



What makes you continuously use chatbot services? Evidence from chinese online travel agencies

Lin Li¹ · Kyung Young Lee² · Emmanuel Emokpae³ · Sung-Byung Yang¹ 

Received: 17 February 2020 / Accepted: 9 December 2020 / Published online: 21 January 2021
© Institute of Applied Informatics at University of Leipzig 2020

Abstract

Thanks to artificial intelligence, chatbots have been applied to many consumer-facing applications, especially to online travel agencies (OTAs). This study aims to identify five quality dimensions of chatbot services and investigate their effect on user confirmation, which in turn leads to use continuance. In addition, the moderating role of technology anxiety in the relationships between chatbot quality dimensions and post-use confirmation is examined. Survey data were gathered from 295 users of Chinese OTAs. Partial Least Square (PLS) was used to analyze measurement and structural models. Understandability, reliability, assurance, and interactivity are positively associated with post-use confirmation and technology anxiety moderates the relationships between four chatbot quality dimensions and confirmation. Confirmation is positively associated with satisfaction, which in turn influences use continuance intention. This study examines how chatbot services in OTAs are considered by users (human-like agents vs. technology-enabled services) by investigating the moderating role of technology anxiety.

Keywords Chatbot service quality · Online travel agency · Extended post-acceptance model of IS continuance · Technology anxiety · Artificial intelligence · Human-machine interaction

JEL classification Z3 · Z32

This article is part of the Topical Collection on Artificial Intelligence (AI) and Robotics in Travel, Tourism and Leisure

Responsible Editor: Marianna Sigala

✉ Sung-Byung Yang
sbyang@khu.ac.kr

Lin Li
lilin@khu.ac.kr

Kyung Young Lee
kyunglee@dal.ca

Emmanuel Emokpae
emmanuel.emokpae@dal.ca

¹ School of Management, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02447, Republic of Korea

² Rowe School of Business, Dalhousie University, 6100 University Ave, Halifax, NS B3H 4R2, Canada

³ Faculty of Computer Science, Dalhousie University, 6050 University Ave, Halifax, NS B3H 4R2, Canada

Introduction

The fundamental concept of chatbots dates as far back as 1921 when robotics was introduced. Their value became obvious in the year of 2000 when the SmarterChild intelligent agent with AOL Instant Messenger facilitated stock and weather searches (Ask et al. 2016). Gradually, this concept gained more recognition in 2011 with the launch of Apple's Siri, a voice assistant (VA) with the chatbot technology, and those VAs have been rapidly adopted in diverse fields (Johnson et al. 2012). Over the years, the hype about chatbots has developed due to the significant improvement in artificial intelligence (AI) algorithms, such as natural language processing (NLP) and machine learning (ML) technologies (Rahman et al. 2017). These AI algorithms enabled machines to better recognize information and learn from data by forecasting outcomes (Uliyar 2017). Further customization of AI-enabled chatbots is now being made to better understand human communications (either written or spoken) and to communicate back to people in

the same natural language that ought to be understood by human beings as if they were human agents (Uliyar 2017).

Chatbots therefore have been applied to various industries and more than 85% of customer interactions are expected to be replaced by chatbots by the end of the year 2020 (Wirtz et al. 2018). Especially in the tourism and hospitality industry, chatbots have been applied to facilitate services like bookings/reservations, recommendations, and other services (Nica et al. 2018; Ukpabi et al. 2019). China is one of the leading countries in this field in terms of its various chatbot-enabled applications. For example, major online travel agencies (OTAs) in China, such as [Ctrip.com](#), [Qunar.com](#), and [Figgy.com](#), have already introduced chatbot services on their websites to better help their customers book travel packages and flight tickets. It is acknowledged that when chatbots are implemented in these e-Commerce and e-Service websites, key dimensions of chatbot service quality should be identified and considered in the website design process of OTAs (Jain et al. 2018). However, it is found that chatbot quality dimensions and their effect on users' confirmation/satisfaction and intention to continue using this service are yet to be discovered in the current literature. For the quality dimensions, categorizations of quality dimensions vary by different quality assessments. For example, Parasuraman et al. (1988) proposed the quality dimensions in the context of the customer services in the telecommunications and banking industry, whereas other studies focused on organizational databases or systems. The quality dimensions thus have been less examined cohesively and comprehensively in the context of chatbot services. As to the post-acceptance model of information systems (IS) continuance (i.e., the expectation-confirmation model (ECM) of IS continuance), it has been used to examine post-use adoption of smart technologies and new online services (e.g., Kim et al. 2019; Li et al. 2020a; Park 2020), and some studies included antecedents of post-use confirmation in the extended ECM (e.g., Cheng 2014; Hong et al. 2017; Nascimento et al. 2018). However, the ECM has been rarely extended to examine users' post-use confirmation, satisfaction, and use continuance in the context of OTAs by proposing chatbot quality dimensions as antecedents, which demonstrates a research gap on the quality dimensions for the post-acceptance model of IS continuance to a larger extent, underscoring the importance of this study.

While technology-enabled consumer-facing services are prevalent in many industries, for some people, using technology devices to serve themselves is still challenging; these people prefer interacting with real human agents. In the context of travel services as well, whereas some users are quite familiar with using travel websites or mobile apps to get the travel and hospitality services they want, others still want to interact with human travel agents due to some fear or uncomfortableness of dealing with technology-enabled services. While chatbot agent services are developed by OTAs

to automate travel agents' repetitive tasks and reduce service costs, they are also meant to reduce some users' fear and uncomfortableness of using tech-enabled services by providing them with a more human-like interface than previous web/mobile-based self-services (Uliyar 2017). To examine this phenomenon of an individual's fear or uncomfortableness of using technologies (i.e., technophobia), scholars have introduced the concept of *technology anxiety* and investigated their roles in users' technology adoption (Meuter et al. 2003). Besides, there are other theoretical perspectives on *human-likeness* (i.e., *humanness*) that explain the phenomena on human reactions to the technologies which provide human-like service interfaces (e.g., Lankton et al. 2015; Mori 1970; Purington et al. 2017). We, therefore, integrate the literature on both *technology anxiety* and *humanness* and apply them in our research to investigate how users with a high level of technology anxiety consider and react to chatbot services (e.g., as either another type of technology-enabled services or non-technological human-like agents) at their post-adoption stage.

Our research purpose is threefold. First, this study attempts to identify quality dimensions of chatbot services perceived by users (i.e., understandability, reliability, responsiveness, assurance, and interactivity) which have rarely been examined in the OTA context based on IS literature. Second, drawing on the extended post-acceptance model of IS continuance (i.e., the extended expectation-confirmation model), we examine how these quality dimensions of chatbot services influence user confirmation and satisfaction, which in turn leads to use continuance intention. Third, we will examine how technology anxiety moderates the relationships between chatbot quality dimensions and post-use confirmation in order to see if high technology anxiety may lead to a stronger or a weaker relationship, which can inform us whether users consider chatbot services as human-like agents or technology-enabled services. This research would contribute to the body of knowledge on human-computer interactions with AI-driven services and the IS continuance model in the context of chatbot services of Chinese OTAs. Practically, the results of this study can provide quality assurance specialists, e-service providers, and chatbot developers with guidelines to better understand chatbot users in enhancing their service adoption in the tourism and hospitality sector.

Theoretical background

Chatbot services in the tourism industry

The rise of chatbot services could be directly proportional to rising demand for a more convenient, quick, on-demand, and less pressured self-service (Terpening and Littleton 2016), while their growth and penetration are dependent on

advancement in AI, NLP, ML technologies, and chatbot development platforms (Rahman et al. 2017). Chatbot services support both text-based and spoken interactions between humans and machines to facilitate conversation. Both forms are supposed to be designed to provide natural feelings in conversation, so that users should feel that they are having conversations with human agents but not with machines (Prasetya et al. 2018). These conversations are triggered mainly by users' inputs (e.g., questions or wake-up calls) and consist of smaller task-oriented dialogues (McTear et al. 2016). In addition, chatbot systems are equipped with contextual awareness technologies (Pearl 2016), which enable the machine to wait until it receives a message before taking its turn, by observing when a user is done typing, interpreting typed words, linguistics, and managing miscommunication (Skantze 2007).

Chatbots have been studied from several points of view. First, technical aspects of chatbots have been examined, such as the technologies for speech conversation systems (Abdul-Kader and Woods 2015), the development of chatbots using a reinforcement learning algorithm (Serban et al. 2017), and programming methods for chatbots (Long et al. 2019). Second, some research has primarily focused on human-chatbot interactions, such as how the use of chatbots can increase customer purchase (Luo et al. 2019) and how willing users are to collaborate and interact with chatbots (Ciechanowski et al. 2019). Third, as chatbot technologies are being applied to customer services, some efforts have been made to investigate the usability of the chatbot service (Kang and Kim 2017) and its impact on customer satisfaction (Chung et al. 2020). These attempts have recently been made in various industries, such as healthcare (Nadarzynski et al. 2019) and finance (Quah and Chua 2019). Especially, some empirical studies on the adoption of chatbot services have also recently been conducted in the context of instant messengers and social media (Kahiga 2019; Zarouali et al. 2018).

Among many industries that can benefit from chatbot services, we focus on the tourism industry for the following reasons. First, the tourism industry is considered one of the fields that benefit the most from chatbot services, along with the finance and retail industries (FlowXO 2020). Second, among these industries, the tourism industry seems to have the most percentage of people who are using web (online or mobile) services for their needs (over 80%), while some statistics show that the finance and retail sectors have about 70% of web service users, depending on the products or services they offer (Dubrova 2020; Milenkovic 2020; Osman 2020). Thus, as more percentage of users receive their services through online or mobile channels in the tourism industry, investigation of the role of chatbot services in the tourism context would provide valuable insights to both academicians and practitioners.

As to prior research on chatbot services in the tourism and hospitality industry, several attempts have been made only at the conceptual level. To name a few, Ukpabi et al. (2019)

integrated organizational theories to study the firm-level adoption of chatbot services in the tourism sector; Zlatanov and Popesku (2019) introduced the current applications of AI technologies including chatbots in the tourism and hospitality industry; Buhalis and Yen (2020) explained that the benefits of using chatbot services outweigh the challenges for hotels; Ivanov (2020) elaborated the influences of automation technologies (e.g., chatbot services) on tourism and hospitality jobs; and Tussyadiah (2020) pointed out the directions for future studies on automation technologies in a tourism sector. However, these studies have neither identified the quality dimensions nor empirically examined their impacts on users' adoption of chatbot services. Moreover, the capability and capacity of chatbot services in OTAs keeps accelerating. For example, it is argued by some key informants of chatbot service providers that chatbots operated by OTAs or airlines can provide faster services to travellers during peak travel seasons, with significantly improved accuracy and booking completion rates, and thus about 75% of travellers' typical post-sales inquiries, such as booking- and post-sale-related questions, could be handled by the chatbots (Gupta 2019). In sum, although more and more customers are using chatbot-enabled OTAs these days (Ivanov and Webster 2019), an attempt to verify the key quality dimensions of chatbot services and their roles in facilitating users' continuous use has rarely been made especially in the field of tourism and hospitality, which underscores the importance of this study.

Extended post-acceptance model of IS continuance

In order to answer the research question of what makes customers continue to use chatbot-enabled OTAs, the extended post-acceptance model of IS continuance is adopted in this study (Bhattacharjee 2001). This model uses theoretical underpinnings of the expectation-confirmation model (ECM) to explain the mechanism of how consumers of information technology (IT) products or services decide to continue use (or repurchase) a product they have previously adopted (purchased) (Liao et al. 2010; Oliver 1980). Key concepts embedded in this model include users' *post-use confirmation*, *perceived usefulness*, *satisfaction*, and *intention to continue to use an IS*. Briefly, this model posits that when (or before) users start to use a product (or a service), they usually have a certain level of performance expectation from the product. After a while, they can access if their expectation about the performance of the product has been met (or exceeded). If so, they perceive that their expectation from the product has been *confirmed*, which leads to their *satisfaction* with the use of the product. In addition, users' post-use assessment of the instrumentality (i.e., *usefulness*) of the product should also influence users' *satisfaction* with the product. Then, this satisfaction should eventually make users *continue to use* the product or repurchase it (Bhattacharjee 2001).

While the original ECM of IS includes perceived usefulness to explain users' post-use assessment of the instrumentality of a product, this study focuses only on users' post-use confirmation of users' initial expectations from chatbot-based OTAs and extends the ECM with key chatbot quality dimensions for post-use confirmation. The definitions in the context of this study are as follows. *Post-use confirmation* is defined as the extent to which a user's initial expectation about the performance of chatbot-based OTAs has been met. *Satisfaction* is defined as the user's positive emotional state from an appraisal of the jobs done by chatbot-based OTAs, while *use continuance intention* refers to the user's intention to continue using the chatbot-based OTAs (Bhattacharjee 2001).

For the last two decades, the ECM of IS has been applied to numerous empirical studies on the use continuance of consumer electronic products and online services. Recently, this model has been applied to quite a few empirical studies on the post-use adoption of new online services and smart technologies. For example, Park (2020) used the ECM to examine users' acceptance of smart wearable devices; Li et al. (2020a) extended the ECM to investigate users' adoption of an augmented reality game app, Pokémon GO; and Kim et al. (2019) examined users' continuous intention on accommodation apps based on the ECM. Notably, several studies that applied the ECM have extended the model by adding key antecedents for the theoretical loop of confirmation–satisfaction–use continuance. To name a few, Lee and Chen (2014) used DeLone and McLean's (2003) three quality dimensions (i.e., information, system, and service qualities) as key influencing factors for users' confirmation and found significant relationships among those quality measures and confirmation in the context of the users' continuance of m-Commerce services; Cheng (2014) identified four quality dimensions specific to online learning and examined their impacts on user's post-use confirmation, as well as the relationships among confirmation, satisfaction, and use continuance in the context of an e-Learning service for nurses; Susanto et al. (2016) extended the ECM by investigating the role of users' trust in the system, perceived privacy, and perceived security in the use continuance of smartphone banking services; Hong et al. (2017) and Nascimento et al. (2018) both investigated the use continuance of smartwatches by examining the role of users' innovativeness, habit, perceived value, perceived usability, and enjoyment, based on the ECM. These studies not only investigated the post-use adoption of smart devices, but also included antecedents of post-use confirmation (or satisfaction) in their extended ECM. Yet, little empirical effort has been made to extend the ECM and examine the factors affecting users' post-use confirmation, satisfaction, and use continuance in the context of OTAs. Therefore, this study extends the ECM by proposing five chatbot quality

dimensions, which is relevant in the context of chatbot-enabled OTAs.

