How Service Quality Influences Customer Acceptance and Usage of Chatbots?

By Lars Meyer-Waarden*, Giulia Pavone, Thanida Poocharoentou, Piyanut Prayatsup, Maëlis Ratinaud, Agathe Tison, and Sarah Torné

The present study aims to investigate consumers' acceptance of and intention to reuse a chatbot in the context of automated customer service in the airline industry. In particular, we identify the most valuable factors that affect acceptance of an intention to reuse a chatbot by integrating the theoretical framework SERVQUAL. The main results show that reliability and perceived usefulness are the most important criteria that affect the intention to reuse the chatbot. Contrary to our expectations, empathy does not have any significant effect. The study suggests that in the case of an interaction with a chatbot for a purpose that may involve an economic transaction, customers prefer the chatbot for its utilitarian value, as reliability and usefulness are considered to be more important than empathy. Moreover, tangible elements play an important role in increasing the perceived ease of use.

1. Introduction

Rapidly improving digital technologies change the nature of service, customers' service experiences, and customers' relationships with firms (van Doorn et al. 2017; Rust and Huang 2014; Ostrom et al. 2015). For example, service ro-

Lars Meyer-Waarden is Professor of Marketing at Toulouse School of Management, TSM Research-UMR 5303 CNRS, University of Toulouse 1 Capitole, France, E-Mail: lars.meyer-waarden@tsm-education.fr
* Corresponding Author.

Giulia Pavone is a PhD candidate and teaching assistant at Toulouse School of Management, TSM Research-UMR 5303 CNRS, University of Toulouse 1 Capitole, France, E-Mail: giulia. pavone@tsm-education.fr

Thanida Poocharoentou is a student within the Master International Marketing of Innovative Technologies from Toulouse School of Management, University of Toulouse 1 Capitole, 2 Rue du Doyen-Gabriel-Marty, 31042, Toulouse, France, E-Mail: thanida.poocharoentou@tsm-education.fr

Piyanut Prayatsup is a student within the Master International Marketing of Innovative Technologies from Toulouse School of Management, University of Toulouse 1 Capitole, 2 Rue du Doyen-Gabriel-Marty, 31042, Toulouse, France, E-Mail: piyanut.prayatsup@ tsm-education.fr

Maëlis Ratinaud is a student within the Master International Marketing of Innovative Technologies from Toulouse School of Management, University of Toulouse 1 Capitole, 2 Rue du Doyen-Gabriel-Marty, 31042, Toulouse, France, E-Mail: maelis.ratinaud@tsm-education.fr

Agathe Tison is a student within the Master International Marketing of Innovative Technologies from Toulouse School of Management, University of Toulouse 1 Capitole, 2 Rue du Doyen-Gabriel-Marty, 31042, Toulouse, France, agathe. tison@tsm-education.fr

Sara Torné is a student within the Master International Marketing of Innovative Technologies from Toulouse School of Management, University of Toulouse 1 Capitole, 2 Rue du Doyen-Gabriel-Marty, 31042, Toulouse, France, E-Mail: sara. torne@tsm-education.fr

bots, in combination with cameras, sensors, speech recognition, big data, analytics, artificial intelligence (AI) and mobile and cloud technology, considerably impact service industries (Wirtz et al. 2018) and attract interest by the academic communities (van Doorn et al. 2017; Huang and Rust 2018; Čaić et al. 2018). An increasing number of companies are integrating AI and service robots, such as chatbots, in customer service (Huang and Rust 2018), which are computer programmes that are able to mimic human-human communication by using natural language processing and machine learning techniques (Hill et al. 2015; Araujo 2018). Chatbots interact in conversations, but instead of a human on the other end, a computer is communicating based on AI (Wünderlich and Paluch 2017). Originally, chatbots were designed to execute just simple tasks, but today their degree of complexity has increased and they are able to execute more complicated tasks, such as giving health, financial or shopping recommendations (Araujo 2018).

As of 2017, over 100,000 chatbots have been created on Facebook Messenger and consumers are increasingly interacting with them through social media and instant messaging apps (Araujo 2018). In this regard, Gartner (2018) predicts that by 2020, 85 % of all consumer interactions in customer service will be handled without a human agent. Nevertheless, consumers still seem to be reluctant to use chatbots, and they remain sceptical about the quality of their service (Forrester 2019). As 85 % of organisations intend to establish an AI-based service chatbot to offer an automated customer dialogue (Gartner 2018), it is crucial to understand how users perceive this new form of technology-mediated communication (Wünderlich and Paluch 2017). Considering that AI is rapidly reshaping customer service, there is the need to comprehend which is the most effective and beneficial way to implement these technologies in order to guarantee a higher perceived service quality and a positive customer experience. In this context, we investigate customer interactions with an existing chatbot in France, known as Flybot, which acts as a travel assistant helping users to book their trips. Launched in October 2017, Flybot works through Facebook Messenger. In a few minutes, the chatbot is able to propose to the user the best flight(s) for the destination (s)he is interested in and also offers travel tips or cheaper dates. The aim of our study is therefore to investigate consumers' acceptance of and intention to reuse the chatbot by using extended Technology Acceptance Theories (Davis et al. 1989; Davis 1989; Venkatesh and Davis 2000; Kulviwat et al. 2007; Ostrom et al. 2019; Venkatesh et al. 2012), with one added component, namely trust (Benbasat and Wang 2005; Gefen et al. 2003; Wirtz et al. 2018), which are enhanced with the widely applied SERVQUAL model (Parasuraman et al. 1985). In fact, we analyse how cognitive, social antecedents, as well as the perceived service quality's criteria that define the interaction with the chatbot (tangible elements, competence, reliability, responsiveness, empathy and credibility), affect the perceived ease of use, usefulness, trust and intention to reuse the chatbot in the context of automated customer service in the airline industry.

This study offers insights and contributions at both theoretical and managerial levels, as most of the research in the domain of chatbots comes from engineering and computer science and not from marketing literature. At a theoretical level, our study offers three main contributions to service research, shedding light on consumers' perceptions and acceptance of AI-based chatbot service agents. Firstly, as investigations in marketing are missing, van Doorn et al. (2017) recommend research on service frontline experiences with a high automated social presence but low human social presence as conceptualised in the context of chatbots or service robots. Considering that little research in the service field has empirically investigated the usage of conversational agents (Wünderlich and Paluch 2017; Hill et al. 2015; Chung et al. 2018), this is one of the first studies where perceived service quality and acceptance is analysed after a real experience with an existing chatbot. Thus, our empirical approach allows us to get meaningful insights related to the real usage of a chatbot in service settings (booking of a long haul flight). Secondly, we integrate for the first time two well-established theories widely used in service research and information systems research, SERVQUAL (Parasuraman et al. 1985) and Technology Acceptance Theories (Venkatesh et al. 2012; Ostrom et al. 2019; Kulviwat et al. 2007; Davis et al. 1989; Davis 1989), respectively, and test them in the new, emerging context of AI-based service agents (chatbots). This approach allows us to measure traditional service quality dimensions in the innovative context of AI-technology based services (chatbots), thus investigating both consumers' perceptions of the service quality and consumers' beliefs related to the technological components of the service. Finally, by highlighting the relationship between usability and aesthetic and its effect on customers' intention to reuse the chatbot, our results integrate and harmonise the literature between two different fields: service research (Parasuraman et al. 1985; Hausman and Siekpe 2009; Cronin and Taylor 1994; Buttle 1996; Asubonteng et al. 1996) and human-robot interactions (Minge and Thüring 2018; Mahlke 2007; Hill et al. 2015; Hartmann et al. 2007). Thus, the results of the study offer insights from a wider perspective by inviting a dialogue between these different disciplines and opening further research in related fields.

This article is structured as follows: after presenting our theoretical framework and hypotheses, the data and methodology are shown. This is followed by the results, as well as their discussion. We then highlight the theoretical and managerial implications. Finally, the main limits and future research directions are presented.

