. Later, when the reliability of multi-item scales was assured, the next step consisted in grouping all the items of the same scale into a new single variable.

In order to measure whether respondents actually perceived the presentation they have been assigned to as we aimed, a manipulation check was required. Thus, based on a seven-point Likert scale adapted from Kim & Kramer (2015), respondents were asked how much they perceived Trinity as a "partner" and as a "companion" (Cronbach's $\alpha = .895$). Similarly, the manipulation check for the servant condition asked them to indicate on a seven-point Likert scale adapted from Kim & Kramer (2015), how much they perceived Trinity as a "servant" and as an "attendant" (Cronbach's $\alpha = .646$). Eventually, respondents have been asked to evaluate how they perceived the behaviour of the voice assistant, based on a validated seven-point differential semantic scale composed by five items that has been adapted from Kiesler (1983) (Cronbach's $\alpha = .874$). Unexpectedly, although Trinity was perceived as more "partner" and "companion" when presented as a partner than as a servant, the manipulation check was not statistically significant. This means that we cannot assume that descriptors intended to present Trinity as a partner are valid for the entire population. Nevertheless,

this failure result can be compensated by what emerged from the servant manipulation check. In fact, both presentation checks aimed at verifying that Trinity is perceived in two different ways under the two conditions of presentation (partner vs. servant). Although we did not find that partner descriptors made Trinity be perceived as more partner in the partner condition than in the servant, we found that servant descriptors made her be perceived as more servant in the servant condition than in the partner condition. Hence, it is the same meaning that has been differently phrased. Thus, we can extend results to the entire population. In conclusion, partner presentation manipulation was successful for our sample and mean differences in the perceived hierarchical role depending on presentation phrasing can be extended to the entire population. Moreover, Trinity was perceived as more "dominant" in the dominant condition than in the servant and the manipulation check was statistically significant.

Results

Since the analysis shows no statistically significant effects on the dependent variable, we run a two-way ANCOVA by controlling for four metric covariates. In fact, they are variables that cannot be manipulated but are related to the dependent variable and should be accounted for, since they may mislead the effect of other variables. Thus, ANCOVA will allow to compare satisfaction means among groups while controlling for these extraneous variables. First of all, the Levene's test was significant ($p = .004$) which means that we had to reject null hypothesis, that is variances are equal across groups. Although the basic assumption of the analysis of variance is violated (homoscedasticity), we proceeded with the interpretation of results.

Unfortunately, the model fit was not significant $F(7,186) = 1.12$, $p = .351$. Consequently, there are no significant mean differences on user satisfaction due to presentation or behaviour. Thus, although sample means are different, this finding cannot be extended to the entire population. Also, covariates resulted in a non-statistically significant effect on user satisfaction., i.e. $F_{\text{Privacy Concerns}}(1,186) = .104$, $p = .747$, $F_{\text{Online Purchase Intention}}(1,186) = 3.20$, $p = .075$., $F_{\text{Familiarity}}(1,186) = .568$, $\alpha = .452$ and $F_{\text{Online Purchase Via Smart Speaker Frequency}}(1,186) = 1.09$, $\alpha = .299$. Thus, control variable alone cannot influence the level of user satisfaction. It was not found a significant main effect of presentation on user satisfaction ($F(1,186) = .306$, $p = 0.581$), which means that presentation do not predict user satisfaction regardless of behaviour. Accordingly, the Partial Eta Squared explains that only 0.2% of variance in user satisfaction is explained by the presentation alone. Thus, we can state that user satisfaction mean does not significantly differ between two groups of presentation, even controlling for covariates. Nevertheless, when Trinity is presented as a partner ($M = 5.21$; $SD = .166$), users have been more satisfied by the interaction compared to when it is presented as a servant ($M = 5.08$; $SD = .165$). As far as behaviour, it has a not statistically significant main effect on user satisfaction ($F(1,186) = .708$,

$p < .401$). The partial eta squared indicated that behaviour alone explains only the 0.4% of variance in user satisfaction. Thus, we can state that user satisfaction mean does not significantly differ between two groups of behaviour, even controlling for covariates. As well as presentation, also behaviour showed mean differences in user satisfaction when Trinity behaves in a dominant manner ($M = 5.24$; $SD = .16$) compared to when it behaves in a submissive manner ($M = 5.04$; $SD = .17$). Lastly, we checked for the p-value of the interaction term in order to verify the validity of our main hypothesis (H1). According to what the study aimed at demonstrating, effects should be viewed in the context of a two-way interaction between presentation and behaviour. It was not found a significant interaction between these two variables on satisfaction, $F(1,186) = 1.83$, $p = .178$. Thus, smart voice assistant behaviour does not moderate the effect of presentation on user satisfaction. Nevertheless, a voice assistant presented as a partner has been rated much more in satisfaction when it acts in a dominant manner ($M = 5.46$; $SD = .249$) compared to voice assistant presented as a partner when it acts in a submissive manner ($M = 4.95$; $SD = .217$). Similarly, a voice assistant presented as a servant has been rated much more in satisfaction when it acts in a submissive manner ($M = 5.14$; $SD = .227$) compared to voice assistant presented as a servant when it acts in a dominant manner ($M = 5.02$; $SD = .237$). Accordingly, the Partial Eta Squared showed that the interaction term "*PRES.*BEH.*" explains a larger percentage of variance in user satisfaction (η= 1.0%) than "*PRES.*" (η= 0.2%) and *"BEH.*" (η= 0.4%) alone. As a result, we can infer that when there is a fit between presentation and behaviour the level of satisfaction will be higher than when there is a misfit. Nevertheless, this effect is not statistically significant and, as such, cannot be extended to the entire population ($p > .05$). By comparison, there was no significant difference between partner presentation when it acts in a submissive manner and servant presentation when it acts in a dominant manner. Even, there was not much difference between means, although a voice assistant constructed as a submissive partner scored slightly higher. In other words, partner engaging in a submissive behaviour ($M = 4.95$, $SE = .217$) and servant assistants engaging in a dominant behaviour ($M = 5.02$, $SE = .237$) generally elicit in users the same degree of satisfaction. Eventually, a clarification is due: based on statistical significance, we cannot make inferences about the entire population, but we should be content with the validity of the hypothesis for the selected sample.