Chatbot quality dimensions

To propose the quality dimensions relevant to chatbot services in OTAs, the literature on information systems (IS) and service quality was reviewed. The effort for identification of IS quality dimensions started from Zmud's (1978) article on information dimensionality, where he suggested relevance, accuracy, factuality, quantity, reliability, and readability as quality features of information, and empirically validated the measurement properties of them. Then, since customers' perception of service quality had been considered an important factor for companies to differentiate their products and services from competitors' ones (Parasuraman et al. 1988) and the topic of information quality had become more and more important for both academia and practice (Lee et al. 2002), several seminal articles on IS quality were published. Although the scope of this study is not to list all quality dimensions of those seminal articles, comprehensively review, and integrate those articles, to introduce a few, Wang and Strong (1996) proposed four categories of data quality of IS (i.e., intrinsic, contextual, representational, and accessibility) with 15 sub-dimensions in the context of an organizational database, which became an important theoretical background for the studies on IS quality including Lee et al. (2002) and Nelson et al. (2005); Parasuraman et al. (1988) proposed and empirically validated the five dimensions of service quality (i.e., tangibles, reliability, responsiveness, assurance, and empathy) in the context of customer services in the appliance repair, banking, and telecom industries; Lee et al. (2002) came up with a 2×2 model of classifying the quality dimensions that belong to the quadrants made by the products vs. services and the confirming specifications vs. customer expectations criteria and tested their measurement properties in the context of information embedded in organizational systems. Later, DeLone and McLean (2003) updated their initial articles on the IS success model (DeLone and McLean 1992) and proposed system, information, and service quality dimensions that fit the context of e-Commerce services, and Nelson et al. (2005) proposed nine dimensions of information and system quality and validated their measurement properties with data collected from various industries.

The followings are the two takeaways from reviewing these seminal articles. First, some quality dimensions belonged to multiple categories (out of information-related, system-related, or service-related). For example, *accessibility* and *reliability*, once considered either data or information qualities in some studies (Lee et al. 2002; Wang and Strong 1996; Zmud 1978), are considered either system or service qualities in other studies (Nelson et al. 2005; Parasuraman et al. 1988). It is believed that these different categorizations

of quality dimensions by different articles are due to the specificity of the target of quality assessment (i.e., different types of information/data presented, service provided, or technical aspect). Second, therefore, most of these seminal articles on IS quality proposed that the relative importance and salience of those quality dimensions should be different according to the type of IS and the use context. For example, Parasuraman et al. (1988) suggested that service quality dimensions can be adopted based on the services being investigated, DeLone and McLean (2003) mentioned that authors' selection of quality dimensions depends on the research context and focus, and Nelson et al. (2005) also suggested that the relative importance of those quality dimensions may not be applicable beyond the context of the research. These points on the selection of quality dimensions were adopted in this study and a set of quality dimensions specific to this research context were come up with: chatbot services according to the features embedded in the technology and the context of use.

Therefore, among over twenty information-, system-, and service-quality dimensions identified in the six seminal articles that have been reviewed, it is started with the five service quality dimensions (i.e., tangibles, reliability, responsiveness, assurance, and empathy) proposed by Parasuraman et al. (1988) since their research context (the customer services in the telecommunications and banking industry) is similar to the context of chatbot services of this study; the other studies focused mostly on organizational databases or systems. Three dimensions from Parasuraman et al. (1988), namely *reliability*, *responsiveness*, and *assurance*, which are believed to be salient for chatbot-related smart services, were firstly selected.

Then, as chatbots are supposed to be equipped with NLP technologies that support the contextual *understanding* of human dialogues and enable *interactive* conversations with humans, which are relatively new but important quality dimensions of smart services that the extant studies on IS quality failed to propose, recent studies on the quality dimensions of smart services were further explored. Recent studies on smart end-user technologies have suggested the *understandability* and *interactivity* as key technological quality dimensions for smart services (Cho et al. 2019; McKinney et al. 2002), which are also believed to be relevant for chatbot services because these two dimensions properly cover chatbots' capability of *understanding human dialogues* and *providing interactive conversations* when servicing users.

Five quality dimensions for chatbot services are therefore selected in the study: (1) *understandability*, (2) *reliability*, (3) *responsiveness*, (4) *assurance*, and (5) *interactivity*. The followings provide their conceptual definitions and differences among them with the selected reasons for the study. First, the perceived *understandability* of a chatbot is selected as an important quality dimension. Over the years, many researchers have identified chatbots' top priority as being accurate at

emulating human conversation. For example, a specific test designed for the quality of chatbots evaluates machines' ability to show intelligent behavior to understand human conversation (Park et al. 2018). Nguyen (2019) also tried to examine whether a chatbot could improve the consumer experience when the chatbot possesses the capability of understanding. Research also has shown that chatbot agents that understand and use humans' humor are ranked more likable, cooperative, and capable as well as provide better solutions and performance than those that do not understand it (Sensuse et al. 2019; Thies et al. 2017). The dimension of *understandability* in this study is distinct from the information understandability (or ease of understanding) as one of the information quality dimensions proposed by Lee et al. (2002) or Wang and Strong (1996). While the understandability of information in Lee et al. (2002) presents an information reader's perception that the information presented by IS is understandable, the *understandability* of a chatbot service in this study refers to user's perception that a chatbot service understands human's dialogues, the context of a conversation, and the nuance of human language.

Second, *reliability*, *responsiveness*, and *assurance* are also selected as quality dimensions of chatbot services, which have been widely acknowledged in prior literature by adapting their conceptual definitions to better fit the context of chatbot services. *Reliability* is defined as a user's perception that a chatbot service "has the ability to perform the promised service dependably and accurately" (Parasuraman et al. 1988, p.23), and it is relevant for chatbot services because providing reliable performance and information to users are considered critical when using chatbot-based services (Chung and Park 2019). AlHagbani and Khan (2016) designed a simple Arabic chatbot to find out that the reliability of chatbot services could increase the acceptance rate of chatbots in the Arabic world's online communities. Kalia et al. (2017) discovered that the reliability of chatbots can be ensured if a meaningful response is provided in a conversation. Sensuse et al. (2019) also found that chatbots that are reliable could enhance the effectiveness of job performance and motivate further development. *Responsiveness* is defined as a user's perception that a chatbot service shows "capability to help users and provide prompt service to users," while *assurance* is defined as a user's perception that a chatbot service has "knowledge and ability to inspire trust and confidence" to users (Parasuraman et al. 1988, p.23). *Responsiveness* and *assurance* are originally the service quality dimensions for human agents rather than those for technology-enabled services (DeLone and McLean 2003; Parasuraman et al. 1988). However, it is argued that these two quality dimensions are also relevant for chatbot services because chatbot-enabled service agents are supposed to act like humans in that they should provide users with (1) prompt (or *responsive*) services without delay in their responses to human's requests as well as (2) trustworthy answers to make

users feel *assured* that the services and responses they get from chatbots are not very different from those by human service agents. As to *responsiveness*, Meerschman and Verkeyn (2019) revealed that responsiveness is one of the important chatbot quality attributes and can be used for ensuring the quality of chatbots. Danilava et al. (2013) found that responsiveness could significantly improve the chatbot design in terms of creating interaction profiles. Nguyen (2019) further argued that responsiveness could significantly improve customer support chatbot systems. For *assurance*, Pereira and Díaz (2018) stated that quality assurance is a major dimension for the unsophisticated script-based conversational chatbots, and Lee and Park (2019) further discovered that assurance is a salient factor in service quality which influences user satisfaction and use intention of chatbots in the financial service industry. Most recently, Li et al. (2020b) built a chatbot, Jennifer, to examine the public information generated from reputable sources during the COVID-19 outbreak and found that the assurance quality of chatbots could be secured when the information is provided from reputable sources. However, both *responsiveness* and *assurance* have rarely been examined together in the context of chatbots, especially for OTAs, which underscores the importance of this study.

Third, the perceived *interactivity* of chatbot services can be defined as a user's perception that her/his communications with a chatbot service resemble the dialogues s/he has with human agents (with multiple times of interactions), so that s/he feels in control of personal needs when using it (Cho et al. 2019; Heeter 1989). *Interactivity* is proposed as the last quality dimension for chatbot services since it has been considered an important factor for end user-facing systems to provide personalized services and increase user engagement (Neuhof et al. 2015). Moreover, interactivity is found to be an important factor in increasing the humanness of chatbot-based systems (Go and Sundar 2019), the level of disclosure of sensitive topics (Sannon et al. 2018), and individuals' evaluations of the movie (Sundar et al. 2016). The distinction between *interactivity* and *understandability* can be explained that the former focuses on users' perception of their control over the interactive communication with chatbots while the latter focuses on the chatbot's capability of understanding the conversation. Further, the distinction of *assurance* from *interactivity* is that the assurance dimension focuses on user's feelings of after-use trust that the responses they get from chatbots are as trustworthy as those they would have received from human agents. Overall, five dimensions (*understandability*, *reliability*, *responsiveness*, *assurance*, and *interactivity*) of chatbot services capture users' perception of chatbot enabled-services' capability of (1) comprehending human's conversations and providing (2) accurate, (3) spontaneous, (4) trustworthy, and (5) interactive services. Table 1

presents a summary of the key studies on each of the five chatbot quality dimensions identified in this study.

Technology anxiety and humanness of technology

Extant studies have proposed a concept of *technology anxiety*, which entails the degree to which individuals have difficulty or fear of understanding and using technologies (Meuter et al. 2003). Based on Meuter et al. (2003), The *technology anxiety* of chatbot services is defined as a user's perception that s/he feels intimidation, unfamiliarity, and difficulty of using chatbot services. According to extant studies on the impact of technology anxiety on user reactions, those who have high technology anxiety negatively react to any new technology and avoid using it or computer-based services but seek human agents when they need any service (Meuter et al. 2003). Recent studies have provided evidence that technology anxiety (discomfort with technology or a part of technology unreadiness) is negatively associated with users' adoption or use continuance of various technologies such as e-learning, mobile health, and self-service technologies (Chen et al. 2013; Deng et al. 2014; Kotrlik and Redmann 2009). Other studies have also found that technology anxiety negatively moderates the relationships between key antecedents and the adoption of technological devices (Kim and Forsythe 2008; Lee and Yang 2013; Yang and Forney 2013). For example, Yang and Forney (2013) found that mobile shopping consumers with a low level of technology-use anxiety tend to have a high level of perceived facilitating conditions and a stronger relationship between facilitating conditions and their performance expectations from mobile shopping.

This study attempts to examine if, in the case of chatbot services, technology anxiety works similar to or different from the context of using other types of conventional technologies. As mentioned above, chatbot technologies are often applied to online (mobile) services as human-like customer service agents, so that chatbot-enabled services are supposed to somehow replace real human agents for the customers who want to be served by human agents instead of using the self-service sections in the service providers' websites or mobile apps (e.g., self-service menus for 'booking online,' 'my account,' or 'pay bill'). Thus, it is argued that individuals who have a fear of interacting with machines and prefer human agents to technology-based self-service might have more positive reactions to chatbot services (i.e., pseudo-human agents), even if they still know that the chatbot service is still an AI-based technology service, than those who are comfortable with technology-based self-service menus. In order to explain how the role of technology anxiety may differ in the case of using human-like technologies including chatbot services, three theoretical perspectives are introduced to explain the phenomenon of human reactions to technologies that mimic human behaviors.

Table 1 A summary of key literature on chatbot quality dimensions

Chatbot quality dimension	Research setting	Main findings and implications	Reference
Understandability	<p>An experiment was used to determine whether a chatbot could improve consumer experience or not.</p> <p>A specific test was designed to evaluate the machines' ability for the quality of chatbots.</p> <p>ELISA chatbot was evaluated in terms of information quality and service quality.</p> <p>Exploratory wizard-of-oz (WoZ) studies were conducted to simulate interactions with a hypothetical chatbot.</p>	<p>Understandability does little help in improving customer support systems and simple tasks should be handled by chatbots.</p> <p>Chatbots conduct intelligent behavior to understand human conversations.</p> <p>Chatbots with understandability could provide better solutions and performance.</p> <p>Chatbot agents that can understand and use humans' humor are ranked more likable, cooperative, and capable than those that cannot.</p> <p>The reliability of chatbots could increase the acceptance rate of chatbots in the Arabic world's online communities.</p> <p>The reliability of chatbots can be ensured if a meaningful response is provided in a conversation.</p>	<p>Nguyen (2019)</p> <p>Park et al. (2018)</p> <p>Sensuse et al. (2019)</p> <p>Thies et al. (2017)</p> <p>AlHagbani and Khan (2016)</p>
Reliability	<p>A simple Arabic chatbot was designed to study the chatbot applications and their challenges in chatbot development.</p> <p>A new methodology, Quark was used to develop human-chatbot interactions in a typical management process.</p> <p>ELISA chatbot was evaluated in terms of information quality and service quality.</p>	<p>Chatbots that are reliable could enhance the effectiveness of job performance and motivate further development.</p> <p>Responsiveness could significantly improve chatbot design quality in terms of creating interaction profiles.</p>	<p>Kalia et al. (2017)</p> <p>Sensuse et al. (2019)</p>
Responsiveness	<p>Text-based conversational chatbots were tested in human instant messaging dialogues for the development of a computational model.</p> <p>A chatbot-focused quality model was developed by extracting extant literature and conducting expert interviews.</p> <p>An experiment was conducted to determine if a chatbot could improve the consumer experience.</p>	<p>Chatbots that are reliable could enhance the effectiveness of job performance and motivate further development.</p> <p>Responsiveness could significantly improve chatbot design quality in terms of creating interaction profiles.</p>	<p>Danilava et al. (2013)</p>
Assurance	<p>A survey was used to explore the factors that influence user satisfaction, which leads to the use intention of chatbots in the context of financial service.</p> <p>A chatbot, Jennifer, was built to examine if the public information generated from reputable sources are better organized and shared during the COVID-19 outbreak.</p>	<p>Responsiveness is one of the chatbot quality attributes and can be used for ensuring the quality of chatbots.</p> <p>Responsiveness could significantly improve customer support systems.</p> <p>Assurance is a salient factor in service quality that influences user satisfaction and use intention of chatbots in financial services.</p> <p>Assurance of quality could be secured when the information is provided from reputable sources and information quality should be maintained in a rigorous process in an empathetic manner.</p>	<p>Meerschman and Verkeyn (2019)</p> <p>Nguyen (2019)</p> <p>Lee and Park (2019)</p> <p>Li et al. (2020b)</p>
Interactivity	<p>Classifications were conducted to determine the key dimensions for the success of chatbots.</p> <p>A 2x2x2 experiment was conducted to explore the interactivity with chatbots in choosing a digital camera on an e-Commerce website.</p> <p>An experimental survey was designed to explore the impact of interactivity with conversational agents on stress-related disclosure.</p> <p>A five-condition, between-participants lab experiment on an online movie search website was conducted to find the effects of message interactivity on perceived contingency and user engagement.</p>	<p>Quality assurance is a major dimension for unsophisticated script-based conversational chatbots.</p> <p>Interactivity is an important factor in increasing the humanness of chatbot-based systems.</p> <p>The level of interactions with conversational agents could affect the level of disclosure of sensitive topics.</p> <p>The level of message interactivity including interaction history and chatbot function greatly influences individuals' evaluations of the movie website.</p>	<p>Pereira and Diaz (2018)</p> <p>Go and Sundar (2019)</p> <p>Samon et al. (2018)</p> <p>Sundar et al. (2016)</p>