2. Theoretical framework

2.1. Robots and chatbots in service encounters

Service robots belong to one of the most recent innovations in the customer service domain, which have become more and more popular in customer-oriented businesses (Huang and Rust 2018) and their use is increasing. Characterised by autonomy technology with a physical embodiment, service robots have a higher level of social presence than other service technologies (Jörling et al. 2019). Service robots are system-based autonomous and adaptable interfaces that can be physical or virtual, designed as humanoid (i.e. anthropomorphic) or not and can interact, communicate and deliver services to an organisation's customers (Wirtz et al. 2018). Huang et al. (2007) define chatbots as service robots that are conversational agents that interact with users in a strictly limited domain or on a certain topic with natural language sentences, generally deployed on the Internet for the purpose of seeking information, site guidance and Frequently Asked Questions. According to Lester et al. (2004), chatbots are technologies that exploit natural language, engaging users in text-based information-seeking and task-oriented dialogues for a broad range of applications. Chatbots may differ in their level of intelligence. In this regard, Huang and Rust (2018) identify four different type of intelligences that depend on the nature of the service: mechanical, analytical, intuitive and empathetic. Mechanical intelligence is at the most basic level and concerns the ability to automatically perform repetitive tasks that do not require advanced training (Huang and Rust 2018). In this case, a chatbot is rule-based and it does not understand the external environment. Analytic intelligence refers to the ability to process information to problem solving and learn from it (Sternberg 2005; Huang and Rust 2018). Machine learning and data analytics techniques allow the technology to learn from data and find insights without being programmed, allowing mass-personalisation based on big data. Mechanical and analytical intelligences are still considered "weak AI" because they simulate intelligence but do not have intuition (Huang and Rust 2018). At the highest level of complexity, there are intuitive intelligence, which refers to the ability of the AI to think creatively and adjust effectively to novel situations, and empathetic intelligence, which is the ability to recognise and understand other people's emotions, to affect them and to respond appropriately (Sternberg 2005; Huang and Rust 2018; Goleman 1996). These two intelligences represent the most advanced generation of AI, but they are still far from reality. Chatbots can be distinguished from humans (Wirtz et al. 2018) in that service employees have their own (limited) capabilities, perceptions and weaknesses, showing their heterogeneity across individuals. Human employees need to have a deep understanding of their customers and service processes to deliver heterogeneous, individualised services and to achieve this, learning is needed. Furthermore, connecting employees to Customer Relationship Management (CRM) systems requires time and effort. Service employees represent high costs through training, but then can be a source of competitive advantage. Differentiation in service can be based on better hiring, selection, training, motivation, and organisation of service employees. In contrast, chatbots acquire knowledge quickly and system-wide through CRM systems, as well as AI, to determine optimal solutions. Furthermore, chatbots are free from human error and fatigue, do not show heterogeneity and behave identically, providing homogeneous services in a highly reliable manner (Huang and Rust 2018). Furthermore, they do not feel and express real emotions. In fact, the majority of chatbots are designed at the mechanical and/or analytic level. On the one hand, they offer the advantages of being cost- and time-effective, always available and extremely consistent, but on the other hand, they may fail to satisfy consumers and have a low potential of competitive advantage. In fact, their pre-programmed scripts may pose a risk of not properly responding to the user's requests, leading the customer to frustration and dissatisfaction. For this reason, it is important to consider the level of efficiency and competences of a chatbot to meet customers' expectations. Information provision, processing, bookings and payments are considered the most appropriate areas of service robots (Ivanov and Webster 2019).

2.2. Technology Acceptance and SERVQUAL theories for Chatbot acceptance

We integrate two main theories that are popular in information systems literature and in service literature: Technology Acceptance Theories (Kulviwat et al. 2007; Ostrom et al. 2019) based on an extended TAM (Davis et al. 1989; Davis 1989; Venkatesh and Davis 2000) with one added component, namely trust (Benbasat and Wang 2005; Gefen et al. 2003), and the SERVQUAL model (Parasuraman et al. 1985).

Technology Acceptance Models are used to predict usage and acceptance of new technologies by users (Venkatesh et al. 2012), including service robots and AI (Ostrom et al. 2019). The service robot acceptance model (sRAM) adapts and enhances the TAM by integrating social-emotional and relational elements (Wirtz et al. 2018). By drawing from the Theory of Reasoned Action (Fishbein and Ajzen 1975), Technology Acceptance Models aim to investigate the impact of external variables on internal beliefs, attitudes, and intentions (Venkatesh et al. 2012). TAM studies technology adoption behaviours by evaluating two key elements: the Perceived Usefulness (PU) and the Perceived Ease of Use (PEU). PU is defined as the degree to which a person believes that using a particular system would enhance his or her job performance. PEU refers to how a person believes that using a particular system would be free of efforts (Davis 1989). The TAM postulate that the actual usage of a technology is determined by the behavioural intention, which is jointly determined by the attitude towards using the technology and the PU (Davis et al. 1989). The attitudes towards using the technology and the PU are also affected by the PEU (Davis et al. 1989). The TAM has been widely used because of its parsimony and explication power, including the acceptance of service robots and shopbots (Gentry and Calantone 2002; Wirtz et al. 2018;). By testing three models explaining behavioural intentions to adopt shopbots – namely the Theory of Reasoned Action (TRA-Fishbein and Ajzen 1975), Theory of Planned Behavior (TPB-Ajzen 1991), and TAM (Davis 1989) – Gentry and Calantone (2002) found that TAM explains more variance of shopbot adoption than TRA and TPB.

Thus the authors confirm the appropriateness of using this model to study these service robot technologies. Nevertheless, considering the parsimony of the TAM, researchers have addressed the need to extend the model, integrating variables which may influence acceptance (Venkatesh and Davis 2000; Benbasat and Wang 2005). In this regard, trust, by reducing environmental uncertainty, complexity, and risk, and by enhancing consumer loyalty, is one of the most recognized key factors in online shopping environments integrated in the TAM (Gefen et al. 2003; Benbasat and Wang 2005).

In addition, considering that we are investigating a new technology in the context of consumer service, we integrate a service quality measurement model into an extended Technology Acceptance Model. The extended Technology Acceptance Models are useful to understand the beliefs that drive the acceptance of and the intention to use the technology, while a service quality model is helpful to better identify the service determinants that consumers consider important to evaluate consumers' perceived service quality. In the service literature, there are a number of key instruments available for measuring service quality, such as SERVQUAL (Parasuraman et al. 1985), E-SERVQUAL (Parasuraman et al. 2005), SERVPERF (Cronin and Taylor 1992; 1994) and the hierarchical model (Dagger et al. 2007; Brady and Cronin 2001). We chose the SERVQUAL model (Parasuraman et al. 1985), as it has been the major generic model used to measure and manage service quality across different service settings and various cultural backgrounds. Moreover, this model is highly valued by academics and practitioners to measure the level of customer service satisfaction (Seth et al. 2005). SERVQUAL is a well-established instrument as a result of extensive field-testing and refinement, which can be used comparatively for benchmarking purposes (Dagger et al. 2007). Parasuraman et al. (1988) identify five determinants of perceived service quality: tangibles, reliability, responsiveness, assurance and empathy. In this regard, assurance – defined as knowledge and courtesy of employees and their ability to inspire trust (Parasuraman et al. 1988, p.23) – includes the dimensions of communication, competence, credibility, security, courtesy, understanding and access. As this study investigates the perceived service quality delivered by a chatbot, which, despite its competence, may be perceived as not secure or credible, we have decided to measure separately the dimensions of competence and credibility. While chatbots or AI can be very competent in determined situations, their goals may be perceived as not aligned with the user's goals (Ostrom et al. 2019). This could reduce the perception of safety and credibility related to these technologies. Thus, our conceptual model includes the four dimensions of the updated SERVQUAL, which are tangibles, reliability, responsiveness and empathy, plus the two different dimensions of competence and credibility. This distinction will help us to better understand how each dimension has an impact on the intention to reuse the chatbot.

2.2.1. The effect of tangibles on perceived usefulness and the perceived ease of use

Tangibles refer to the physical evidence of the service and include the physical facilities, the appearance of personnel and the tools or equipment used (Parasuraman et al. 1985). Research distinguishes two types of systems or technology qualities (Mahlke 2007). Instrumental, tangible qualities concern the usability, perceived usefulness (PU) and perceived ease of use (PEU) of the technology. Non-instrumental qualities, on the other hand, concern the visual aesthetics and attractiveness of the technology. The perception of both types of qualities influences user behaviours, such as technology adoption and usage. Minge et al. (2017), Lindgaard and Dudek (2003) and Hassenzahl (2005) demonstrate that a user's judgment of a technology relies on both instrumental qualities, such as usability (PU and PEU), and non-instrumental factors, such as aesthetic features. The results of the literature are mixed about the impact of product appearance and visual aesthetics on consumer behaviours. The influence of perceived usability on the overall service/product appraisal was found to be higher than that of aesthetics (Minge et al. 2017). On the one hand, Minge and Thüring (2018) recently showed that the effect of aesthetics on product usability ("Beautiful is usable") is strong only in the short term, either before or during the early stage of adoption of a product, and it quickly vanishes once users have become acquainted with a product. On the other hand, some studies show that in the context of human-computer interfaces, aesthetics and visual appearance are important determinants of system acceptability and help to overcome a poor usability experience ("Usable gets beautiful"; Hartmann et al. 2008). Furthermore, Sonderegger and Sauer (2010) suggest that products whose aesthetics are perceived as more attractive present higher perceived usability. In particular, colours and layout are direct system features that can impact system usage through PU and PEU (Tractinsky and Lowengart 2007; Li and Yeh 2010; Heijden 2003). Former research tested the "Beautiful is usable" hypotheses only with quite "old" technologies such as ATMs (created in the 1960ies, and their usage is now current and common), cell-phones or websites (Tractinsky 2004). In contrast to ATMs, it is worthwhile to investigate chatbots, which are new and emerging technologies based on AI that can interact much deeper and autonomously with customers in a lot of service encounter situations (whereas ATMs can only perform a limited number of services, such as delivery of cash, cash transfers, account information). Thus, we hypothesize that the tangible characteristics of a chatbot interface, in particular its colours and visual appearance, should be an important factor having a positive effect on customers' PU and PEU ("Beautiful is usable") and might even help to overcome poor usability ("Usable gets beautiful"; Hartmann et al. 2008).