Discussion

We were mainly interested in whether subjects who actually perceive a voice assistant as dominant, when it was previously presented as a partner, would be more satisfied, as well as subjects who actually perceive it as submissive, when it was previously presented as a servant. Unfortunately, results did not support this hypothesis. Nevertheless, when respondents have been exposed to a partner (servant) presentation of the voice assistant followed by a dominant (submissive) interactive

behaviour that perfectly matches the essence of a partner (servant), they have actually scored higher in user satisfaction, than a partner (servant) presentation followed by a submissive (dominant) behaviour. By considering results limited to our sample, we can infer that the interaction between presentation and behaviour positively affects user satisfaction when they are matched. Thus, behaviour moderates the effect of presentation on satisfaction. The moderating role is even more effective when a fit occurs: Partner Presentation – Dominant Behaviour and Servant Presentation – Submissive Behaviour. In fact, the combined action of presentation and behaviour generates an increase in user satisfaction that is higher than the increase generated by both presentation and behaviour alone. In particular, we can see that a servant presentation combined with a submissive behaviour scored higher in user satisfaction than she would have scored alone, although it is still lower than the combined action of partner presentation and dominant behaviour. Only restriction of our study is the impossibility to extend results to the entire population.

We expanded existing research and applied well-established findings to a new context like smart interaction. Nevertheless, the main contribution to literature of my dissertation is the investigation of interaction behavioural or presentation components of smart speaker. In fact, it was never been deepened their combined action and the effect on satisfaction. We discovered that user satisfaction can be increased by the type of presentation and behaviour conveyed to users, when they are matched, based on a fit between the role a speaker is promised to play and the role it actually plays.

Overall, it has to be pointed out that the present study demonstrated a preference of the investigated sample for smart assistants showing a dominant behaviour compared to those showing a submissive behaviour. A similar finding concern presentation of smart assistants. In fact, they have been proved to elicit a higher level of satisfaction in users when it is presented as a partner than as a servant. According to our findings, IT companies should strike for supply consumers with what they want. In a managerial perspective, the possibility to customize a product experience should be an option available on a smart speaker by making it storage personal data and other information of customers in order to create an experience as smooth and enjoying as possible. Another managerial implication concerns the importance of voice in influencing the perception of the smart assistant role. Therefore, managers could manipulate voice characteristics in order to create a track voice that communicate a certain role that is suitable for the target they aim to reach as well as for the brand itself. In fact, we found that users appreciate when smart assistants undertake interaction in accordance with what they are intended to be. From my dissertation, it emerged that a crucial element of the interaction seemed to be voice. Similarly, wording can impact the role perceived by users. Thus, managers could and should indulge in the wording chosen for the introduction of the smart product to the public, the basic description of the product itself and the vocabulary stored in its artificial memory. According to

previous findings, wording manipulation should be adapted to the context in which the smart voice assistant will be marketed. As far as voice characteristics, also the speech phrasing and the wording should fit the objective target. After discovering the partial failure of our analysis, we gathered some interesting insights by collecting feedback from some respondents. Three main issues emerged: some respondents revealed they were not easily able to recall the scenario they were assigned to, when asked (scenarios weakness); some respondents mentioned that there were too many instructions to read and remember that made the survey too long and complex (survey complexity and length); some respondents ran into some system errors while carrying out the questionnaire so that they had to leave it and re-enter it (system errors). The only two problems that may be solved are scenarios weakness and survey complexity and length. The former can be accounted for by a more noticeable manipulation of scenarios or by a cognitively effortless setting of scenarios. For the latter, it may be helpful a more cognitively effortless setting of the overall survey. As far as scenarios weakness, there are many noticeable elements that can be manipulated in presentation and behaviour of voice assistants. Future research could try to test whether they significantly impact the perception of the role played by the smart assistant as hypothesized. For example, some noteworthy suggestions can be derived from Schweitzer, Belk, Jordan, & Ortner (2019). Their findings showed that the appearance of the assistant is important (servant like a dog; partner like a person) as well as the emotions communicated via wording (servant is unemotional; partner is emotional) and the most suitable task for the assistant (basic tasks for servant; complex task for partner).