The first theory on human reactions to human-like machines is called *uncanny valley theory* (UVT) (Mori 1970), which posits that an individual's reaction to a human-like machine first shows a positive relationship with the human-likeness of the machine to a certain point (e.g., industrial robot arms for assembling parts or AI-driven chatbots with human-like voices or sentences), but this positive reaction becomes abruptly negative as the machine comes with some atypical or imperfect human looks and behaviors and fails to show an appropriate level of appearance like humans either in reality or experimental settings (e.g., creepy looking humanoids or chatbots with human-looking animated avatars) (Ciechanowski et al. 2019; Kätsyri et al. 2015; Urgen et al. 2018). This negative reaction can be positive again as the human-likeness of the machine becomes near-perfect, making the human reaction chart against the human-likeness of the machine a valley-like shape (Mori 1970; Urgen et al. 2018). Some studies have shown empirical evidence of this UVT by confirming the sequence of positive-negative-positive relationships between human reactions and the machine's human-likeness in certain cases (e.g., Kätsyri et al. 2015). In the present study, however, the first positive and linear association between users' reaction and machines' (chatbots') human-likeness (before the 'valley' part of the theory) was focused on, as the features added to the online chatbot services in the context of the present study (i.e., chatbots on OTAs), supporting interactive conversations with users via text or voice channels with no likeness of human appearance, can be seen as ones of typical entry-level human-like features of machines (in terms of the x-axis of the UVT curve) without any atypical or imperfect human-likeness, not going beyond the point that users feel eerie or uneasy. Moreover, a recent study found that people tend to experience less uncanny effect when using simpler text-based chatbots than when using chatbots serving with animated (obviously robot-looking) avatars (Ciechanowski et al. 2019). As such, it is believed that the human-like interactive and contextual language processing capability of chatbot services may positively influence the way users react to the chatbot technology.

Second, the literature on *technology humanness* is introduced, referred to as users' perceived similarity of a technology device to humans in their motions (behaviors) and physical appearances, which is operationalized on a continuum from 'system-like' to 'human-like' (Kamide et al. 2014; Lankton et al. 2015). The literature on technology humanness has examined the relationship between users' familiarity with a technology device and humanness of it (Kamide et al. 2014), the impact of human-like trust in a device on users' adoption (Carter and Liu 2018), and the relationship among users' perceived technology-driven social presence, perceived humanness, perceived usefulness, and the enjoyment of a technology device (Lankton et al. 2015). Overall takeaways from these

studies are that individuals' perception of the humanness of the same technology devices varies due to their different perception of social presence enabled or provided by the device (Lankton et al. 2015), and users' perceived humanness of a device is positively associated with users' trust in and emotional reactions to the device, which leads to the adoption (Kamide et al. 2014; Lankton et al. 2015). Among several theories used to explain users' perceived technology humanness, the theory of social presence is focused on (Rettie 2003; Short et al. 1976), since the social presence of a technology device, which refers to users' perception of interpersonal interactions with the device, is related to the interactive nature of technology-based services such as social media and chatbot services (Lankton et al. 2015).

The last theoretical perspective is the *personification* of machines by humans. Recent studies have investigated humans' tendency to personify human-like technologies, such as smart voice assistant systems (Lee et al. 2019; Purington et al. 2017). In addition to feeling some degree of the humanness of a technology device, people start to treat the device as a human companion, although they are well aware that the human-like functions of the device are 100% realized by technological artifacts such as AI and NLP technologies (Lopatovska and Williams 2018). This theoretical perspective also suggests that the degree of personification of a machine varies among different individuals and key technological features that trigger users' personification of a machine are interactive and conversational functions embedded in a technology (Purington et al. 2017). Some key findings from this literature related to our study are that people are generally more comfortable interacting with machines that can mimic what humans do (e.g., human-like conversations), and their degree of personification is related to users' satisfaction with the machine (Purington et al. 2017).

All these theories imply that some technology devices can be no longer seen as 100% technology devices or machines but can be, at least, partially personified by users, depending on how humans see them. Therefore, the role of technology anxiety on users' post-use assessment of chatbot services should be different from conventional technology-based online services, because chatbot services come with human-like features. To examine the role of technology anxiety in chatbots, this study proposes technology anxiety as a key moderating factor for the relationships between chatbot quality dimensions and users' post-use confirmation. Depending on how users see the chatbot-enabled services with different levels of technology anxiety, either as human-like agents or as another new type of technology-enabled self-services, the way they assess their post-use confirmation against the service quality dimensions should be different. A more detailed argument about the moderating role of technology anxiety will be elaborated in the next chapter.

Research model and hypotheses development

Figure 1 illustrates our research model. Briefly, chatbot quality dimensions are positively associated with users' post-use confirmation, and technology anxiety positively moderates the relationships between chatbot quality dimensions and confirmation. Confirmation is positively associated with user satisfaction, which is also positively associated with the use continuance intention of chatbot services.

The relationships between chatbot quality dimensions and post-use confirmation

A chatbot needs to have a robust database of information that allows it to generate suitable and most appropriate responses to users, so its NLP system converts human languages to relevant information whilst responding to users' inputs with relevant information (Hill et al. 2015). This gives it human-like characteristics, which may eliminate initial distrust users often have towards computer-based systems (Zamora 2017). A number of recent studies on chatbots have verified a positive relationship between users' understandability and user's positive reactions. For example, Kuligowska (2015) highlighted the ability of understanding human's conversation by commercial chatbot services as one of the important evaluation criteria for chatbots. If users find that most information provided by a chatbot is understandable, they are more likely to confirm that their initial expectation of the performance has been met. Therefore, we hypothesize that:

H1: Understandability of a chatbot service is positively associated with user's post-use confirmation.

Reliability is considered one of the important factors for information service functions provided by customer-facing IS services and found to be an important predictor of user satisfaction (Kettinger and Lee 1994). In the context of

chatbots, the reliability of the information provided by chatbot services is regarded as an important factor for users (Chung and Park 2019). As such, if chatbot-based OTAs provide dependable services with accurate information (Parasuraman et al. 1988), current users will be more likely to think that their initial expectation of chatbot performance has been confirmed. Therefore, we hypothesize that:

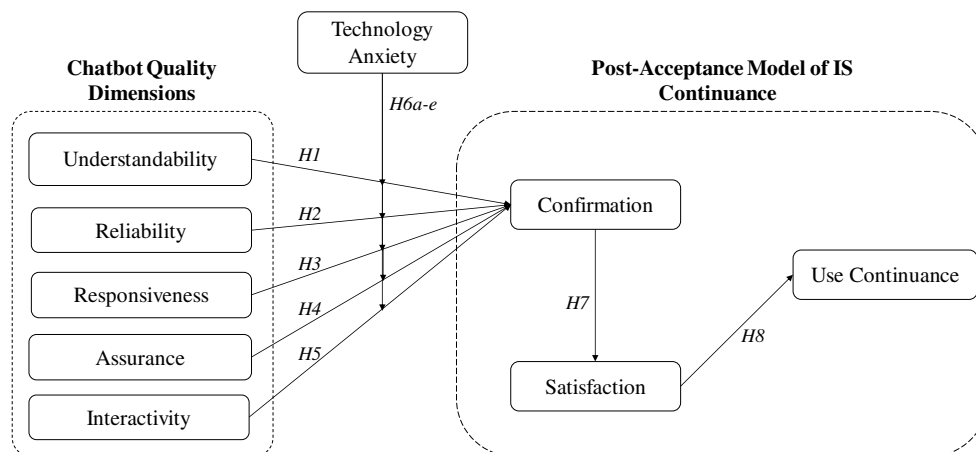
H2: Reliability of a chatbot service is positively associated with user's post-use confirmation.

Responsiveness and assurance are originally the service quality dimensions for human agents rather than those for IT or IS (DeLone and McLean 2003; Parasuraman et al. 1988). However, we argue that these two quality dimensions are also relevant for chatbot services because chatbot-enabled online agents are supposed to act just like human agents. Thus, we define responsiveness as the extent to which a chatbot service shows a willingness to help and provides prompt services to users. If users perceive that their OTA chatbot answers promptly to their questions and provides immediate services for travel booking and destination recommendations, then it is more likely that they will confirm their initial expectation for using the service. Therefore, we hypothesize that:

H3: Responsiveness of a chatbot service is positively associated with user's post-use confirmation.

As previously defined, users perceive assurance from a chatbot if the chatbot service has "knowledge and ability to inspire trust and confidence" to users (Parasuraman et al. 1988, p.23). If users feel that the information given by an OTA chatbot and the transaction they do with the chatbot are trustworthy enough for them to rely on, and they are also confident that those information and transactions are as good as those given or handled by human agents (or other means of online tools such as web-based hotel/travel bookings), they

Fig. 1 Research model



will find that their initial expectation with the chatbot service will be better confirmed. Therefore, we posit that:

- H4: Assurance of a chatbot service is positively associated with user's post-use confirmation.

For interactivity, we found some evidence that interactivity plays an important role in users' reactions to smart services. Cho et al. (2019) suggested that the interactivity of smart wearable devices is an important influencing factor for users' positive reactions to the use of the device. Moreover, Shin et al. (2013) found that the interactivity of smart TVs improves users' positive attitude towards the device. Since a chatbot service is also a kind of smart services, a high degree of interactivity should influence users' positive reactions, positively influencing their evaluation of their post-use confirmation. Thus, if users of OTA chatbot services find that the chatbot shows a seamless interaction without delay and errors and makes them engaged in the conversation just like they are having a conversation with human agents, they will evaluate that their initial expectation for using the service is confirmed. Therefore, we hypothesize that:

- H5: Interactivity of a chatbot service is positively associated with user's post-use confirmation.

The moderating role of technology anxiety

As discussed above, users may see the chatbot service as either 'a human-like service' that can appropriately replace human's customer service work or 'another type of technology-enabled services.' If users see chatbot services as human-like agents, we expect that those with a high level of technology anxiety would more positively influence the relationships between chatbot quality dimensions and post-use confirmation of their initial expectation. But, if they treat chatbot services as just one type of technology-driven services, which is not different from other self-service technologies, the relationships between chatbot quality dimensions and their assessment of their post-use confirmation would be negatively affected by their anxiety to the chatbot technology.

Technology anxiety has been used as a moderating factor in the studies on the adoption of consumer electronics. To list a few, Yang and Forney (2013) found that technology anxiety mitigates the relationship between facilitating conditions of mobile shopping use and performance expectancy, and it also moderates the relationship between social influence and intention to use mobile shopping. In addition, Kim and Forsythe (2008) found that users' attitude towards a virtual try-on technology is more strongly related to the use of the technology when the users have a low level of technology anxiety. Finally, Lee and Yang (2013) found that consumers'

perceived interpersonal service quality (against self-service technology) in retail stores and their patronage intention is negatively moderated by technology anxiety. As such, it is found that technology anxiety in general negatively moderates the relationships between user's perception of service quality/facilitating conditions/attitude and their adoption/post-use reactions.

To the best of our knowledge, however, there has been no empirical study that examines the moderating effect of technology anxiety on the relationships between chatbot service quality dimensions and users' post-use confirmation. Moreover, a recent study on technology anxiety in the context of chatbot users did not find a significant relationship between technology anxiety and intention to use chatbots (Lee and Park 2019). Due to its double-sided nature of being considered either 'pseudo-human agents' or 'technology-driven services,' it is still uncertain that users' perceived technology anxiety has a positive or a negative moderating role in the relationships between chatbot service quality dimensions and their post-use confirmation. However, based on the positive relationships between the human likeness of a technology device and users' trust in the device as well as users' feeling of social presence while using human-like chatbot services (Kamide et al. 2014; Lankton et al. 2015), we argue that technology anxiety may have a positive moderating role in the relationships between chatbot service quality dimensions and users' post-use confirmation. In other words, a user who has a high level of technology anxiety towards chatbots might not have had a very high initial expectation from chatbot services, but when s/he finds that the services provided by chatbots are understandable, reliable, responsive, trustworthy, and interactive, their post-use confirmation become stronger than those who have less technology anxiety towards chatbots due to their higher perception of human-likeness of technology (Purington et al. 2017). Therefore, we propose that:

- H6: Technology anxiety positively moderates the relationships between service quality dimensions (understandability, reliability, responsiveness, assurance, and interactivity) and post-use confirmation of a chatbot service.