H1: *Tangibles (colours, visual appearance) have a positive effect on the chatbot's PU.*

H2: *Tangibles (colours, visual appearance) have a positive effect on the chatbot's PEU.*

Nevertheless, the influence of perceived usefulness on the overall appraisal of the chatbot was found to be higher than that of aesthetics (Minge et al. 2017). Furthermore, Minge and Thüring (2018) show that the effect of aesthetics on product usability is strong only in the short term, either before or during the early stage of adoption of a product, and it quickly vanishes once users have become acquainted with a product.

2.2.2. The effect of competence on perceived usefulness

Competence refers to the required skills and knowledge that the customer service agent needs to have in order to perform the service (Parasuraman et al. 1985). It includes the knowledge and skills of the contact and operational support staff, as well as the research capability of the organisation (Parasuraman et al. 1985). Research shows that competent service performance increases positive consumer responses to service encounters (Leo and Chandon 1997). Price et al. (2006) suggest that competence may be even more important in the case of brief, non-personal encounters, such as in the case of service robots. Wirtz et al. (2018) argue that in service settings characterised by complex cognitive/analytical and simple emotional/social tasks, consumers seek a competent and reliable core service with convenient customer service. Regarding chatbots, competence refers to their ability to effectively answer a request based on their knowledge, skills and the adequacy of their communication. Considering that chatbots are often rule-based and follow predetermined scripts, they may fail to properly answer customers' requests. In this regard, consumers may feel frustrated and consider the interaction useless and a loss of time (Forrester 2019). Therefore, we suggest that the competence of a chatbot has a positive effect on its perceived utilities (PU). Thus:

H3: *Competence has a positive effect on the chatbot's PU.*

2.2.3. The effect of reliability on perceived usefulness

Reliability refers to the consistency of the performance and the dependability of the service, which needs to be delivered in the right way and at the designated time (Parasuraman et al. 1985). Research shows that reliability is also important to consumers' favourable evaluations in the context of information systems (Butler and Gray 2006) and self-service technologies (Collier and Kimes 2013). In the context of digital services, reliability is defined as the correct technical functioning of a website and the accuracy of service promises and product information (Zeithaml et al. 2000). Above all, when interacting with a new service technology, customers may be especially concerned about the reliability of the new service and may perceive its performance as uncertain (Evans and Brown 1988; Dabholkar 1996). This should also be true in the case of consumers interacting with a service chatbot. In this context, perceived reliability should play a fundamental role in increasing perceived usefulness (PU) and usage intentions, thus reflecting the capability of the technology to perform the promised service accurately (Parasuraman et al. 1985). In particular, we suggest that when a chatbot is able to deliver the proper service in a reliable way, its PU increases. Thus:

H4: *Reliability has a positive effect on the chatbot's PU.*

2.2.4. The effect of responsiveness on perceived usefulness

Responsiveness, another dimension of the perceived service quality (Parasuraman et al. 1985), is defined as the willingness of employees to provide a service, which involves timely responses, immediate answers and prompt service. If a provider improves responsiveness, the perceived quality of its service increases (Asubonteng et al. 1996). In line with the social exchange theory, users' perceptions of responsiveness are important antecedents of perceived usefulness-PU (Gefen and Keil 1998). Thus, we propose that in the case of chatbots, which are able to promptly answer consumers' requests, responsiveness is also an important antecedent of PU. Thus:

H5: *Responsiveness has a positive effect on the chatbot's PU.*

2.2.5. The effect of empathy on perceived usefulness and trust

Empathy is defined as the caring, individualised attention the firm provides its customers (Parasuraman et al. 1988). Some researchers have recently started to study empathy in human-chatbot service interactions. For instance, Liu and Sundar (2018) show that the expression of sympathy and empathy during the interaction with a chatbot are favoured over unemotional provisions of advice. This is in line with the "Computer-Are-Social-Actors (CASA)" paradigm that suggests users tend to expect the same social rules of human-human interactions, such as politeness and empathy, in human-computer interactions (Nass and Moon 2000). Research shows that empathy encourages information sharing between the buyers and the sellers by reducing uncertainty, thus leading to greater usefulness and trust (Kwon and Suh 2004; Aggarwal et al. 2005). This should also be true in the case of chatbots, which may be perceived as more useful (PU) and trustworthy if they express an adequate degree of empathy. Thus, we expect that when chatbots show empathy, they strengthen the relationship with the users by increasing the PU and trust:

H6: Empathy has a positive effect on the chatbot's PU.

H7: Empathy has a positive effect on trust in the chatbot.

2.2.6. The effect of credibility on trust

Drawing from previous research, our extended TAM integrates trust as an important variable, which affects the intention to use a new technology (Gefen et al. 2003, Gentry and Cantalone 2012). Trust is the is defined as the individual willingness to rely on the actions of a trustee and to depend based on the beliefs in ability, benevolence, and integrity. Users rely on trust to support their decisions to use new technologies related with a degree of uncertainty and intangibility (Gefen et al, 2003;Venkatesh et al., 2012). Trust is a crucial concept that needs to be considered when investigating transactional buyer-seller relationships, both offline and online (Gefen et al. 2003). The level of trust is an indicator of the willingness and amount of risk that one is willing to take by accepting vulnerability (Schoorman et al. 2007). Vulnerability is an important factor that defines trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" (Lee and See 2004 p.54). Trust between users and informational/transactional websites is based on perceptions of risk, ease of use and credibility (Corritore et al. 2003). Credibility comprises the objective and subjective components of the believability of a source or message. It involves trustworthiness, expertise, believability and honesty (Parasuraman et al. 1985). Credibility plays an important role in positively affecting behavioural intentions to use online services (Wang et al. 2003). Credibility is strongly linked to trust based on a partner's expertise and reliability (Wang et al. 2003). We therefore suggest that credibility has a positive effect on trust towards the chatbot. In fact, we assume that the higher the credibility of a chatbot, due to good information, impartiality, qualification and expertise, the higher the consumer's trust toward the chatbot will be.

H8: Credibility of a chatbot has a positive effect on its trust.

2.2.7. The effect of perceived ease of use and perceived utilities on the intention to reuse

In line with a large amount of literature about TAM, perceived ease of use (PEU) has a positive effect on perceived utilities (PU) (Venkatesh et al. 2012; Venkatesh and Davis 2000, Davis 1989). In the context of service interactions, we assume that if a chatbot is easy to use, the user will only have to expend minimal efforts to obtain a service. This, in turn, will increase the PU of the chatbot. Therefore, consistent with Venkatesh et al. (2012); Venkatesh and Davis (2000) and Davis (1989), we hypothesize that PEU has a positive effect on PU:

H9: PEU of a chatbot has a positive effect on its PU.

Moreover, research shows that the perceived utilities (PU) and the perceived ease of use (PEU) have a positive effect on the intention to use a technology (Venkatesh et al. 2012; Venkatesh and Davis 2000; Davis 1989). In fact, the more users believe that the chatbot enhances their performance without requiring big efforts, the more they will be keen on using it, which in turn should increase their intention to reuse. Thus, we propose:

H10: PU has a positive effect on the intention to reuse the chatbot.

H11: PEU has a positive effect on the intention to reuse the chatbot.

2.2.8. The effect of trust on the intention to reuse the chatbot

Trust is a crucial concept that needs to be considered when investigating transactional buyer-seller relationships, both offline and online (Gefen et al. 2003). Also in the context of electronic markets and in social media networks (including chatbot situations), trust is an important predictor of positive economic outcomes (Ba and Pavlou 2013), as there is absence of human interaction (Wang et al. 2003). Trust towards the e-vendor is vital for the consumer to feel protected from harmful behaviours, such as unfair pricing, inaccurate information, violations of privacy, unauthorised use of personal information and unauthorised tracking of transactions (Gefen et al. 2003). Trust plays also an important role in increasing consumers' intended use of a website. In the context of interactions with automated systems, trust is crucial in order to encourage automation use (Hoff and Bashir 2015). Moreover, as suggest-

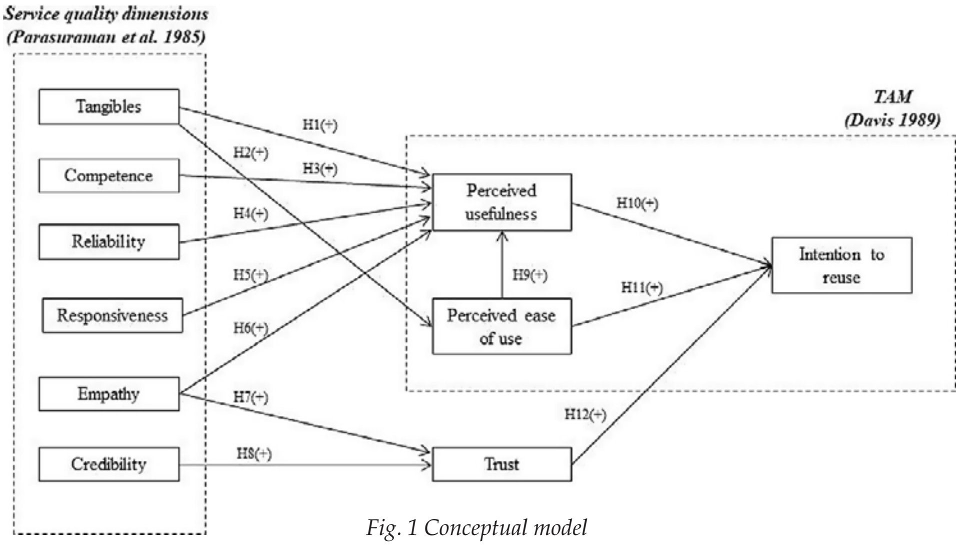

Fig. 1 Conceptual model

ed by Benbasat and Wang (2005), trust plays also a key role in increasing consumers' intentions to use online recommendation agents. Thus, we propose that in the context of chatbots, trust has a positive effect on the intention to use it again after the first interaction (Venkatesh et al. 2012):

H12: *Trust in a chatbot has a positive effect on the intention to reuse.*

All the hypotheses are formalised in the conceptual model below (*Fig. 1*).