The relationships among confirmation, satisfaction, and use continuance

Ever since Bhattacharjee (2001) proposed and tested the post-acceptance model of IS continuance based on the ECM, many studies have found evidence of significant relationships among confirmation, user satisfaction, and use continuance. Recent studies have also used this model and provided empirical supports for the theoretical relationships among these three variables in the context of smart consumer technologies. To name a few, Susanto et al. (2016) found that

confirmation is positively related to user satisfaction, and satisfaction is positively associated with continuance use intention in the context of smartphone banking. Other studies also found significant relationships among confirmation, satisfaction, and use continuance intention as presented in the post-acceptance model of IS continuance in various contexts such as mobile health services, offline-to-online (O2O) services, and smartwatches (Akter et al. 2013; Hsu and Lin 2019; Nascimento et al. 2018). Likewise, we argue that these findings can also be applied to the case of a chatbot service. If current chatbot users find that the chatbot service has met their initial expectation of using it, they will be satisfied with the service. Moreover, if they are satisfied with a chatbot service after a while of using it, then they are more inclined to continue using the service. Thus, we posit that:

- H7: User's Post-use confirmation of a chatbot service is positively associated with satisfaction.
- H8: User's satisfaction with a chatbot service is positively associated with the use continuance intention.

Research methodology

Measurements

This study collected survey data from the users of OTAs in China with seven-point Likert scales. Table 2 provides the operational definitions of variables used in our research model with referred studies. Our survey items were either adapted

from the literature or developed based on the conceptual definitions of extant studies.

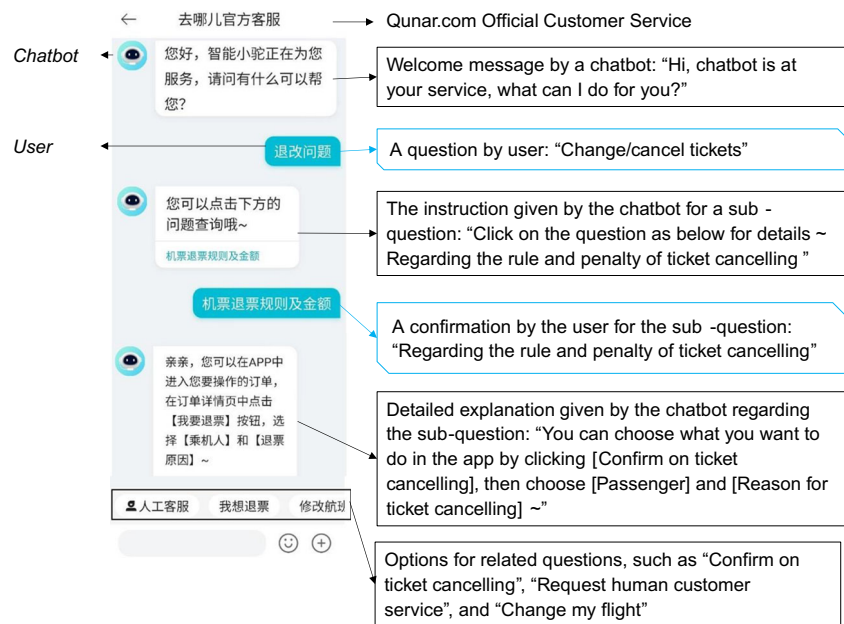
Description of the research site and data collection

Three major OTAs in China were selected as our research sites: Ctrip.com, Qunar.com, and Fliggy.com. The reasons for choosing these three OTAs are as follows. First, the Chinese market is chosen because China has become the world's second-largest OTA market in 2018, and it has grown aggressively with a 27% annual growth rate, more than four times of the U.S.'s growth rate over 2017. Moreover, it is estimated that China's OTA market will be very close to the size of the U.S.'s one by 2022 (Phocuswright Research 2019). Second, Ctrip.com, Qunar.com, and Fliggy.com are the three major OTA chatbot service providers in China by far (Zou 2019) and users are becoming familiar with their chatbot services. Specifically, these three OTAs cover 74.1% of the total market share in China in the first half of 2019 (Zou 2019), which makes them dominate the OTA market in China. Thus, a large target population can be guaranteed with the reliability and relevancy of this study. Third, in order to accurately validate users' reactions to chatbot services, it is important for chatbot users to be aware that the service is enabled by chatbot technologies rather than human agents. Unlike the users of some chatbot services who are not well informed that the service is done by chatbots, the users of Ctrip.com, Qunar.com, and Fliggy.com are well aware that services are enabled by chatbots, not by human agents. When users log on to these OTA websites/apps and look for customer services, the options of human and chatbot services are clearly distinguishable and shown with different icons. When users

Table 2 Operational definitions

Construct	Operational definition	References
Understandability	The extent to which users perceive that a chatbot service understands human's dialogues, the context of a conversation, and the nuance of human language	Park et al. (2018); Thies et al. (2017)
Reliability	The extent to which users perceive that a chatbot service has the ability to perform the promised service dependably and accurately	Parasuraman et al. (1988)
Responsiveness	The extent to which users perceive that a chatbot service shows a willingness to help users and provides prompt services to users	Parasuraman et al. (1988)
Assurance	The extent to which users perceive that a chatbot service has knowledge and ability to inspire trust and confidence to users	Parasuraman et al. (1988)
Interactivity	The extent to which users perceive that their communication with a chatbot service resembles the dialogues they have with human agents (with multiple times of interactions), so that feel in control of their personal needs when using it	Cho et al. (2019); Heeter (1989)
Technology anxiety	The extent to which users feel intimidation, unfamiliarity, and difficulty of using a chatbot service	Meuter et al. (2003)
Confirmation	The extent to which users' initial expectation about the performance of chatbot-based OTAs has been met	Bhattacharjee (2001)
Satisfaction	The extent to which users perceive that their positive emotional state comes from an appraisal of the jobs done by chatbot-based OTAs	Bhattacharjee (2001)
Use continuance	The extent to which users perceive that they intend to continue using the chatbot-based OTAs	Bhattacharjee (2001)

Fig. 2 A user's typical exchange of information in a conversation with a chatbot in OTAs



intend to use chatbot services and click the icon of them, a message pops up indicating that the service is provided by chatbots and asks for customers' questions using a chatbot profile image.

The typical exchange of information in a conversation between a user and a chatbot is illustrated in Fig. 2. To be specific, a chatbot welcomes a user with a message as a starter and different kinds of questions are given as options. After the user selects a specific category, more detailed instructions for a sub-question are provided to guide the user, such as what products can be taken on a plane and how to get refunded a ticket. Typically, a conversation would be completed in three to four turns, and each turn includes a request from an OTA user and a response from the chatbot; if the user is a VIP (very important person) user, her/his simple text of "buy me a ticket to Beijing," for example, could result in a direct issue of a flight ticket that matches her/his preference, as previous conversations can be used as a reference by chatbots (Gupta 2019).

As to our survey respondents, they were acknowledged that the customer services are done by chatbots on the chatbot-enabled OTA. In addition, the screening question of "have you used the chatbot services in major travel sites?" was asked to make sure our survey participants were actual users of the chatbot-enabled services. Furthermore, survey participants were required to think about chatbot services while answering the questions. Because only the actual chatbot users were allowed to participate in the survey and chatbot services were solely pinpointed specifically for our research setting, we thus controlled survey respondents' knowledge that the services were provided by chatbots.

The *So Jump* (<http://www.sojump.com>) survey platform was chosen to collect questionnaires, as it is one of the

largest and the most commonly used online professional questionnaire platforms to conduct surveys in China (Liu et al. 2019). It allows participants to answer questionnaires easily on either mobile phones or websites. The survey was conducted in October 2019 for a duration of two weeks, and 326 survey responses were collected. A total of 295 responses were used for our data analysis after removing 31 responses due to their incompleteness and aberration with singular answers (Meijer and Sijtsma 1995). The demographic information of respondents presented in Table 3. The questionnaire items for the constructs used in this study are presented in the Appendix Table 8.

Results

Measurement model

A confirmatory factor analysis (CFA) using SmartPLS 2.0 was conducted to test our measurement model. As shown in Table 4, composite reliability (CR) and Cronbach's α values were used to measure internal reliability. CR and Cronbach's α values of all the constructs were more than the acceptable threshold values of 0.6 (Fornell and Larcker 1981). As to convergent validity, it is measured by factor loadings and average variance extracted (AVE) values. Factor loading scores ranged from 0.521 to 0.869, which surpassed the acceptable threshold value of 0.5 (Bagozzi et al. 1991), and AVE values ranged from 0.483 to 0.700, which exceeded the recommended value of 0.5 (Fornell and Larcker 1981) except for technology anxiety (0.483), support the presence of convergent validity, considering that AVE values could be a relatively conservative estimation of the convergent validity

Table 3 Demographics of respondents

Demographics	Frequency	Percent
Gender		
Female	125	42.4%
Male	170	57.6%
Age		
< 16	15	5.1%
17–21	156	52.9%
22–30	100	33.9%
31–45	19	6.4%
46–64	4	1.4%
> 65	1	0.3%
Education		
High school	8	2.7%
Studying at the undergraduate level	23	7.8%
Bachelor's degree	217	73.6%
Studying at the graduate level	38	12.9%
Master's degree and above	9	3.1%
Income (Renminbi: RMB, Chinese Yuan)		
< 3000 RMB	82	27.8%
3000–5000 RMB	112	37.9%
5000–7000 RMB	61	20.7%
7000–9000 RMB	24	8.1%
9000–11,000 RMB	7	2.4%
> 11,000 RMB	9	3.1%
Occupation		
Student	44	14.9%
Working	238	80.7%
Unemployed	7	2.4%
Other	6	2.0%
Usage frequency (indicating how often do users use chatbot services)		
Less than once a month	63	21.4%
Once to twice a month	172	58.3%
Once to twice a week	34	11.5%
Three to four times a week	13	4.4%
More than five times a week	2	0.7%
Other	11	3.7%
Average time per use (indicating, on average, how long do users use chatbot services)		
Less than 5 min	21	7.1%
5 to 10 min	108	36.6%
10 to 20 min	95	32.2%
20 to 30 min	46	15.6%
30 min to 1 h	19	6.4%
More than 1 h	3	1%
Other	3	1%
Total	295	100%

(Fornell and Larcker 1981). In addition, discriminant validity was accessed by comparing the square root values of AVE

with the inter-construct correlation coefficients (Fornell and Larcker 1981). As shown in Table 5, all the square roots values of AVE, highlighted in bolds, were larger than those of inter-construct correlation coefficients, indicating that discriminant validity is adequately achieved.

Since our survey data collection is cross-sectional and self-reported, it could have a risk of common method bias (CMB). To examine the CMB of the study, we conducted several tests. First, following the suggestion of Podsakoff et al. (2003), Harman's one-factor test was conducted by including all items in a principal component exploratory factor analysis without rotation. Based on the technique of Podsakoff et al. (2003), if one factor is detected to explain most of the covariance among the measurements, the threat of common method bias is high. Our test results showed that the dominant factor only accounts for 39.8% of the total variance, which is far below the threshold of 50% (Podsakoff et al. 2003), indicating that the threat of CMB is not serious in this study. Second, based on the recommendation of Liang et al. (2007), a common method factor that contains indicators of all the principle constructs was developed and variances of all indicators related to principal constructs and the common method factor were calculated. As shown in Table 6, our results suggested that the average of the square of substantive factor loadings ($R^2 = 0.6415$) was much larger than that of the square of common method factor loadings ($R^2 = 0.0176$), indicating that CMB is not a concern in this study. Third, the correlation matrix (shown in Table 5) indicated that the correlation values are smaller than 0.757 ($r < 0.757$), which does not imply highly correlated variables, whereas CMB emerges when extremely high correlations exist ($r > 0.90$) (Pavlou et al. 2007). The above three tests consistently suggest that this study is robust against the CMB concern.

Structural model

A partial least squares (PLS) method was used to test our hypotheses using SmartPLS 2.0. PLS is appropriate to use for a multi-path model when the sample size is small (Chin 1998). A bootstrapping technique and a PLS algorithm are used to test whether our hypothesized relationships are positive or negative and significant or not. Results of the structural model are shown in Fig. 3 and Table 7, including explained variances of endogenous variables (R^2), the path-coefficients (β), the level of significance (p value) based on t -values, and the effect sizes (ES s) of moderating effects.

As shown in Fig. 3 and Table 7, six out of the seven main hypothesized paths were supported except for $H3$, indicating the relationship between responsiveness and confirmation is not significant ($\beta = 0.072$, t -value = 1.173). Among the other dimensions of chatbot quality, understandability was positively related to confirmation ($\beta = 0.145$, t -value = 1.744),

Table 4 Reliability and convergent validity

Construct	Factor loading	Cronbach's α	Composite reliability (CR)	Average variance extracted (AVE)
Understandability (UN)	0.866 0.862 0.791	0.785	0.875	0.700
Reliability (REL)	0.821 0.843 0.836	0.780	0.872	0.695
Responsiveness (RES)	0.804 0.710 0.797	0.674	0.815	0.595
Assurance (ASS)	0.804 0.828 0.806	0.743	0.854	0.660
Interactivity (INT)	0.818 0.824 0.810	0.825	0.884	0.657
Confirmation (CF)	0.821 0.869 0.796	0.773	0.868	0.688
Satisfaction (SAT)	0.858 0.746 0.837 0.758	0.812	0.877	0.642
Use continuance (UC)	0.815 0.786 0.852 0.786	0.825	0.884	0.657
Technology anxiety (TA)	0.720 0.800 0.710 0.521	0.670	0.786	0.483

marginally supporting *H1*, and reliability was also positively related to confirmation ($\beta = 0.261$, t -value = 2.732), supporting *H2*. So are the relationships between assurance and confirmation ($\beta = 0.159$, t -value = 1.972) and between

interactivity and confirmation ($\beta = 0.156$, t -value = 2.025), supporting both *H4* and *H5*. As expected, confirmation was positively associated with satisfaction ($\beta = 0.757$, t -value = 24.864) and satisfaction was positively associated with use

Table 5 Construct correlations and discriminant validity

Construct	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
UN (1)	0.837								
REL (2)	0.754	0.834							
RES (3)	0.674	0.678	0.771						
ASS (4)	0.736	0.712	0.666	0.812					
INT (5)	0.690	0.752	0.669	0.724	0.811				
CF (6)	0.615	0.649	0.557	0.612	0.615	0.829			
SAT (7)	0.678	0.689	0.585	0.637	0.659	0.757	0.801		
UC (8)	0.608	0.650	0.552	0.586	0.594	0.637	0.746	0.811	
TA (9)	-0.440	-0.430	-0.528	-0.392	-0.375	-0.291	-0.321	-0.327	0.695