3. Methodology

3.1. Research design

We test our model with a travel chatbot in France, called Flybot. With more than 600,000 unique users, including more than 150,000 monthly users, it brings together the largest community of travellers on Facebook Messenger. Within five minutes, Flybot is able to determine the best flights for the criteria given (distance, price, flight time). We conducted an online survey in December 2018 on Facebook and LinkedIn. The participants were asked to search for a flight ticket with Flybot on Messenger. As Flybot works solely in French, only French participants participated in the study. After the interaction with the chatbot Flybot (users had to simulate a booking request for a long haul flight Toulouse-Bangkok based on the best price criterion), questionnaires were administered to the participants. In total, 146 responses were collected. Of the sample, 73.3 % were women and 26.7 % were men. Eighty-five percent of the respondents were between 18 and 25 years old. Even if this biased sample limits generalizability to older generations, samples drawn from students are useful as this generation represents a promising market segment for new technologies (including chatbots) since they tend to be more attracted to digital technologies (McMillan and Morrison 2006; Barbosa et al. 2018; Ashraf et al. 2014).

3.2. Measures

In order to measure the construct, we adapt existing scales from the literature to the service travel chatbot context. To measure competence, we adapt scales from Sirdeshmukh et al. (2002) and van Dolen et al. (2002). To measure reliability, we adapt the scale from Tybout et al. (2005) and Park and Park (2008). To measure responsiveness, we use the scale of Gorn et al. (2004). To measure empathy, we adapt the scale of Hausman (2004). For credibility, we use the scale of Bower (2001). For tangibles, we adapt the scale from Parasuraman et al. (1985). To measure the perceived ease of use, we adapt the scale from Rauniar et al. (2014). To measure the perceived usefulness, we adapt the scale from Deshpande and Zaltman (1992). For trust, we use the scale of Crosby et al. (1990), and to measure the intention to reuse, we use the scale of Harris and Goode (2004) and Zeithaml et al. (1996). All detailed scales and items can be seen in the appendix (*Tab. A1*)

We conduct exploratory factor analyses and confirmatory factor analyses. The reliability (ρ) and convergent validity are satisfactory for each item (ρ > .7, Conv. Val. > .5; see *Tab. A2* in the appendix). The discriminant validity is satisfactory (Corr $(A, B)^2$ < Conv. Val. (A) and Corr $(A, B)^2$ < Conv. Val. (B); see *Tab. A3* in the appendix). The measurement model achieved good fit according to usual fit indices: RMSEA < .08, CFI > .90, and TLI > .90 (see *Tab. A4* in the appendix).

4. Results

In order to test the hypotheses, we conduct structural equation and mediation analysis using the software R 3.6.1. The results are summarized in table 1 below.

Hypothesis	β	p	Hypotheses
H1: Tang → PU	.14	.049	Accepted
H2: Tang → PEU	.603	.001	Accepted
H3: Comp → PU	-.31	.016	Rejected
H4: Rel → PU	1.22	.001	Accepted
H5: Resp → PU	-.08	.164	Rejected
H6: Emp → PU	.004	.962	Rejected
H7: Emp → Trust	-.07	.309	Rejected
H8: Cred → Trust	.96	.001	Accepted
H9: PEU → PU	.06	.306	Rejected
H10: PU → IRU	.63	.001	Accepted
H11: PEU → IRU	-.01	.847	Rejected
H12: Trust → IRU	.13	.452	Rejected

Notes: Tang = Tangibles; PU = Perceived Usefulness; PEU = Perceived Ease of Use; Comp = Competence; Rel = Reliability; Resp = Responsiveness; Emp = Empathy; Cred = Credibility; IRU = Intention to Reuse

Tab. 1: Results of the structural equation model

Regarding *H1*, the results show that tangibles (colours, visual appearance) have a significant and positive effect on PU (β = .14, p = .049). *H1* is thus accepted. Moreover, tangibles also have a highly significant and positive effect on PEU (β = .603, p < .001). *H2* is thus accepted. Competence has a significant negative effect on PU (β = -.31, p = .016). *H3* is thus rejected. Reliability has a highly significant and positive effect on PU (β = 1.22, p < .001). *H4* is thus accepted. Responsiveness does not have a significant effect on PU (p = .164). *H5* is thus rejected. Empathy does not have a significant effect on PU (p = .962). *H6* is thus rejected. Empathy does not have a significant effect on trust (p = .309). *H7* is thus rejected. Credibility has a highly significant and positive effect on trust (β = .96, p < .001). *H8* is thus accepted. The PEU does not have a significant effect on PU (p = .306). *H9* is thus rejected. The PU has a highly significant and positive effect on the intention to reuse (β = .63, p < .001). *H10* is thus accepted. The PEU does not have a significant effect on the intention to reuse (p = .847). *H11* is thus rejected. Trust does not have a significant effect on the intention to reuse (p = .452). *H12* is thus rejected.

4.1. Mediation analysis

In the mediation analysis, we test three different constructs: the impact of competence on the intention to reuse via PU, the impact of tangibles on the intention to reuse via PU and the impact of reliability on the intention to reuse via PU. The statistical analysis reveals that all three mediations are not significant. The links of competence, tangibles, and reliability on intention to reuse the chatbot are thus all direct and not mediated by perceived usefulness.

Mediation	B	95% confidence interval		Significant
		Lower	Upper	
Comp → PU → IRU	-.2	-21.71	21.31	No
Tang → PU → IRU	.09	-3.54	3.73	No
Rel → PU → IRU	.75	-24.36	25.87	No

Notes: Comp = Competence; PU = Perceived Usefulness; IRU = Intention to reuse; Tang = Tangibles; Rel = Reliability

Tab. 2: Results of the mediation analysis

5. Discussion of the results

By integrating SERVQUAL variables into extended Technology Acceptance Models, this study aims to understand the most relevant factors that drive chatbot acceptance and the intention to reuse.

Our results are in line with the literature (Minge et al. 2017; Mahlke 2007; Lindgaard and Dudek 2003; Hassenzahl 2005) and show two types of system or technology qualities that influence chatbot user behaviours: technology adoption and technology usage. On the one hand, we demonstrate the importance of instrumental qualities of a chatbot, such as chatbot (Davis 1989; Venkatesh and Davis 2000; Kulviwat et al. 2007; Venkatesh et al. 2012). On the other hand, non-instrumental qualities, which concern the visual aesthetics and attractiveness of the chatbot, have a significant impact on its adoption. In particular, the colours and the appearance of the chatbot have a significant positive effect on its perceived usefulness and a strong positive effect on its perceived ease of use. In line with existing research, a tangible, physical environment plays an important role in generating positive consumer evaluations of the service experience and subsequent behavioural intentions (Wakefield and Blodgett 1999). In line with previous research, our study confirms that in virtual environments, tangible elements such as aesthetics (e.g. visual elements of the interface) play an important role in positively affecting the perceived usefulness and perceived ease of use of chatbots ("Beautiful is usable") and, subsequently, in determining a consumer's intention to reuse (Hausman and Siekpe 2009). Nevertheless, the influence of perceived usefulness on the overall appraisal of the chatbot was found to be higher than that of aesthetics (Minge et al. 2017). This is in line with Minge and Thüring (2018) who show that the effect of aesthetics on product usability is strong only in the short term, either before or during the early stage of adoption of a product, and it quickly vanishes once users have become acquainted with a product.

Furthermore, our study suggests that the strongest determinant of the perceived usefulness of the chatbot is the agent's reliability. Thus, we confirm that, in the context of customer service through a chatbot, consumers expect a service to be performed accurately and in a timely manner. This result is in line with the literature (Yang and Jun 2002), which shows that, in digital contexts, reliability is the most important dimension of the perceived service quality.

In line with the literature (Wang et al. 2006), we also show that the credibility of the chatbot agent has a direct and positive effect on trust. The service agent needs to be perceived as an expert that is credible, impartial, well-informed and qualified. Nevertheless, our study does not reveal any significant effect of trust on the intention to reuse the chatbot, suggesting that the simple fact of trusting the chatbot is not enough to increase the intention to reuse it again if users perceive it to be useful for their purposes.

In contrast to the literature, responsiveness, empathy and perceived ease of use do not have any effect on perceived usefulness. In this regard, previous literature shows that in highly involved service settings (i.e. economic transactions), higher degrees of social interaction relative to functional content can be perceived in a negative way by consumers (Köhler et al. 2011). Chatbots should thus be designed to provide customers with relevant, reliable and functional content about the service, thus enhancing the customer's ability to use the firm's services. Furthermore, empathy seems not to be a relevant factor in an automated, routine interaction that is driven by economic purposes rather than socio-relational purposes (such as in the context of Flybot focusing on transactional purposes). Even with media increasingly emphasising ideas of human-robot relationships, customer relationships with service robots/chatbots seem to be a distinct phenomenon because they differ from traditional customer-firm links and put less focus on human skills, such as empathy (van Doorn et al. 2017). While chatbots are increasingly able to perform standardised tasks, as well as analyse big data, it is likely there will be some human characteristics that technology will have difficulty replacing. Namely, service contexts characterised by strong needs for empathy (e.g. those faced by professors, doctors, psychologists, social workers), where developing original and creative solutions is required (e.g. designers, engineers) or necessitates high levels of social intelligence (e.g. managerial positions), are less at risk for automation.

In contrast to the literature, a counterintuitive result is that perceived competence of the chatbot has a negative effect on the PU. Considering that chatbots are often rule-based and follow predetermined scripts, they may often fail to properly answer customers' requests. Therefore, users might not consider competence as an important factor when interacting with a chatbot. Moreover, research shows that higher intelligence and competence are generally attributed to robots that are more animate and humanlike (van Doorn 2017; Bartneck et al. 2009). Thus, the lack of anthropomorphic and humanlike cues of the chatbot used in this study may have affected this counterintuitive result.

Moreover, in contrast to the literature, our results show that the perceived ease of use of chatbots does not lead to more perceived usefulness. Thus, we suggest that consumers tend to find the chatbot useful not because of its empathy or ease of use, but because of the concrete functionality and reliability of the technology, which is considered more important to help consumers to execute tasks or receive the service they are looking for.

6. Theoretical contributions

On a theoretical level, our study offers three main contributions to service research and sheds light on consumers' perceptions and acceptance of AI-based service agents. Firstly, considering that little research in the service field has empirically investigated the usage of conversational agents (Wünderlich and Paluch 2017; Hill et al. 2015; Chung et al. 2018), this is one of the first studies where perceived service quality and acceptance have been analysed after a real experience with an existing chatbot. Thus, our empirical approach allows us to get meaningful insights related to the real usage of this technology in service settings and increases the ecological validity of our results.

Secondly, we integrate for the first time two well-established theories widely used in service research and information system research, SERVQUAL (Parasuraman et al. 1985) and Technology Acceptance Models (Venkatesh et al. 2012; Ostrom et al. 2019; Kulviwat et al. 2007; Davis et al. 1989; Davis 1989), respectively, and test them in the new, emerging context of AI-based service agents (chatbots). This approach allows for the measurement of traditional service quality dimensions in the innovative context of AI-technology based services, thus investigating both consumers' perceptions of the service quality and consumers' beliefs related to the technological components of the service.

Finally, by highlighting the relationship between usability and aesthetic in affecting customers' intention to reuse the chatbot, our results integrate and harmonise the literature between two different fields: service research (Parasuraman et al. 1985; Hausman and Siekpe 2009; Cronin and Taylor 1994; Buttle 1996; Asubonteng et al. 1996) and human-robot interactions (Minge and Thüring 2018; Mahlke 2007; Hill et al. 2015; Hartmann et al. 2007). Thus, the results of the study offer insights from a wider perspective by inviting the dialogue between these different disciplines and opening further research in related fields.

7. Managerial implications

By adopting a consumer perspective, our results also offer interesting insights to managers who want to reshape their customer service through AI and service chatbots. In particular, we have identified the factors that mostly affect the perceived service quality in customer-chatbot interactions and lead to higher acceptance and increased reuse of automated service robots. The most important criterion to increase the users' intention to reuse a chatbot is its perceived usefulness. In the case of customer experience through a chatbot, the reuse of the service is strongly driven by the perception of the chatbot's ability to efficiently perform the task and deliver the service required. Users thus have to believe in the existence of a positive use-performance-efficiency relationship by interacting with the chatbot (Venkatesh et al. 2012; Ostrom et al. 2019; Kulviwat et al. 2007; Davis 1989). In order to increase perceived usefulness and, subsequently, to have a real competitive advantage, managers should improve chatbots' efficiency and reliability to perform the promised services dependably and accurately (Parasuraman et al. 1985). Perceived credibility is also considered an important factor. In particular, in order to be trusted, the chatbot should be perceived as an expert that is well-informed and qualified. Even if the influence of a chatbot's perceived usefulness on the overall service quality judgment is higher than that of aesthetics (Minge et al. 2017), our results show that consumers give particular attention to its tangible aesthetic characteristics. In particular, colours and appearance play an important role in increasing the chatbot's perceived usefulness. Thus, we suggest that the visual elements of the chatbot need to be carefully designed in order to make the interaction and the service delivery more pleasant for the users. Nevertheless, managers have to ensure that the effect of aesthetics on the chatbot's perceived usefulness is strong only in the short term, as it quickly vanishes once users have become acquainted with a product or service (Minge and Thuring 2018).

Finally, in contrast to general assumptions, our results show that empathy does not play a major role in human-chatbot interactions. In fact, users seem to value the chatbot more for its utilitarian functions than for its socio-relational objectives. Thus, we suggest that in the case of chatbot interactions characterised by transactional purposes, the focus should be more on the reliability, accuracy of the responses and usefulness rather than on the relational aspects through empathy. Nevertheless, managers should be careful not to immediately generalise our results to other contexts, as they may have different specificities. For instance, research shows that in the case of highly involved service contexts (e.g. chatbots and conversational agents used for medical and health advice), expressions of empathy and emotional support are favoured over unemotional provisions of advice (Liu and Sundar 2018; Gelbrich et al. 2017). Thus, when implementing a conversational agent or chatbot, managers should carefully consider the context of usage and the purposes of the technology.

8. Limitations and future research directions

Our study presents some limitations that need to be addressed. First, SERVQUAL is a concise multiple-item scale with good reliability and validity that firms can use to better understand the service expectations and perceptions of consumers and, as a result, improve service. SERVQUAL has been designed as a generic measure, applicable across a broad spectrum of services. As such, it provides a basic skeleton through its expectations/perceptions format with encompassing statements for each of the service quality dimensions. However, SERVQUAL is a contentious scale that has been widely used but critiqued, as its dimensions are not universal and it is known to be unable to capture the contextual information that significantly influences users' perceptions (Yarimoglu 2014; Buttle 1996). The validity of the SERVQUAL model as a generic instrument for measuring service quality across different service sectors has also been raised. A simple revision of SERVQUAL items is not enough for measuring service quality across different service settings, as customers' assessments may vary in offline and online contexts. Furthermore, SERVQUAL focuses on the process of service delivery, not the outcomes of the service encounter. Therefore, its predictive quality is questioned. In the context of human-chatbot interactions, a more granular approach to measure users' perceptions and future intention behaviours might be required, such as E-SERVQUAL for online contexts (Parasuraman et al. 2005), SERVPERF (Cronin and Taylor 1992, 1994), the extended hierarchical model (Dagger et al. 2007; Brady and Cronin 2001;), the SNSQUAL for social network service quality (Phillips et al. 2016) or the ARTQUAL to assess aesthetic environments (Ijadi Maghsoodi et al. 2019).

Second, research suggests that humanlike (affective) nonverbal behaviour is more effective in transporting a chatbot's communicative message than robot-specific nonverbal behaviour (Rosenthal et al. 2018). Furthermore, service robots can be designed as humanoid (i.e. anthropomorphic) to simulate a human or non-humanoid appearance (Wirtz et al. 2018). Future research about this should integrate how humanlike versus robot-specific service robot designs and behaviours moderate positively or negatively the impact of SERVQUAL variables on the intention to reuse a chatbot. Thus, extending the research to include variables such as warmth and competencies (van Doorn et al. 2017) should mediate the impact of a chatbot's service production on service outcomes (e.g. service experience, satisfaction, engagement, loyalty). Then, service manipulability (e.g. the degree to which service experiences can be customised by consumers) should also be considered, as service customisation is key to success (Bitner et al. 2000).

Third, research about the impact of chatbots' versus human service providers' positive or negative outcomes, as well as the attribution of the responsibilities on satisfaction and the intention to reuse, would be a promising research outlook (Jörling et al. 2019).

Fourth, considering that Flybot can interact only in French, the sample is composed solely of French respondents. Thus, the results are not generalizable across different nationalities. Moreover, the majority of the participants are students between 18 and 25 years old. Considering that younger generations are more familiar with technologies, an older sample could present different results. For this reason, we suggest replicating the study with participants from different nationalities and different generations.