Table 6 Results of common method bias test

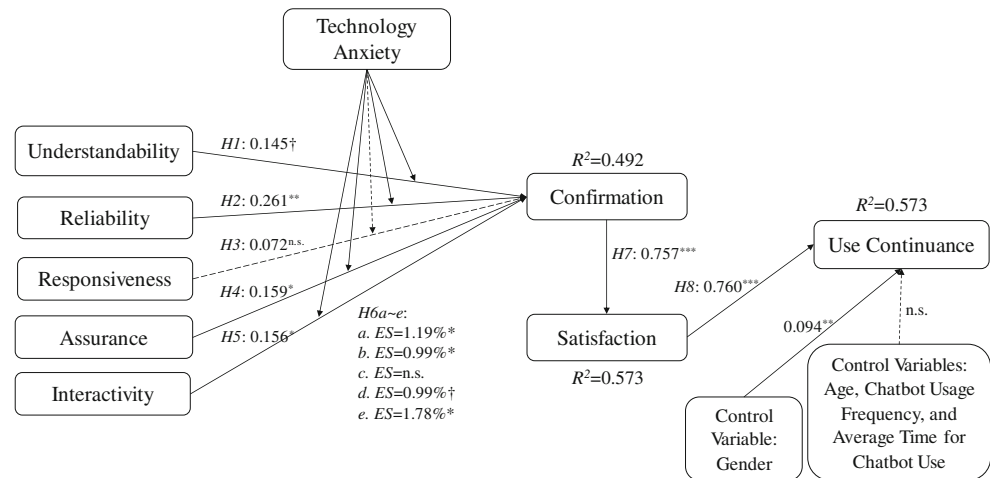
Construct	Indicator	Substantive factor loading (<i>R1</i>)	<i>R1</i> ²	Common method factor loading (<i>R2</i>)	<i>R2</i> ²
UN	UN1	0.849	0.7208	0.008	0.0001
	UN2	0.968	0.9370	-0.115	0.0132
	UN3	0.684	0.4679	0.115	0.0132
REL	REL1	0.796	0.6336	0.033	0.0011
	REL2	0.763	0.5822	0.089	0.0079
	REL3	0.943	0.8892	-0.125	0.0156
RES	RES1	0.837	0.7006	-0.001	0.0000
	RES2	0.972	0.9448	-0.374	0.1399
	RES3	0.439	0.1927	0.354	0.1253
ASS	ASS1	0.632	0.3994	0.188	0.0353
	ASS2	0.818	0.6691	0.018	0.0003
	ASS3	0.992	0.9841	-0.210	0.0441
INT	INT1	0.840	0.7056	-0.042	0.0018
	INT2	0.824	0.6790	0.004	0.0000
	INT3	0.790	0.6241	0.037	0.0014
CF	CF1	0.910	0.8281	-0.104	0.0108
	CF2	0.797	0.6352	0.081	0.0066
	CF3	0.784	0.6147	0.019	0.0004
SAT	SAT1	0.875	0.7656	-0.014	0.0002
	SAT2	0.641	0.4109	0.116	0.0135
	SAT3	0.909	0.8263	-0.078	0.0061
	SAT4	0.765	0.5852	-0.014	0.0002
UC	UC1	0.773	0.5975	0.051	0.0026
	UC2	0.641	0.4109	-0.117	0.0137
	UC3	0.909	0.8263	-0.089	0.0079
	UC4	0.765	0.5852	0.160	0.0256
TA	TA1	0.644	0.4147	-0.075	0.0056
	TA2	0.691	0.4775	-0.087	0.0076
	TA3	0.743	0.5520	0.014	0.0002
	TA4	0.764	0.5837	0.165	0.0272
Average		0.7919	0.6415	0.0002	0.0176

Table 7 The summary of the hypotheses testing

Hypothesis	Path coefficient (β)	<i>t</i> -value	Result
<i>H1</i> : Understandability \Rightarrow Confirmation	0.145 [†]	1.744	Marginally Supported
<i>H2</i> : Reliability \Rightarrow Confirmation	0.261 ^{**}	2.732	Supported
<i>H3</i> : Responsiveness \Rightarrow Confirmation	0.072 ^{n.s.}	1.173	Not Supported
<i>H4</i> : Assurance \Rightarrow Confirmation	0.159 [*]	1.972	Supported
<i>H5</i> : Interactivity \Rightarrow Confirmation	0.156 [*]	2.025	Supported
<i>H6a</i> : Understandability x TA \Rightarrow Confirmation	0.154 [*]	2.08	Supported
<i>H6b</i> : Reliability x TA \Rightarrow Confirmation	0.149 [*]	1.978	Supported
<i>H6c</i> : Responsiveness x TA \Rightarrow Confirmation	0.091 ^{n.s.}	0.737	Not Supported
<i>H6d</i> : Assurance x TA \Rightarrow Confirmation	0.130 [†]	1.789	Marginally Supported
<i>H6e</i> : Interactivity x TA \Rightarrow Confirmation	0.196 [*]	2.401	Supported
<i>H7</i> : Confirmation \Rightarrow Satisfaction	0.757 ^{***}	24.864	Supported
<i>H8</i> : Satisfaction \Rightarrow Use Continuance	0.760 ^{***}	22.469	Supported

[†] $p < 0.1$; ^{*} $p < 0.05$; ^{**} $p < 0.01$; ^{***} $p < 0.001$ (two-tailed); n.s. = not significant

Fig. 3 Results of the structural model



continuance ($\beta = 0.760$, t -value = 22.469), supporting $H7$ and $H8$, respectively. In addition, four control variables (gender, age, chatbot usage frequency, and the average time for chatbot use) were added to the structural model and it was found that none of them except for gender are significant. Overall, 49.2% (0.492) of the variance in confirmation, 57.3% (0.573) of the variance in satisfaction, and 57.3% (0.573) of the variance in use continuance intention were explained by exogenous and control variables.

The moderating effect of technology anxiety (TA) was also tested using the procedure introduced in Chin et al. (2003) with a calculation of effect sizes (Cohen 2013), path-coefficients (β), and the level of significance (t -value) of the interaction term (predictor variable \times moderator variable) in a PLS analysis. The results of the effect size calculation are also shown in Fig. 3. We found a very small but not negligible increase in the R^2 of confirmation (Henseler and Fassott 2010), by introducing TA as a moderating variable in the relationship between each one of service quality dimensions and confirmation. Results showed that interaction terms for all moderating relationships except the one with responsiveness are significant at the $\alpha = 0.05$ or 0.1 levels with the path-coefficients (β) and t -values calculated from the PLS algorithm and bootstrapping analysis: understandability \times TA ($\beta = 0.154$, t -value = 2.08), reliability \times TA ($\beta = 0.149$, t -value = 1.978), responsiveness \times TA ($\beta = 0.091$, t -value = 0.737), assurance \times TA ($\beta = 0.130$, t -value = 1.789), and interactivity \times TA ($\beta = 0.196$, t -value = 2.401). As such, hypotheses $H6a$, $H6b$, and $H6e$ were supported while $H6d$ was marginally supported. Based on Chin et al.'s (2003) interpretation of the moderating effect with a PLS analysis, we found that introducing the interaction factor (TA \times each of chatbot quality dimensions) increases the path-coefficient of the direct relationships between each service quality dimension and confirmation, although TA itself is negatively correlated with all the other variables (see Table 5). For example, in $H6e$, TA increased the path-efficient (the sensitivity or slope) for the

positive and significant effect of interactivity on confirmation by $\beta = 0.196$. These results imply that TA positively moderates the relationships between four service quality dimensions (except for responsiveness) and users' post-use confirmation. The higher levels of users' technology anxiety, the stronger the relationships between four chatbot quality dimensions (i.e., understandability, reliability, assurance, and interactivity) and their post-use confirmation, contradicting the findings from extant studies on TA of conventional (less human-like) technologies, which argued that technology anxiety is a mitigating factor for the relationships between influencing factors and users' technology adoption (Kim and Forsythe 2008; Lee and Yang 2013; Yang and Forney 2013).

Discussion

Research findings

Among the five chatbot quality dimensions, *understandability*, *reliability*, *assurance*, and *interactivity* are positively associated with confirmation ($H1$, $H2$, $H4$, and $H5$ supported), although only marginal support for the relationship between understandability and confirmation is found in our study, indicating that a more rigorous attempt to empirically confirm this relationship in a future study. For *reliability* and *assurance*, our findings are consistent with Rosen and Karwan (1994) who highlighted the importance of reliability and assurance in the service settings of restaurants, healthcare, and bookstore. In addition, the significant relationship between *understandability* and post-use confirmation is consistent with Kuligowska's (2015) who suggested that chatbots' ability to understand human's conversation is an essential evaluation criterion for commercial chatbot services. A positive relationship between *interactivity* and confirmation is also found, implying that interactivity, which is considered an advanced feature of smart services including smartwatches and chatbots,

can influence users' positive reactions and eventually increase their satisfaction and continuance use of chatbot services. Our findings are consistent with Cho et al. (2019) and Shin et al. (2013) who highlighted an important role of interactivity in facilitating users' attitudes and reactions to smart devices.

Interestingly, *responsiveness* is not significantly related to confirmation (*H3* not supported). This finding could be interpreted that in the Chinese OTA context, human services are relatively fast in speed compared to other countries, and thus users may not regard responsiveness as one of their key priorities for confirming their post-use expectation. For example, while the average waiting time to talk with a human representative in online customer service is 10 to 12 min in the United Kingdom, the waiting time in China is usually less than five minutes (Fishman et al. 2017). In addition, even the current responsiveness of OTA chatbot services of the three research sites are already very high (the average of three measured items of responsiveness is 5.41/7.00, the highest among all five quality measures, and the averages of the other four variables ranged from 4.6 to 5.1), so that users of the three OTAs do not feel that the variation in responsiveness can make a big difference for their evaluation for post-use confirmation. Accordingly, responsiveness and the speed of chatbot's responses may not be prioritized from the Chinese users' perspective.

Moreover, as service quality can be used to measure the relative importance of multiple dimensions in influencing consumers perceptions (Parasuraman et al. 1988), we conclude that in the context of chatbot-based OTA services in China, reliability ranks first, followed by assurance, interactivity, and understandability in the relative importance in chatbot quality dimensions. It is consistent with the findings of Parasuraman et al. (1988) that reliability turns out to be the most critical dimension, followed by assurance. Rosen and Karwan (1994) also found consistent findings that reliability is the most critical among all quality dimensions in teaching, restaurant, healthcare, and bookstore service settings.

Our moderating test results indicate that the higher levels of users' technology anxiety, the stronger the relationships between chatbot quality dimensions (i.e., understandability, reliability, assurance, and interactivity) and their post-use confirmation of chatbot-based OTAs (*H6a*, *H6b*, and *H6e* supported, and *H6d* marginally supported). These results are in line with the proposed hypotheses that technology anxiety should have a positive moderating impact on the relationships between service quality dimensions and users' post-use confirmation due to its' human-likeness possibly perceived by users. These results are contradictory to extant studies, which have found that technological anxiety mitigates the relationships between users' attitude/perceptions regarding technologies and use (or performance expectancy) (Kim and Forsythe

2008; Yang and Forney 2013), and statistically counter-intuitive to some extent since the correlation coefficients between technology anxiety and all other variables are negative. We further investigated our results based on the interpretation of moderation effects with a PLS analysis by Chin et al. (2003), but did find that technology anxiety strengthens the relationships between four abovementioned quality dimensions and confirmation. These results suggest that users' technology anxiety works differently for chatbot-enabled OTAs from traditional technology-enabled products or services. More specifically, if users with a high level of technology anxiety find that a chatbot-enabled OTA service is highly reliable, understandable, trustworthy (with high assurance), and interactive, they might consider it a human-like service with conversational interfaces, rather than a pure technology-enabled service with technological interfaces (e.g., find-select-click interfaces). Moreover, due to the perceived social presence of the services enabled by chatbots' capability of understanding and interacting with users like human agents (Lankton et al. 2015), they will rate post-use confirmation of their initial expectation from the chatbot-enabled service higher than those who have not much of technological anxiety. Our finding implies that chatbot services may somehow replace real human agents on OTAs at the current stage at least for those who have a high level of technology anxiety.

The positive relationships among confirmation, satisfaction, and use continuance intention (*H7* and *H8* supported) correspond to the extended post-acceptance model of IS continuance (i.e., ECM) proposed and tested by Bhattacharjee (2001). As expected, our findings are consistent with extant studies; if users find that their initial expectation of using chatbot-enabled OTA is confirmed, they will be satisfied with the service, which will eventually lead to use continuance intention.

Theoretical implications

This paper provides four major theoretical implications. First, this study contributes to the literature on the extended post-acceptance model of IS continuance (i.e., the ECM) since it extends the model by adding chatbot quality dimensions into the model and examined their roles in the chatbot service setting, which has been under-investigated. More specifically, this study introduced the five chatbot quality dimensions as antecedents, which we believe are relevant in the context of chatbot-enabled OTAs. The roles of these service quality antecedents are evaluated simultaneously with the theoretical loop of confirmation, satisfaction, and continuance use intention, as well as technological anxiety as a moderating factor. By theoretically extending the ECM, followed by the empirical analysis in the context of Chinese chatbot-enabled OTA services, this study contributes to the literature on the post-

acceptance model of IS continuance. In addition, this study contributes to the literature on chatbot user experience (UX) by identifying salient quality dimensions for contemporary chatbot services. By reviewing several seminal articles on IS quality, this study proposes the five service quality dimensions, provides conceptual definitions of them in the context of chatbot services, and empirically validates their roles in facilitating chatbot users' post-adoption confirmation. Since this study is among one of the first attempts to define the dimensions of chatbot service quality, the identification and empirical validation of these five dimensions (i.e., understandability, reliability, responsiveness, assurance, and interactivity) could be considered as one of our key theoretical contributions and shed a light onto the literature on chatbots and UX. Although our test results suggest that responsiveness is not significantly related to confirmation, the rest four dimensions are significantly related to post-use confirmation in the context of chatbot-enabled Chinese OTAs. These results on the relationships between the five quality dimensions and post-adoption confirmation of users might not be very surprising in that intuitively any IS-enabled service with good quality should result in a high level of users' post-use confirmation. However, it is worth to note that this study identifies two new service quality dimensions specific to human-like IT-enabled services (i.e., interactivity and understandability) on top of the three conventional service quality dimensions (i.e., reliability, responsiveness, and assurance) identified by extant studies, and empirically validates the differences among those dimensions. Therefore, these five quality dimensions, possibly with a few more quality dimensions specific to other contexts of the human-like IT-enabled service, can be further used in future studies.