Fifth, respondents used the chatbot during a limited timeframe and only one time. Longitudinal studies are recommended, where other contexts are included and where the customers may have different needs and purposes. For instance, in the context of health services, customers may find other criteria, such as empathy, more valuable. Finally, other mediation and moderation effects could be investigated by taking into account variables that are important in the service literature, such as customer satisfaction with the service delivered by the chatbot. A longitudinal (pre-use vs. post-use) three-factor, mixed design research design would be recommended to verify if the main stream research "Beautiful is usable" holds (Tractinsky 2004; Kurosu and Kashimura 1995), or if the opposing view is valid, which claims that "Usable gets beautiful" (Minge and Thuring 2018; Tuch et al. 2010). Furthermore it could be tested if the effect of aesthetics on product usability is strong only in the short term, and if it quickly vanishes once users have become acquainted with a product.

Finally, if media increasingly emphasises ideas of human-robot relationships, future research should investigate similarities and differences between these bonds and determine if traditional theories (e.g. social exchange theory, relationship norms) can be applied to explain customer-robot relationships or if novel theories are needed (van Dorn et al. 2017). For that, the service robot acceptance model (sRAM) with social-emotional and relational elements (Wirtz et al. 2018) would be a promising research avenue.

9. Conclusion

In the future, technology will continue to play a major role in the numerous service experiences that engage customers on a social level and enable true relationships between service robots and humans. By adopting a consumer perspective, we identify the most important service quality determinants that characterise the interaction with a service robot/chatbot in the context of flight booking online. For the first time in this research field, we integrate and apply two well-known and widely accepted theoretical models, extended TAM and SERVQUAL, to the context of a chatbot in customer service production. The results suggest that customers prefer the chatbot for its usefulness and reliability. Empathy and trust do not have any significant impact. Thus, we argue that in some contexts, such as in highly involved service settings characterised by economic transactions, consumers prefer chatbots for their utilitarian value and their reliability, not for their socio-relational abilities. The results also show that the tangible elements, in particular the colours and the visual appearance, play an important role in affecting the perceived usefulness and perceived ease of use. On the other hand, the perceived ease of use does not have any effect on the intention to reuse the chatbot. The study opens the way for new potential research about customer preferences towards chatbots and offers potential insights to managers who want to introduce this technology in their customer service.

Appendix

Authors (year)	Concepts and items of scales	Cronbach alpha
Van Dolen et al. (2002) Sirdeshmukh et al. (2002)	**Competence:** A1) Flybot is efficient A2) Flybot is thorough A3) Flybot meets my needs A4) Flybot performs as I expected A5) Flybot competently handles my request	.923
Tybout et al. (2005) Park et al. (2008)	**Reliability:** B1) Flybot is useful B2) Flybot is reliable B3) Flybot gives useful information B4) Flybot gives real information	.863
Gorn et al. (2004)	**Responsiveness:** C1) Flybot responds quickly C2) Flybot responds immediately	.851
Hausman (2004)	**Empathy:** D1) Flybot is sympathetic D2) Flybot is honest D3) Flybot is attentive	.826
Bower (2001)	**Credibility:** F1) Flybot is credible F2) Flybot is impartial F3) Flybot is well-informed F4) Flybot is qualified F5) Flybot is an expert	.899
Parasuraman et al. (1985)	**Tangibles:** G1) Flybot has attractive Messenger colours G2) Flybot has attractive website colours G3) Flybot has an attractive appearance	.911
Davis (1989)	**Perceived ease of use:** H1) Flybot is adaptable H2) Flybot is understandable H3) Flybot is easy to use	.752
Davis (1989)	**Perceived usefulness:** I1) Flybot gives useful information I2) Flybot gives exact information I4) Flybot is efficient	.889
Crosby et al. (1990)	**Trust:** J1) Flybot engages me J2) Flybot puts my interests first J3) Flybot keeps its promises J4) Flybot gives perfect service quality	.899
Zeithaml et al. (1996)	**Intention to reuse:** K1) I will keep using Flybot K2) I will give positive comments on Flybot to others K3) I will recommend Flybot	.907

Tab. A1: Measurement scales and Cronbach alpha

Constructs	α	ρ	Conv. val.	Loadings
Competence	.923	.926	.714	
A1 – Efficient				.792
A2 – Thorough				.771
A3 – Meets needs				.910
A4 – Performs as expected				.874
A5 – Handles requests				.870
Reliability	.863	.869	.626	
B1 – Useful				.693
B2 – Reliable				.828
B3 – Useful information				.849
B4 – Real information				.786
Responsiveness	.851	.852	.743	
C1 – Quick				.839
C2 – Immediate				.884
Empathy	.826	.842	.643	
D1 – Sympathy				.752
D2 – Honest				.699
D3 – Attentive				.935
Credibility	.899	.899	.643	
F1– Credible				.835
F2 – Impartial				.687
F3 – Well-informed				.827
F4 – Qualified				.838
F5 – Expert				.811
Tangibles	.911	.913	.778	
G1 – Flybot Colours				.832
G2 – Website colours				.848
G3 – Flybot Appearance				.960
Perceived ease of use	.752	.769	.625	
H1 – Adaptability				.787
H2 – Understandable				.794
H3 – Easy to use				.793
Perceived usefulness	.869	.869	.678	
I1– Useful information				.852
I2 – Exact information				.840
I4 – Efficient				.764
Trust	.899	.903	.702	
J1 – Engagement				.844
J3 – Keeps promises				.914
J4 – Perfect service quality				.872
Intention to reuse	.907	.917	.789	
K1 – Keep using				.749
K2 – Positive comments				.946
K3 – Recommendation				.954

Tab. A2: Reliability (α and ρ) and convergent validity

	M	SD	Comp	Rel	Resp	Emp	Cred	Tang	PEU	PU	Trust	IRU
Comp	5.20	1.25	**.714**									
Rel	5.24	1.10	.545	**.626**								
Resp	6.14	.98	.023	.047	**.743**							
Emp	4.93	1.23	.007	.176	.095	**.643**						
Cred	4.92	1.09	.153	.605	.043	.298	**.643**					
Tang	5.87	1.01	.033	.035	.098	.132	.077	**.778**				
PEU	5.70	1.04	.269	.177	.130	.164	.258	.225	**.625**			
PU	5.16	1.19	.422	.541	.033	.219	.640	.084	.256	**.694**		
Trust	5.09	1.16	.169	.546	.033	.285	.613	.088	.271	.686	**.702**	
IRU	4.31	1.52	.303	.397	.033	.148	.408	.030	.135	.370	.376	**.789**

Notes: Notes: Comp = competence; Rel = reliability; Resp = responsiveness; Emp = empathy; Cred = credibility; Tang = tangibles; PEU = perceived ease of use; PU = perceived usefulness; IRU= intention to reuse.

Tab. A3: Discriminant validity

χ^2	DF	RMSEA	CFI	TLI
839	481	.072	.915	.901

Tab. A4: Indices of fit

References

Aggarwal, P., Castleberry, S. B., Ridnour, R., & Shepherd, C. D. (2005). Salesperson empathy and listening: Impact on relationship outcomes. *Journal of Marketing Theory and Practice*, *13*(3), 16–31. https://doi.org/10.1080/10696679.2005.11658547.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2018.03.051.

Ashraf, A. R., Thongpapanl, N. (Tek), & Auh, S. (2014). Cultural Contexts?: The Case of Online Shopping Adoption. *Journal of International Marketing*, *22*(3), 68–93.

Asubonteng, P., Mccleary, K. J., & Swan, J. E. (1996). SERVQUAL revisited: a critical review of service quality. *THE JOURNAL OF SERVICES MARKETING*, *10*(6), 62–81.

Ba, S. & Pavlou, P. A. (2013). Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Quarterly*, *26*(3), 243–268. https://doi.org/10.2307/4132332.

Barbosa, B., Filipe, S., Santos, C. A., & Simões, D. (2018). Are Millennials Ready for the Internet of Things? in "Smart Marketing With the Internet of Things", 300p, ISBN13: 9781522557630, DOI 10.4018/978-1-5225-5763-0. *In "Smart Marketing With the Internet of Things", 300p, ISBN13: 9781522557630, DOI 10.4018/978-1-5225-5763-0.*

Bartneck, C., Kanda, T., Mubin, O., & Al Mahmud, A. (2009). Does the design of a robot influence its animacy and perceived intelligence? *International Journal of Social Robotics*, *1*(2), 195–204.

Benbasat, I. & Wang, W. (2005). Trust In and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, *6*(3), 72–101. https://doi.org/10.17705/1jais.00065.

Bitner, M. J., Brown, S. W., & Meuter, M. L. (2000). Technology Infusion in Service Encounters. *Journal of the Academy of Marketing Science*, *28*(1), 138–149.

Bower, A. B. (2001). Highly attractive models in advertising and the women who loathe them: The implications of negative affect for spokesperson effectiveness. *Journal of Advertising*, *30*(3), 51–63. https://doi.org/10.1080/00913367.2001.10673645.

Brady, M. K. & Cronin Jr., J. (2001). Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach. *Journal of Marketing*, *65*, 34–49. https://doi.org/10.1007/s10840-017-0265-3.

Butler, B. S. & Gray, P. H. (2006). Reliability, Mindfulness, and Information Systems ^L tflfl I Lx_Tl lr Issues & Opinions Reliability, Mindfulness, and Information Systems1. *Source: MIS Quarterly*, *30*(2), 211–224. Retrieved from http://www.jstor.org/stable/25148728%0Ahttp://www.jstor.org/stable/25148728?seq=1&cid=pdf-reference#references_tab_contents%0Ahttp://about.jstor.org/terms.

Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, *30*(1), 8–32. https://doi.org/10.1108/03090569610105762.

Čaić, M., Odekerken-Schröder, G., & Mahr, D. (2018). Service robots: value co-creation and co-destruction in elderly care networks. *Journal of Service Management*, *29*(2), 178–205. https://doi.org/10.1108/JOSM-07-2017–0179.

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2018). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, (September), 1–9. https://doi.org/10.1016/j.jbusres.2018.10.004.

Collier, J. E. & Kimes, S. E. (2013). Only If It Is Convenient. *Journal of Service Research*, *16*(1), 39–51. https://doi.org/10.1177/1094670512458454.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*(6), 737–758.

Craig, C. S. & Douglas, S. P. (2005). *International Marketing Research Third edition* (L. John Wiley & Sons, Third Ed.).

Cronin, J. J., & Taylor, S. A. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing*, *56*(3), 55. https://doi.org/10.2307/1252296.

Cronin, J. J. & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. *Journal of Marketing*, *58*(1), 125. https://doi.org/10.2307/1252256.

Crosby, L. A., Evans, K. R., & Cowles, D. (1990). Relationship Quality in Services Selling: An Interpersonal Influence Perspective. *Journal of Marketing*, *54*(3), 68–81. https://doi.org/10.1177/002224299005400306.

Dabholkar, P. A. (1996). Consumer evaluations of new technology-based self-service options: An investigation of alternative models of service quality. *International Journal of Research in Marketing*, *13*(1), 29–51.

Dagger, T. S., Sweeney, J. C., & Johnson, L. W. (2007). A hierarchical model of health service quality: Scale development and investigation of an integrated model. *Journal of Service Research*, *10*(2), 123–142. https://doi.org/10.1177/1094670507309594.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.1016/S0305-0483(98)00028–0.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). *USER ACCEPTANCE OF COMPUTER TECHNOLOGY?: A COMPARISON OF TWO THEORETICAL MODELS *. 35*(8).

Deshpande, R. & Zaltman, G. (1992). Factors Affecting the Use of Market Research Information: A Path Analysis. *Journal of Marketing Research*, *19*(1), 14. https://doi.org/10.2307/3151527.

Evans, K. R. & Brown, S. W. (1988). Strategic options for service delivery systems. *In: C.A. Ingene and G.L. Frazier, (Eds.).Proceedings of the AMA Summer Educators' Conference (American Marketing Association, Chicago), 207- 212*. https://doi.org/10.1007/978-3-642-36172-2_200957.

Fishbein, M. & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley. – References – Scientific Research Publish. *1975*.

Forrester (2019). The Forrester New Wave¿: Conversational AI For Customer Service, Q2 2019. Retrieved June 15, 2019 from https://www.forrester.com/report/The+Forrester+New+Wave+Conversational+AI+For+Customer+Service+Q2+2019/-/E-RES144416.

Gartner (2018). Gartner Predicts a Virtual World of Exponential Change. Retrieved June 15, 2019 from https://www.gartner.com/smarterwithgartner/gartner-predicts-a-virtual-world-of-exponential-change/.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and Tam in online shopping: an integrated model. *MIS Quarterly*, *27*(1), 51–90. https://doi.org/10.2307/30036519.

Gefen, D. & Keil, M. (1998). The Impact of Developer Responsiveness on Perceptions of Usefulness and Ease of Use" An Extension of the Technology Acceptance Model. *The DATA BASE for Advances in Information Systems*, *29*(2), 35–49. Retrieved from https://wattsupwiththat.com/2011/08/05/climate-variability-in-east-africa-el-nino-southern-oscillation/.

Gelbrich, K., Hagel, J., & Orsingher, C. (2017). How Anthropomorphized Helpers Increase Customers Outcomes in Smart Service Usage. *Paper Presented at the 26th Annual Frontiers in Services Conference, New York, June*, 22–25. https://doi.org/10.1163/_q3_SIM_00374.

Gentry, L. & Calantone, R. (2002). A Comparison of Three Models to Explain Shop-Bot Use on the Web. *Psychology and Marketing*, *19*(11), 945–956. https://doi.org/10.1002/mar.10045.

Goleman, D. (1996). Emotional intelligence. Why it can matter more than IQ. *Learning*, *24*(6), 49–50. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84902967348&partnerID=40&md5=9847c270b35e3c22a65b143ed9303d9c.

Gorn, G. J., Chattopadhyay, A., Sengupta, J., & Tripathi, S. (2004). Waiting for the Web: How Screen Color Affects Time Perception. *Journal of Marketing Research*, *41*(2), 215–225. https://doi.org/10.1509/jmkr.41.2.215.28668.

Harris, L. C. & Goode, M. M. H. (2004). The four levels of loyalty and the pivotal role of trust: A study of online service dynamics. *Journal of Retailing*. https://doi.org/10.1016/j.jretai.2004.04.002.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2007). Investigating attractiveness in web user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems – CHI '07*, 387. https://doi.org/10.1145/1240624.1240687.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction*, *15*(4), 1–30. https://doi.org/10.1145/1460355.1460357.

Hassenzahl, M. (2005). Hedonic, emotional, and experiential perspectives on product quality. *In C. Ghaoui (Ed.), Encyclopedia of Human Computer Interaction. IGI Global*., 266–272. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2008/1624.

Hausman, A. (2004). Modeling the patient-physician service encounter: Improving patient outcomes. *Journal of the Academy of Marketing Science*. https://doi.org/10.1177/0092070304265627.

Hausman, A. V. & Siekpe, J. S. (2009). The effect of web interface features on consumer online purchase intentions. *Journal of Business Research*, *62*(1), 5–13. https://doi.org/10.1016/j.jbusres.2008.01.018.

Heijden, H. Van Der. (2003). Factors influencing the usage of websites: The case of a generic portal in The Netherlands. *Information and Management*, *40*, 541–549.

Hill, J., Randolph Ford, W., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2015.02.026.

Hoff, K. A. & Bashir, M. (2015). Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *57*(3), 407–434. https://doi.org/10.1177/0018720814547570.

Huang, J., Zhou, M., & Yang, D. (2007). Extracting chatbot knowledge from online discussion forums. *IJCAI International Joint Conference on Artificial Intelligence*, 423–428. https://doi.org/10.1016/j.tins.2003.11.005.

Huang, M.-H. & Rust, H. (2018). Artificial intelligence in service. *Journal of Service Research*, *21*(2), 2018. https://doi.org/10.1177/1094670517752459.

Ijadi Maghsoodi, A., Saghaei, A., & Hafezalkotob, A. (2019). ARTQUAL: A comprehensive service quality model for measuring the quality of aesthetic environments and cultural centers. *International Journal of Quality and Reliability Management*. https://doi.org/10.1108/IJQRM-01-2019–0004.

Ivanov, S. & Webster, C. (2019). Perceived Appropriateness and Intention to Use Service Robots in Tourism. In C. Springer (Ed.), *Information and Communication Technologies in Tourism*, 237–248. https://doi.org/10.1007/978-3-030-05940-8.

Jörling, M., Böhm, R., & Paluch, S. (2019). Service Robots: Drivers of Perceived Responsibility for Service Outcomes. *Journal of Service Research*, (November), 0–60. https://doi.org/10.1177/1094670519842334.

Köhler, C. F., Rohm, A. J., de Ruyter, K., & Wetzels, M. (2011). Return on Interactivity: The Impact of Online Agents on Newcomer Adjustment. *Journal of Marketing*, *75*(2), 93–108. https://doi.org/10.1509/jmkg.75.2.93.

Kulviwat, S., Bruner II, G. C., Kumar, A., Nasco, S. A., & Clark, T. (2007). Toward a unified theory of consumer acceptance technology. *Psychology & Marketing*, *24*(12), 1059–1084.

Kurosu, M. & Kashimura, K. (1995). Apparent usability vs. inherent usability: Experimental analysis on the determinants of the apparent usability. In CHI '95: *Conference companion on human factors in computing systems*. 292–293.

Kwon, I.-W. G. & Suh, T. (2004). Factors Affecting the Level of Trust and Commitment in Supply Chain Relationships. *Journal of Supply Chain Management*, *40*(2), 4–14. https://doi.org/10.1111/j.1745-493X.2004.tb00165.x.

Lee, J. D. & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *Human Factors*, *46*(1), 50–84.

Leo, J. & Chandon, P. (1997). Service encounter dimensions – a dyadic perspective?: Measuring the dimensions of service. *International Journal of Service Industry Management*, *Vol. 8* (Iss. 1), 65–86.