Second, another important theoretical contribution of the study is that, based on the theoretical perspectives of uncanny valley theory (Mori 1970), technology humanness (Lankton et al. 2015), and the personification of technology (Purington et al. 2017), this study introduces technology anxiety as a moderating factor for the relationships between chatbot quality dimensions and post-use confirmation and finds an interesting result that technology anxiety strengthens the relationship between users' service quality and post-use confirmation in the context of chatbot-enabled OTA services. The literature on technology anxiety and computer anxiety has almost been consistent about the negative impact of technology anxiety on users' reactions to technology and adoption of it (Igarria and Parasuraman 1989; Meuter et al. 2003; Parasuraman 2000). Technology anxiety has either a mitigating effect on the relationship between users' evaluation of a technology-related product and their adoption or a negative direct relationship with users' adoption of technology if the targets of use are apparently 100% technology-oriented products or services (Meuter et al. 2003; Yang and Forney 2013). These negative impacts of technology anxiety are quite intuitive in that when

someone (i.e., a consumer) has anxiety about using technology, s/he tries to avoid technology-enabled services and seeks human alternatives for her/his needs, if any. However, as discussed above, since a chatbot-enabled customer service agent is designed to mimic human's conversation and the ways real humans interact with their customers to provide the perception of social presence to users (Rettie 2003; Short et al. 1976), it could be seen as a human-like agent (even though the users already acknowledge that the chatbot is supported by AI technology) by some users. Thus, those who have a high level of technology anxiety with possibly a lower initial expectation in our sample using chatbot-enabled OTAs (rather than more technology-enabled alternatives such as OTA mobile apps) could have found that their expectation is well confirmed with a high level of perceived service quality dimensions. With the advancement of an AI-driven NLP technology and its applications, as consumer-facing chatbot services can be developed to more human-like ones and users' perception of those advanced technologies in the future will change accordingly, technology anxiety will work differently in the way the results of this study implied. Perhaps, consumers would finally happen to do not care whether those agents on the other side of their service interface are humans or chatbots and could even personify the chatbots to treat them as human agents (Purington et al. 2017), as long as they can find what they are looking for from their counterparts in a timely and convenient manner. Some users with less technology anxiety will be comfortable using technology services such as mobile apps or online websites, while others with more technology anxiety will try to find human or human-like agents to get their services done. Therefore, we believe that the results of the positive moderating role of technology anxiety on the relationship between quality dimensions and post-use confirmation in this study can shed a light on the literature on the humanness or the personification of technology devices. As this study did not measure the humanness or the degree of personification of chatbot services, we suggest that future research can look into the relationships among users' technology anxiety, their perceived humanness (or personification of human-like technologies), service quality dimensions (with a new list of dimensions that fit the context of the study), and their adoption in various AI-driven technology contexts (e.g., humanoids, cyber service agents with the human voice and appearance, etc.).

Third, this study contributes to the new research stream of smart tourism. Among the three aspects of smart tourism research introduced by Gretzel et al. (2015), this study falls into the part of 'smart experience,' which focuses on enhancing users' experience of tourism services with real-time interaction, personalization, and context-awareness supported by new technologies (Buhalis and Amaranggana 2015; Neuhofer et al. 2015). Although the tourism industry is one of the most benefited industries by chatbot services with a

handful of recent conceptual studies (e.g., Buhalis and Yen 2020; Ukpabi et al. 2019; Zlatanov and Popesku 2019), it still has not been active in empirically investigating users' adoption of chatbot services, except very few attempts (e.g., Melián-González et al. 2019), not to mention in identifying and empirically validating the chatbot quality dimensions. This study, therefore, applies the theoretical model of the relationships among chatbot quality dimensions, technology anxiety, and users' post-use reactions into the context of smart tourism to provide empirical evidence on how current tourism service users assess the quality of AI-enabled smart tourism services (i.e., chatbot services) and react to the service considering their level of technology anxiety to these human-like services. Our results imply that chatbots in the tourism industry that can apparently provide reliable and responsive services with quality assured and real-time interactions will confirm the expectations of users and make them continuously use. Therefore, we believe that our investigation of tourism service consumers' perception of service quality, reaction to the technology by additionally including the moderating role of technology anxiety, their post-use confirmation/satisfaction, and adoption intention will contribute to the body of knowledge on smart tourism.

Practical implications

For practices, we argue that our research on chatbots has the potential to contribute to the audience of chatbot developers, user experience (UX) designers, quality assurance specialists, and e-Commerce/e-Service providers, and to create end-user awareness on what factors should be considered for better performance of chatbot services when interacting with the system. First, this study proposes the five quality characteristics (i.e., understandability, reliability, responsiveness, assurance, and interactivity) of chatbot services that possibly influence use continuance through confirmation and satisfaction. Among the five characteristics, understandability, reliability, assurance, and interactivity are significantly related to confirmation, and further positively related to use continuance through satisfaction. Therefore, several aspects of the chatbot service (e.g., the four quality dimensions found significant in this study) should be tested before implementation by developers and designers to make sure that clients and end-users have an engaging experience for continued use of this product. By answering the question of which chatbot quality dimensions are relatively more important among five dimensions, this study would help chatbot developers be well informed about which features of a chatbot to focus more on. Our results suggest reliability and assurance are the top two essential aspects, followed by understandability and interactivity in terms of relative importance. We suggest OTA managers, UX designers, and developers to continuously pay attention to make

chatbots more reliable, trustworthy, and understandable, providing more interactive functions to satisfy users' needs and sense of control at the same time.

Second, our results suggest that responsiveness is not significantly related to post-use confirmation, which may imply that chatbot users in our sample do not consider responsiveness a major factor influencing their confirmation for chatbot services in the Chinese OTA context. However, as discussed, this result may have been due to the fact that the chatbot-based OTAs in our sample are already responsive enough with its mean value of the variable 5.41/7.00. Thus, we did not see many variations in the responsiveness of chatbot services in our sample. However, this result does not necessarily mean that the feature of responsiveness in chatbot services is not an important dimension in other contexts of chatbot services. Although this result could suggest to chatbot developers, UX designers, quality assurance specialists, and e-Commerce/e-Service providers especially in China that they should focus more on understandability, reliability, assurance, and interactivity, rather than focusing on the responding speed, the role of responsiveness could still be an important factor in other chatbot contexts, especially where the responding time of the chatbot service is not properly achieved.

Third, although most people know that chatbot services are facilitated by AI-enabled NLP technologies, our result with the role of technology anxiety also suggests that, at least in the context of OTA users in China, chatbots could be considered differently from other obviously 100% technology-enabled services. Our result does not necessarily mean that currently available chatbot services can replace human agents very nicely; it is acknowledged that there are still a considerable proportion of questions that chatbots cannot answer in a proper manner. According to a survey with 500 participants, 60% and 56% of the respondents in the U.S. and the UK, respectively, prefer human agents to chatbots for complex inquiries in their service engagements (CGS 2018). Our positive moderating role of technology anxiety in the relationships between quality dimensions and post-use confirmation could have been due to the fact that there are quite a few groups of people who want avoid other technology-enabled alternatives (e.g., mobile apps) because of their high level of technology anxiety; their probably lower initial expectation for the chatbot service is well confirmed after they recognize that the OTA chatbot services were good enough in each dimension of service quality. Nevertheless, this result still has an important practical implication that, with the advancement of chatbot technologies that increase so-called its humanness, e-Commerce and e-Service providers can essentially churn the users who seek human agents (due to their high level of technology anxiety) into chatbot users at least for service inquiries that are not very complex, which can further help them save cost for their customer services.

Limitations and future research

There are five major limitations of this study. First, a specific context was chosen in this study – chatbot-based OTAs. Even though our findings could provide insights of chatbot quality dimensions into a vast audience of practitioners, our results may be difficult to generalize to all chatbot services. In addition, only the Chinese market was investigated. The results from our hypothesized relationships may be different from different contexts or countries. Future studies therefore should test this model across a variety of industries and countries to improve the generalization of our research model and gain various implications for more audiences. Moreover, the role of responsiveness should be further investigated in other contexts of chatbot services (e.g., hotel and restaurant settings in other countries) to increase the external validity of our research model.

Second, related to the first limitation, as this study covers only a single cultural context (China), the cultural difference with regard to users' reactions to human-like technologies is not fully taken into consideration. Therefore, it might be problematic to generalize the results of our study over and above the context of Chinese chatbot users. However, we believe that providing more information about the use and adoption situation in China compared to Western countries can help readers comparatively interpret the results of our study. The followings are some interesting findings we have learned from industry data. First, despite the fact that the Asia-Pacific region is expected to have the highest growth rate in the chatbot market (Chatbot Market 2019) and China has become the world's second-largest OTA market in 2018 (Phocuswright Research 2019), the chatbot usage rate in the Asia-Pacific region (21%) is still reported to be less than a half of that in the North American region (45%) (Andre 2020). This penetration information can be interpreted that chatbot users (our research participants) in China could belong to the early adopter groups based on Roger's innovation diffusion categories (Rogers 2010), compared to North American chatbot users (the early majority groups). Therefore, our survey respondents could be considered more innovative than North American users, implying that current Chinese chatbot users will show more innovative tendencies when it comes to reacting to human-like technologies, regardless of whether they have a high or a low level of technology anxiety. Second, Chinese consumers have relatively less trust in online products due to various product quality safety issues (Chatbots Magazine 2017). Despite some anecdotal evidence that the situation has recently been improved, their lack of trust in products and services provided online still encourages Chinese consumers to ask more questions before making purchases online. Chatbot services, if their quality dimensions are acceptable and better than expected, should be more welcomed by Chinese online consumers than Westerners, so that

the relationships between chatbot service quality dimensions and users' post-use confirmation (or other adoption variables) should be stronger than those of Westerners. Based on our research model and the consideration of cultural differences, a future study could further investigate the relationships among service quality, technology anxiety, humanness, and users' post-adoption variables in a more appropriate manner.

Third, only technology anxiety was examined to see whether there is a moderating effect on the relationships between chatbot quality dimensions and post-use confirmation. Practically, there might be other kinds of anxiety that influence our proposed relationship, such as social anxiety, indicating the difficulty to have conversations or interactions with strangers (Schlenker and Leary 1982). Thus, the role of various kinds of anxiety could be further examined to discover what are the obstacles that hinder users' confirmation, satisfaction, and use continuance towards chatbot services.

Fourth, although the original ECM of IS includes perceived usefulness to explain users' post-use assessment of instrumentality of a product, this study focuses only on key influencing chatbot quality dimensions for post-use confirmation, satisfaction, and use continuance using the extended post-acceptance model of IS continuance model. Although our research purpose and scope are to investigate what features of chatbot quality influence users' post-use confirmation, instead of the evaluation of instrumentality of the service, which could be well captured by service quality dimensions we included, in future research, it would be worthwhile to include more variables, such as perceived usefulness and perceived security to provide some additional interesting implications.

Lastly, although we have added four control variables into our structural model testing, more potentially important factors for users' continuance intention are not taken into consideration. For example, the Internet usage experience or the types of devices they use for chatbot services (e.g., smartphone, tablet PC, or normal PC) could be a factor that could have potentially influenced our dependent variable. Future studies therefore could control more relevant variables to increase the validity of our research model.

Conclusion

This study extends the post-acceptance model of IS continuance in the context of chatbot-enabled OTAs by identifying several quality dimensions of chatbot services (i.e., understandability, reliability, responsiveness, assurance, and interactivity) as antecedents for user's post-acceptance confirmation, as well as by proposing technology anxiety toward chatbots as a moderating factor for these relationships. This study found that, with the exception for responsiveness, most quality dimensions of chatbot services are significantly related to confirmation, which in turn leads to users' continuance

intention. It also revealed that technology anxiety positively moderates the relationships between chatbot quality dimensions and post-use confirmation, suggesting some users may treat chatbot services as human-like agents. Practically, the results of this study provide quality assurance specialists, e-service providers, and chatbot developers with guidelines to better understand chatbot users in enhancing their service adoption in the tourism and hospitality sector. Future studies could extend this model across a variety of industries and cultural backgrounds to improve the generalizability of our research model and gain various implications for more audiences. In addition, future studies could extend our research model by further investigating the relationships among other service quality dimensions, which are salient in other

technology contexts, technology anxiety, other types of anxiety (e.g., social interaction anxiety), humanness, and other post-adoption variables (e.g., recommendation intention) to provide additional implications.

Acknowledgements This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2019S1A3A2098438).

Compliance with ethical standards

Disclosure No potential conflict of interest was reported by the authors.

Appendix

Table 8 Questionnaire items

Construct	Items	References
Use continuance	1. My intentions are to continue using this chatbot service over other alternative means of communication or searching tools on this OTA. 2. All things considered, I expect to continue using this chatbot service often in the future. 3. I can see myself increasing the use of this chatbot service if possible.	Bhattacharjee (2001); Hamari et al. (2016)
Satisfaction	1. I like to use the chatbot service from this online travel website. 2. I am pleased with the experience of using this chatbot service. 3. I think that using the chatbot service on this travel website is a good idea. 4. Overall, I am satisfied with the experience of using this chatbot service.	Fang et al. (2011)
Confirmation	1. My experience with using this chatbot service was better than what I expected. 2. The service level provided by this chatbot service was better than what I expected. 3. Overall, most of my expectations regarding the usage of this chatbot service were confirmed.	Hong et al. (2006)
Technology anxiety	1. I have avoided chatbot services because they are unfamiliar to me. 2. I hesitate to use chatbot services for fear of making mistakes I cannot correct. 3. I have difficulty understanding most technological matters relating to chatbot services. 4. I am not able to keep up with important technological advances, such as the development of chatbot services.	Meuter et al. (2003); Thatcher and Perrewé (2002)
Understandability	1. I feel that what I'm saying to this chatbot system is well understood by the system. 2. I feel that the words in my questions are well understood by this chatbot service. 3. I feel that this chatbot system understands my intentions when I ask a question to it.	Oulasvirta et al. (2006)
Reliability	1. This chatbot service is dependable. 2. When I have problems, this chatbot service is sympathetic and reassuring. 3. I feel I could rely on this chatbot for its services regarding my travel needs.	Parasuraman et al. (1988)
Responsiveness	1. This chatbot provides prompt services that meet my expectations. 2. This chatbot service responds to my requests promptly. 3. This chatbot provides services exactly when I need them without any delay.	Parasuraman et al. (1988)
Assurance	1. I trust this chatbot service. 2. I feel safe and assured to have a conversation with this chatbot service. 3. The chatbot service of this travel website has enough knowledge to answer my travel questions.	Parasuraman et al. (1988)
Interactivity	1. I can be in control of my personal needs through this chatbot service. 2. I perceive this chatbot service to be sensitive to my personal needs. 3. This chatbot service provides an opportunity for me to give my responses.	Cho et al. (2019)

References

- Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 72–80. <https://doi.org/10.14569/IJACSA.2015.060712>.
- Akter, S., Ray, P., & D'Ambra, J. (2013). Continuance of mHealth services at the bottom of the pyramid: the roles of service quality and trust. *Electronic Markets*, 23(1), 29–47. <https://doi.org/10.1007/s12525-012-0091-5>.
- AlHagbani, E. S., & Khan, M. B. (2016). Challenges facing the development of the Arabic chatbot. In *Proceedings of the First International Workshop on Pattern Recognition, International Society for Optics and Photonics*. Tokyo, Japan, 100–110. <https://doi.org/10.1117/12.2240849>.
- Andre, L. (2020). 79 Critical chatbot statistics: 2020 data analysis & market share. Retrieved from <https://financesonline.com/chatbot-statistics/>
- Ask, J. A., Facemire, M., & Hogan, A. (2016). The state of chatbots. *Forrester.com Report*, 20.
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly*, 36(3), 421–458. <https://doi.org/10.2307/2393203>.
- Bhattacharjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly*, 25(3), 351–370. <https://doi.org/10.2307/3250921>.
- Buhalis, D., & Amaranggana, A. (2015). Smart tourism destinations: Enhancing tourism experience through personalisation of services. In *Proceedings of the Tussyadiah & A. Inversini (Eds.), Information and Communication Technologies in Tourism 2015*. Heidelberg: Springer. 377–389. https://doi.org/10.1007/978-3-319-14343-9_28.
- Buhalis, D., & Yen, E. C. S. (2020). Exploring the use of chatbots in hotels: Technology providers' perspective. In *Proceedings of Information and Communication Technologies in Tourism 2020*, Surrey, United Kingdom, 231–242. https://doi.org/10.1007/978-3-030-36737-4_19.
- Carter, L., & Liu, D. (2018). Technology humanness, trust and e-government adoption. In *Proceedings of Australasian Conference on Information Systems 2018*. Sydney, Australia, 114–120. <https://doi.org/10.5130/acis2018.cp>.
- CGS. (2018). Chatbots deliver speed, but consumers still want humans. Are we moving too quickly to automation? Retrieved from: https://www.cgsinc.com/sites/default/files/media/resources/pdf/CGS_Consumer%2BCustServ%2Binfographic%2B2018.pdf
- Chatbot Market. (2019). Chatbot market by component (solutions and services), usage (websites and contact centers), technology, deployment model, application (customer support and personal assistant), organization size, vertical, and region - global forecast to 2024. Retrieved from <https://www.marketsandmarkets.com/Market-Reports/smart-advisor-market-72302363.html>
- Chatbots Magazine. (2017). China, WeChat, and the origins of chatbots. Retrieved from <https://chatbotsmagazine.com/china-wechat-and-the-origins-of-chatbots-89c481f15a44>
- Chen, S. C., Liu, M. L., & Lin, C. P. (2013). Integrating technology readiness into the expectation-confirmation model: an empirical study of mobile services. *Cyberpsychology, Behavior and Social Networking*, 16(8), 604–612. <https://doi.org/10.1089/cyber.2012.0606>.
- Cheng, Y. M. (2014). Extending the expectation-confirmation model with quality and flow to explore nurses' continued blended e-learning intention. *Information Technology & People*, 27(3), 230–258. <https://doi.org/10.1108/ITP-01-2013-0024>.
- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In Marcoulides, G. A. (Ed.), *Modern methods for business research* (pp. 295–336). Mahwah: Lawrence Erlbaum Associates.
- Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, 14(2), 189–217. <https://doi.org/10.1287/isre.14.2.189.16018>.
- Cho, W. C., Lee, K. Y., & Yang, S. B. (2019). What makes you feel attached to smartwatches? The stimulus-organism-response (S-O-R) perspectives. *Information Technology & People*, 32(2), 319–343. <https://doi.org/10.1108/ITP-05-2017-0152>.
- Chung, K., & Park, R. C. (2019). Chatbot-based healthcare service with a knowledge base for cloud computing. *Cluster Computing*, 22(1), 1925–1937. <https://doi.org/10.1007/s10586-018-2334-5>.
- Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, 117, 587–595. <https://doi.org/10.1016/j.jbusres.2018.10.004>.
- Ciechanowski, L., Przegalinska, A., Magnuski, M., & Gloor, P. (2019). In the shades of the uncanny valley: an experimental study of human-chatbot interaction. *Future Generation Computer Systems*, 92, 539–548. <https://doi.org/10.1016/j.future.2018.01.055>.
- Cohen, J. (2013). Statistical power analysis for the behavioral sciences. *Routledge*. <https://doi.org/10.4324/9780203771587>.
- Danilava, S., Busemann, S., Schommer, C., & Ziegler, G. (2013). Why are you silent? Towards responsiveness in chatbots. In *Proceedings of Avec le Temps! Time, Tempo, and Turns in Human-Computer Interaction. Workshop at CHI 2013*, Paris, France.
- DeLone, W. H., & McLean, E. R. (1992). Information systems success: the quest for the dependent variable. *Information Systems Research*, 3(1), 60–95. <https://doi.org/10.1287/isre.3.1.60>.
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of Management Information Systems*, 19(4), 9–30. <https://doi.org/10.1080/07421222.2003.11045748>.
- Deng, Z., Mo, X., & Liu, S. (2014). Comparison of the middle-aged and older users' adoption of mobile health services in China. *International Journal of Medical Informatics*, 83(3), 210–224. <https://doi.org/10.1016/j.ijmedinf.2013.12.002>.
- Dubrova, D. (2020). Chatbot for travel industry: benefits, use cases, and a development guide. Retrieved from <https://theappsolutions.com/blog/how-to/chatbot-for-travel-business/>
- Fang, Y. H., Chiu, C. M., & Wang, E. T. (2011). Understanding customers' satisfaction and repurchase intentions. *Internet Research*, 21(4), 479–503. <https://doi.org/10.1108/10662241111158335>.
- Fishman, T. D., Eggers, W. D., & Kishnani, P. (2017). AI-augmented human services. Retrieved from <https://www2.deloitte.com/us/en/insights/industry/public-sector/artificial-intelligence-technologies-human-services-programs.html>
- FlowXO (2020). What industry are using chatbots today? Retrieved from <https://flowxo.com/what-industries-are-using-chatbots-today/>
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>.
- Go, E., & Sundar, S. S. (2019). Humanizing chatbots: the effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, 97, 304–316. <https://doi.org/10.1016/j.chb.2019.01.020>.
- Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. (2015). Smart tourism: foundations and developments. *Electronic Markets*, 25(3), 179–188. <https://doi.org/10.1007/s12525-015-0196-8>.
- Gupta, R. (2019). Chatbots providing their worth in travel e-commerce. Retrieved from <https://www.chinatravelnews.com/article/131644>

- Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology*, 67(9), 2047–2059. <https://doi.org/10.1002/asi.23552>.
- Heeter, C. (1989). Implications of new interactive technologies for conceptualizing communication. In Salvaggio, J. L., Bryant, J. (Eds.). *Media Use in the Information Age: Emerging Patterns of Adoption and Consumer Use* (pp. 217–235).
- Henseler, J., & Fassott, G. (2010). Testing moderating effects in PLS path models: An illustration of available procedures. In: Esposito Vinzi V., Chin W., Henseler J., Wang H. (Eds.). *Handbook of Partial Least Squares. Springer Handbooks of Computational Statistics* (pp. 713–735). https://doi.org/10.1007/978-3-540-32827-8_31.
- Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: a comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, 49, 245–250. <https://doi.org/10.1016/j.chb.2015.02.026>.
- Hong, S., Thong, J. Y., & Tam, K. Y. (2006). Understanding continued information technology usage behavior: a comparison of three models in the context of mobile internet. *Decision Support Systems*, 42(3), 1819–1834. <https://doi.org/10.1016/j.dss.2006.03.009>.
- Hong, J. C., Lin, P. H., & Hsieh, P. C. (2017). The effect of consumer innovativeness on perceived value and continuance intention to use smartwatch. *Computers in Human Behavior*, 67, 264–272. <https://doi.org/10.1016/j.chb.2016.11.001>.
- Hsu, C. L., & Lin, J. C. C. (2019). Understanding continuance intention to use online to offline (O2O) apps. *Electronic Markets*, 30(4), 883–897. <https://doi.org/10.1007/s12525-019-00354-x>.
- Igbaria, M., & Parasuraman, S. (1989). A path analytic study of individual characteristics, computer anxiety and attitudes toward micro-computers. *Journal of Management*, 15(3), 373–388. <https://doi.org/10.1177/014920638901500302>.
- Ivanov, S. (2020). The impact of automation on tourism and hospitality jobs. *Information Technology & Tourism*, 22(2), 205–215. <https://doi.org/10.1007/s40558-020-00175-1>.
- Ivanov, S., & Webster, C. (2019). *Robots, artificial intelligence and service automation in travel, tourism and hospitality*. Emerald Publishing Limited. <https://doi.org/10.1108/9781787566873>.
- Jain, M., Kumar, P., Kota, R., & Patel, S. N. (2018). Evaluating and informing the design of chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference (IDS)*, ACM, Hong Kong, Hong Kong. 895–906. <https://doi.org/10.1145/3196709.3196735>.
- Johnson, K., Li, Y., Phan, H., Singer, J., & Trinh, H. (2012). The innovative success that is Apple, Inc. *Marshall University Marshall Digital Scholar*, 418–460.
- Kahiga, A. (2019). *A model for adoption of chatbots in Kenya: A case study of Zuku Telegram Bot*. [Doctoral dissertation, University of Nairobi]. University of Nairobi. <http://erepository.uonbi.ac.ke/handle/11295/107169>.
- Kalia, A. K., Telang, P. R., Xiao, J., & Vukovic, M. (2017). Quark: A methodology to transform people-driven processes to chatbot services. In *Proceedings of International Conference on Service-Oriented Computing*, Springer, Cham. 53–61. https://doi.org/10.1007/978-3-319-69035-3_4.
- Kamide, H., Kawabe, K., Shigemi, S., & Arai, T. (2014). Relationship between familiarity and humanness of robots: quantification of psychological impressions toward humanoid robots. *Advanced Robotics*, 28(12), 821–832. <https://doi.org/10.1080/01691864.2014.893837>.
- Kang, H. J., & Kim, S. I. (2017). Evaluation on the usability of chatbot intelligent messenger mobile services: focusing on Google (Allo) and Facebook (M messenger). *Journal of the Korea Convergence Society*, 8(9), 271–276. <https://doi.org/10.15207/JKCS.2017.8.9.271>.
- Kätsyri, J., Förger, K., Mäkäriäinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, 6, 390. <https://doi.org/10.3389/fpsyg.2015.00390>.
- Kettinger, W. J., & Lee, C. C. (1994). Perceived service quality and user satisfaction with the information services function. *Decision Sciences*, 25(5–6), 737–766. <https://doi.org/10.1111/j.1540-5915.1994.tb01868.x>.
- Kim, J., & Forsythe, S. (2008). Adoption of virtual try-on technology for online apparel shopping. *Journal of Interactive Marketing*, 22(2), 45–59. <https://doi.org/10.1002/dir.20113>.
- Kim, S. H., Bae, J. H., & Jeon, H. M. (2019). Continuous intention on accommodation apps: integrated value-based adoption and expectation–confirmation model analysis. *Sustainability*, 11(6), 1578–1595. <https://doi.org/10.3390/su11061578>.
- Kotrlík, J. W., & Redmann, D. H. (2009). A trend study: technology adoption in the teaching–learning process by secondary agriscience teachers – 2002 and 2007. *Journal of Agricultural Education*, 50(2), 64–76. <https://doi.org/10.5032/jae.2009.02062>.
- Kuligowska, K. (2015). Commercial chatbot: performance evaluation, usability metrics and quality standards of embodied conversational agents. *Professionals Center for Business Research*, 2(2), 1–16. <https://doi.org/10.18483/PCBR.22>.
- Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: rethinking trust in technology. *Journal of the Association for Information Systems*, 16(10), 880–918. <https://doi.org/10.17705/1jais.00411>.
- Lee, H. M., & Chen, T. (2014). Perceived quality as a key antecedent in continuance intention on mobile commerce. *International Journal of Electronic Commerce Studies*, 5(2), 123–142. <https://doi.org/10.7903/ijecs.1150>.
- Lee, M. K., & Park, H. (2019). Exploring factors influencing usage intention of chatbot: chatbot in financial service. *Journal of the Korean Society for Quality Management*, 47(4), 755–765. <https://doi.org/10.7469/JKSQM.2019.47.4.755>.
- Lee, H. J., & Yang, K. (2013). Interpersonal service quality, self-service technology (SST) service quality, and retail patronage. *Journal of Retailing and Consumer Services*, 20(1), 51–57. <https://doi.org/10.1016/j.jretconser.2012.10.005>.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: a methodology for information quality assessment. *Information & Management*, 40(2), 133–146. [https://doi.org/10.1016/S0378-7206\(02\)00043-5](https://doi.org/10.1016/S0378-7206(02)00043-5).
- Lee, K., Lee, K. Y., & Sheehan, L. (2019). Hey Alexa! A magic spell of social glue? Sharing a smart voice assistant speaker and its impact on users' perception of group harmony. *Information Systems Frontiers*, 22(3), 563–583. <https://doi.org/10.1007/s10796-019-09975-1>.
- Li, H., Gupta, A., Zhang, J., & Flor, N. (2020a). Who will use augmented reality? An integrated approach based on text analytics and field survey. *European Journal of Operational Research*, 281(3), 502–516. <https://doi.org/10.1016/j.ejor.2018.10.019>.
- Li, Y., Grandison, T., Silveyra, P., Douraghy, A., Guan, X., Kieselbach, T., Li, C. & Zhang, H. (2020b). Jennifer for COVID-19: an NLP-powered chatbot built for the people and by the people to combat misinformation. In *Proceedings of ACL 2020 Workshop on natural language processing for COVID-19 (NLP-COVID)*, Seattle, Washington, 1–9.
- Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: the effect of institutional pressures and the mediating role of top management. *MIS Quarterly*, 31(1), 59–87. <https://doi.org/10.2307/25148781>.