Lester, J., Branting, K., & Mott, B. (2004). Conversational agents: The practical handbook of internet computing. *The Practical Handbook of Internet Computing*, 220–240. https://doi.org/10.1201/9780203507223.

Li, Y. M. & Yeh, Y. S. (2010). Increasing trust in mobile commerce through design aesthetics. *Computers in Human Behavior*, *26*(4), 673–684. https://doi.org/10.1016/j.chb.2010.01.004.

Lindgaard, G. & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers*, *15*(3 SPEC.), 429–452. https://doi.org/10.1016/S0953-5438(02)00063-2.

Liu, B. & Sundar, S. S. (2018). Should Machines Express Sympathy and Empathy?? Experiments with a Health Advice Chatbot. *CYBERPSYCHOLOGY,BEHAVIOR, AND SOCIAL NETWORKING*, *21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110.

Mahlke, S. (2007). Aesthetic and Symbolic Qualities as Antecedents of Overall Judgements of Interactive Products. *People and Computers XX – Engage*, 57–64. https://doi.org/10.1007/978-1-84628-664-3_5.

McMillan, S. J. & Morrison, M. (2006). Coming of age with the internet: A qualitative exploration of how the internet has become an integral part of young people's lives. *New Media and Society*, *8*(1), 73–95. https://doi.org/10.1177/1461444806059871.

Minge, M. & Thüring, M. (2018). Hedonic and pragmatic halo effects at early stages of User Experience. *International Journal of Human Computer Studies*, *109*, 13–25. https://doi.org/10.1016/j.ijhcs.2017.07.007.

Minge, M., Thüring, M., Wagner, I., & Kuhr, C. V. (2017). The meCUE Questionnaire: A Modular Tool for Measuring User Experience. In & T. Z. A. C. Falcão (Ed.), *Advances in Ergonomics Modeling, Usability & Special Populations*, 486, 115–128. https://doi.org/10.1007/978-3-319-41685-4.

Nass, C. & Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153.

Ostrom, A. L., Fotheringham, D., & Bitner, M. J. (2019). Customer Acceptance of AI in Service Encounters: Understanding Antecedents and Consequences. In *Handbook ofService Science,: Vol. II*, 77–103. https://doi.org/10.1007/978-3-319-98512-1_5.

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Voss, C. A. (2015). Service Research Priorities in a Rapidly Changing Context. *Journal of Service Research*, *18*(2), 127–159. https://doi.org/10.1177/1094670515576315.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, *49*(4), 41–50. https://doi.org/10.1177/002224298504900403.

Parasuraman, A., Zeithaml, V.A., & Berry, L. (1988). SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *Jorunal of Retailing*, *64*(September 2014), 12–40.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL a multiple-item scale for assessing electronic service quality. *Journal of Service Research*, *7*(3), 213–233. https://doi.org/10.1177/1094670504271156.

Park, D. H. & Park, S. B. (2008). The multiple source effect of online consumer reviews on brand evaluations: Test of the risk diversification hypothesis. ACR North American Advances. *In NA – Advances in Consumer Research Volume 35, (Eds.) Angela Y. Lee and Dilip Soman, Duluth, MN?: Association for Consumer Research, Pages:*, 744–745. https://doi.org/10.1177/01461672012711012.

Phillips, B., Peak, D., & Prybutok, V. (2016). SNSQUAL: A social networking site quality model. *Quality Management Journal*, *23*(3), 19–36. https://doi.org/10.1080/10686967.2016.11918478.

Price, L. L., Arnould, E. J., & Deibler, S. L. (2006). Consumers ' emotional responses to service The influence of the service provider. *International Journal of Service Industry Management*, *6*(3), 34–63.

Rauniar, R., Rawski, G., Yang, J., & Johnson, B. (2014). Technology acceptance model (TAM) and social media usage: An empirical study on Facebook. *Journal of Enterprise Information Management*, *27*(1), 6–30. https://doi.org/10.1108/JEIM-04-2012–0011.

Rosenthal-von der Pütten, A. M., Krämer, N. C., & Herrmann, J. (2018). The Effects of Humanlike and Robot-Specific Affective Nonverbal Behavior on Perception, Emotion, and Behavior. *International Journal of Social Robotics*, *10*(5), 569–582. https://doi.org/10.1007/s12369-018-0466-7.

Rust, R. T. & Huang, M.-H. (2014). The Service Revolution and the Transformation of Marketing Science. *Marketing Science*, *33*(2), 206–221.

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An integrative lodel of organizational trust: past, present, and future. *Academy of Management Review*, *32*(2), 344–354. https://doi.org/10.5465/AMR.2007.24348410.

Seth, N., Deshmukh, S. G., & Vrat, P. (2005). Service quality models: A review. In *International Journal of Quality and Reliability Management* (Vol. 22). https://doi.org/10.1108/02656710510625211.

Sirdeshmukh, D., Singh, J., & Sabol, B. (2002). Consumer

Trust, Value, and Loyalty in Relational Exchanges. *Journal of Marketing*, *66*(1), 15–37. https://doi.org/10.1509/jmkg.66.1.15.18449.

Sonderegger, A. & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied Ergonomics*, *41*(3), 403–410. https://doi.org/10.1016/j.apergo.2009.09.002.

Sternberg, R. J. (2005). The theory of successful intelligence. *Interamerican Journal of Psychology*, *39*(2), 189–202.

Tractinsky, N. (2004). Toward the study of aesthetics in information technology. In. *Proceedings of 25th International Conference on Information Systems*, 780.

Tractinsky, N. & Lowengart, O. (2007). Web-Store Aesthetics in E- Retailing?: A Conceptual Framework and Some Theoretical Implications Web-Store Aesthetics in E- Retailing?: A Conceptual Framework and Some Theoretical l Implications. *Academy of Marketing Science Review*, *11*(1), 1–19.

Tuch, A., Bargas-Avila, J., & Opwis, K. (2010). Symmetry and aesthetics in website design: It's a man's business. *Computers in Human Behavior*, 26(6), 1831–1837.

Tybout, A. M., Sternthal, B., Malaviya, P., Bakamitsos, G. A., & Park, S. (2005). Information Accessibility as a Moderator of Judgments: The Role of Content versus Retrieval Ease. *Journal of Consumer Research*, *32*(1), 76–85. https://doi.org/10.1086/426617.

van Dolen, W., Lemmink, J., de Ruyter, K., & de Jong, A. (2002). Customer-sales employee encounters: A dyadic perspective. *Journal of Retailing*, *78*(4), 265–279. https://doi.org/10.1016/S0022-4359(02)00067–2.

van Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A. L., Grewal, D., & Petersen, J. A. (2017). Domo Arigato Mr. Roboto?: Emergence of Automated Social Presence in Organizational Frontlines and Customers ' Service Experiences. *Journal of Service Research*, *20*((1)), 43–58. https://doi.org/10.1177/1094670516679272.

Venkatesh, V. & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Venkatesh, V., Thong, J. Y. L., Chan, F. K. Y., Hu, P. J. H., & Brown, S. A. (2011). Extending the two-stage information systems continuance model: Incorporating UTAUT predictors and the role of context. *Information Systems Journal*, *21*(6), 527–555. https://doi.org/10.1111/j.1365-2575.2011.00373.x.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157–178.

Wakefield, K. L. & Blodgett, J. G. (1999). Customer Response to Intangible and Tangible Service Factors. *Psychology and Marketing*, *16*(January 1999), 51–68. https://doi.org/10.1002/(SICI)1520-6793(199901)16:1<51::AID-MAR4>3.0.CO;2–0.

Wang, Y.-S., Lin, H.-H., & Luarn, P. (2006). Predicting consumer intention to use mobile service. *Information Systems Journal*, *16*(2), 157–179. https://doi.org/10.1111/j.1365-2575.2006.00213.x.

Wang, Y. S., Wang, Y. M., Lin, H. H., & Tang, T. I. (2003). Determinants of user acceptance of Internet banking: An empirical study. In *International Journal of Service Industry Management* (Vol. 14). https://doi.org/10.1108/09564230310500192.

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management*. https://doi.org/10.1108/JOSM-04-2018–0119.

Wünderlich, N. V. & Paluch, S. (2017). A Nice and Friendly Chat with a Bot: User Perceptions of AI-Based Service Agents. *ICIS 2017: Transforming Society with Digital Innovation*, *Proceeding*, 11.

Yang, Z. & Jun, M. (2002). Consumer perception of e-service quality: from internet. *Journal of Business Strategies*, *19*(1), 19. Retrieved from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C44&q=Consumer+perception+of+e-service+quality%3A+from+internet+purchaser+and+non-purcha&btnG=.

Yarimoglu, E. K. (2014). A Review on Dimensions of Service Quality Models. *Journal of Marketing Management*, *2*(2), 79–93.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The Behavioral Consequences of Service Quality. *Journal of Marketing*, *60*, 31–46.

Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2000). A conceptual framework for understanding e-service quality: implications for future research and managerial practice. *Marketing Science Institute*, 00–115. https://doi.org/10.1509/jm.75.2.93

Keywords

Chatbot, Customer Service Robots, Artificial Intelligence, Technology Acceptance Model, SERVQUAL