- Liao, C., Palvia, P., & Lin, H. N. (2010). Stage antecedents of consumer online buying behavior. *Electronic Markets*, 20(1), 53–65. <https://doi.org/10.1007/s12525-010-0030-2>.
- Liu, D., Cheng, G., An, L., Gan, X., Wu, Y., Zhang, B., Hu, S., Zeng, Y., & Wu, L. (2019). Public knowledge about dementia in China: a national WeChat-based survey. *International Journal of Environmental Research and Public Health*, 16(21), 4231–4245. <https://doi.org/10.3390/ijerph16214231>.
- Long, J., Yuan, J., & Lee, H. M. (2019). How to program a chatbot: An introductory project and student perceptions. *Issues in Informing Science and Information Technology*, 16, 1–31. <https://doi.org/10.28945/4282>.
- Lopatovska, I., & Williams, H. (2018). Personification of the Amazon Alexa: BFF or a mindless companion. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*. New Brunswick, NJ, 265–268. <https://doi.org/10.1145/3176349.3176868>.
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: machines vs. humans: the impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6), 913–1084. <https://doi.org/10.1287/mksc.2019.1192>.
- McKinney, V., Yoon, K., & Zahedi, F. M. (2002). The measurement of web-customer satisfaction: an expectation and disconfirmation approach. *Information Systems Research*, 13(3), 296–315. <https://doi.org/10.1287/isre.13.3.296.76>.
- McTear, M., Callejas, Z., & Griol, D. (2016). The conversational interface: Talking to smart devices. In *The conversational Interface* (pp. 283–308). Cham: Springer. https://doi.org/10.1007/978-3-319-32967-3_13.
- Meerschman, H., & Verkeyn, J. (2019). *Towards a better understanding of service quality attributes of a chatbot*. [Master's dissertation. Ghent University] Ghent University.
- Meijer, R. R., & Sijsma, K. (1995). Detection of aberrant item score patterns: a review of recent developments. *Applied Measurement in Education*, 8(3), 261–272. https://doi.org/10.1207/s15324818ame0803_5.
- Melián-González, S., Gutiérrez-Taño, D., & Bulchand-Gidumal, J. (2019). Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism*, 1–19. <https://doi.org/10.1080/13683500.2019.1706457>.
- Meuter, M. L., Ostrom, A. L., Bitner, M. J., & Roundtree, R. (2003). The influence of technology anxiety on consumer use and experiences with self-service technologies. *Journal of Business Research*, 56(11), 899–906. [https://doi.org/10.1016/S0148-2963\(01\)00276-4](https://doi.org/10.1016/S0148-2963(01)00276-4).
- Milenkovic, J. (2020). 43 useful online banking statistics: all about mobile money! Retrieved from <https://kommandotech.com/statistics/online-banking-statistics/>
- Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy*, 7, 33–35.
- Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study. *Digital Health*, 5, 1–12. <https://doi.org/10.1177/2055207619871808>.
- Nascimento, B., Oliveira, T., & Tam, C. (2018). Wearable technology: what explains continuance intention in smartwatches? *Journal of Retailing and Consumer Services*, 43, 157–169. <https://doi.org/10.1016/j.jretconser.2018.03.017>.
- Nelson, R. R., Todd, P. A., & Wixom, B. H. (2005). Antecedents of information and system quality: an empirical examination within the context of data warehousing. *Journal of Management Information Systems*, 21(4), 199–235. <https://doi.org/10.1080/07421222.2005.11045823>.
- Neuhofer, B., Buhalis, D., & Ladkin, A. (2015). Smart technologies for personalized experiences: a case study in the hospitality domain. *Electronic Markets*, 25(3), 243–254. <https://doi.org/10.1007/s12525-015-0182-1>.
- Nguyen, T. (2019). *Potential effects of chatbot technology on customer support: A case study*. [Master's dissertation, Aalto University] Aalto University.
- Nica, I., Tazl, O. A., & Wotawa, F. (2018). Chatbot-based tourist recommendations using model-based reasoning. In *Proceedings of the 20th International Workshop on Configuration*, Graz, Austria, 25–30.
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, 17(4), 460–469. <https://doi.org/10.1177/002224378001700405>.
- Osman, M. (2020). Ecommerce statistics for 2020: chatbots, voice, omni-channel marketing. Retrieved from <https://kinsta.com/blog/e-commerce-statistics/>
- Oulasvirta, A., Engelbrecht, K.P., Jameson, A., & Möller, S. (2006). The relationship between user errors and perceived usability of a spoken dialogue system. In *Proceedings of 2nd ISCA/DEGA Tutorial and Research Workshop on Perceptual Quality of Systems*, 61–67.
- Parasuraman, A. (2000). Technology readiness index (TRI): a multiple-item scale to measure readiness to embrace new technologies. *Journal of Service Research*, 2(4), 307–320. <https://doi.org/10.1177/109467050024001>.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40.
- Park, E. (2020). User acceptance of smart wearable devices: an expectation-confirmation model approach. *Telematics and Informatics*, 47, 101318. <https://doi.org/10.1016/j.tele.2019.101318>.
- Park, M., Aiken, M., & Salvador, L. (2018). How do humans interact with chatbots? An analysis of transcripts. *International Journal of Management & Information Technology*, 14, 3338–3350. <https://doi.org/10.24297/ijmit.v14i0.7921>.
- Pavlou, P. A., Liang, H., & Xue, Y. (2007). Understanding and mitigating uncertainty in online exchange relationships: a principal-agent perspective. *MIS Quarterly*, 31(1), 105–136. <https://doi.org/10.2307/25148783>.
- Pearl, C. (2016). *Designing voice user interfaces: Principles of conversational experiences*. O'Reilly Media, Inc.
- Pereira, J., & Díaz, Ó. (2018). Chatbot dimensions that matter: Lessons from the trenches. In *Proceedings of International Conference on Web Engineering*. Cáceres, Spain. 129–135. https://doi.org/10.1007/978-3-319-91662-0_9.
- Phocuswright Research. (2019). How China became the world's 2nd-largest OTA market. Retrieved from <https://www.phocuswright.com/Travel-Research/Research-Updates/2019/How-China-became-the-worlds-2nd-largest-OTA-market>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>.
- Prasetya, S. A., Erwin, A., & Galinium, M. (2018). Implementing Indonesian language chatbot for ecommerce site using artificial intelligence markup language (AIML). In *Prosiding Seminar Nasional Pakar*, 313–322.
- Purington, A., Taft, J. G., Sannon, S., Bazarova, N. N., & Taylor, S. H. (2017). Alexa is my new BFF: Social roles, user satisfaction, and personification of the amazon echo. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. Denver, Col., 2853–2859. <https://doi.org/10.1145/3027063.3053246>.
- Quah, J. T., & Chua, Y. W. (2019). Chatbot assisted marketing in financial service industry. In *International Conference on Services Computing*, Springer, Cham, 107–114. https://doi.org/10.1007/978-3-030-23554-3_8.

- Rahman, A., Al Mamun, A., & Islam, A. (2017). Programming challenges of chatbot: current and future prospective. In Proceedings of the 2017 IEEE region 10 humanitarian technology conference (R10-HTC), Dhaka, Bangladesh, 75–78. <https://doi.org/10.1109/R10-HTC.2017.8288910>.
- Rettie, R. (2003). Connectedness, awareness and social presence. In Proceedings of the 6th Annual International Workshop on Presence. Aalborg, Denmark, 1–7.
- Rogers, E. M. (2010). *Diffusion of innovations*. New York: Simon and Schuster.
- Rosen, D. L., & Karwan, K. R. (1994). Prioritizing the dimensions of service quality: an empirical investigation and strategic assessment. *International Journal of Service Industry Management*, 5(4), 39–52. <https://doi.org/10.1108/09564239410068698>.
- Sannon, S., Stoll, B., DiFranzo, D., Jung, M., & Bazarova, N. N. (2018). How personification and interactivity influence stress-related disclosures to conversational agents. In Proceedings of Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing, Jersey City, NJ, 285–288. <https://doi.org/10.1145/3272973.3274076>.
- Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: a conceptualization model. *Psychological Bulletin*, 92(3), 641–669. <https://doi.org/10.1037/0033-2909.92.3.641>.
- Sensuse, D. I., Dhevanty, V., Rahmanasari, E., Permatasari, D., Putra, B. E., Lusa, J. S., Misbah, M., & Prima, P. (2019). Chatbot evaluation as knowledge application: a case study of PT ABC. In Proceedings of 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE), Bangkok, Thailand, 1–6. <https://doi.org/10.1109/ICITEE.2019.8929967>.
- Serban, I. V., Sankar, C., Germain, M., Zhang, S., Lin, Z., Subramanian, S., Kim, T., Pieper, M., Chandar, S., Ke, N. R., Rajeshwar, S., de Brebisson, A., Sotelo, J. M. R., Suhubdy, D., Michalski, V., Nguyen, A., Pineau, J., & Bengio, Y. (2017). A deep reinforcement learning chatbot. preprint available at arXiv:1709.02349.
- Shin, D. H., Hwang, Y., & Choo, H. (2013). Smart TV: are they really smart in interacting with people? Understanding the interactivity of Korean Smart TV. *Behaviour & Information Technology*, 32(2), 156–172. <https://doi.org/10.1080/0144929X.2011.603360>.
- Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.
- Skantze, G. (2007). *Error handling in spoken dialogue systems: Managing uncertainty, grounding and miscommunication*. Gabriel Skantze.
- Sundar, S. S., Bellur, S., Oh, J., Jia, H., & Kim, H. S. (2016). Theoretical importance of contingency in human-computer interaction: effects of message interactivity on user engagement. *Communication Research*, 43(5), 595–625. <https://doi.org/10.1177/0093650214534962>.
- Susanto, A., Chang, Y., & Ha, Y. (2016). Determinants of continuance intention to use the smartphone banking services: an extension to the expectation-confirmation model. *Industrial Management & Data Systems*, 116(3), 508–525. <https://doi.org/10.1108/IMDS-05-2015-0195>.
- Terpening, E., & Littleton, A. (2016). *The 2016 state of social business: social's shift from innovator to integrator*. Altimeter.
- Thatcher, J. B., & Perrewew, P. L. (2002). An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *MIS Quarterly*, 26(4), 381–396. <https://doi.org/10.2307/4132314>.
- Thies, I. M., Menon, N., Magapu, S., Subramony, M., & O'Neill, J. (2017). How do you want your chatbot? An exploratory wizard-of-oz study with young, urban Indians. In Proceedings of the IFIP Conference on Human-Computer Interaction, Springer, Cham, 441–459. https://doi.org/10.1007/978-3-319-67744-6_28.
- Tussyadiah, I. (2020). A review of research into automation in tourism: launching the annals of tourism research curated collection on artificial intelligence and robotics in tourism. *Annals of Tourism Research*, 81, 102883. <https://doi.org/10.1016/j.annals.2020.102883>.
- Ukpabi, D. C., Aslam, B., & Karjaluo, H. (2019). Chatbot adoption in tourism services: A conceptual exploration. In Ivanov, S., & Webster, C. (Eds.) *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*, 105–121. <https://doi.org/10.1108/978-1-78756-687-320191006>.
- Uliyar, S. (2017). Oracle a primer: oracle intelligent bots - powered by artificial intelligence. Retrieved from: <http://www.oracle.com/us/technologies/mobile/chatbots-primer-3899595.pdf>
- Urgen, B. A., Kutas, M., & Saygin, A. P. (2018). Uncanny valley as a window into predictive processing in the social brain. *Neuropsychologia*, 114, 181–185. <https://doi.org/10.1016/j.neuropsychologia.2018.04.027>.
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5–33. <https://doi.org/10.1080/07421222.1996.11518099>.
- 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 front-line. *Journal of Service Management*, 29(5), 907–931. <https://doi.org/10.1108/JOSM-04-2018-0119>.
- Yang, K., & Forney, J. C. (2013). The moderating role of consumer technology anxiety in mobile shopping adoption: differential effects of facilitating conditions and social influences. *Journal of Electronic Commerce Research*, 14(4), 334–347.
- Zamora, J. (2017). I'm sorry, Dave, I'm afraid I can't do that: Chatbot perception and expectations. In Proceedings of the 5th International Conference on Human Agent Interaction, 253–260. <https://doi.org/10.1145/3125739.3125766>.
- Zarouali, B., Van den Broeck, E., Walrave, M., & Poels, K. (2018). Predicting consumer responses to a chatbot on Facebook. *Cyberpsychology, Behavior and Social Networking*, 21(8), 491–497. <https://doi.org/10.1089/cyber.2017.0518>.
- Zlatanov, S., & Popescu, J. (2019). Current applications of artificial intelligence in tourism and hospitality. In Proceedings of International Scientific Conference on Information Technology and Data Related Research, Belgrade, Serbia, 84–90. <https://doi.org/10.15308/Sinteza-2019-84-90>.
- Zmud, R. W. (1978). An empirical investigation of the dimensionality of the concept of information. *Decision Sciences*, 9(2), 187–195. <https://doi.org/10.1111/j.1540-5915.1978.tb01378.x>.
- Zou, H. J. (2019). Results express: Ctrip's second quarter financial report. Retrieved from <https://www.36kr.com/p/5244947>