Noor, Nurhafihz, Rao Hill, Sally, and Troshani, Indrit (2022) ***Developing a service quality scale for artificial intelligence service agents.*** **European Journal of Marketing, 56 (5) pp. 1301-1336.**

https://doi.org/10.1108/EJM%2D09%2D2020%2D0672

Developing a service quality scale for artificial intelligence service agents

Nurhafihz Noor
Adelaide Business School
The University of Adelaide
hafihz@gmail.com

Sally Rao Hill
Adelaide Business School
The University of Adelaide
sally.raohill@adelaide.edu.au

Indrit Troshani
Adelaide Business School
The University of Adelaide
indrit.troshani@adelaide.edu.au

Accepted for publication, suggested citation:

Noor, N., Rao Hill, S. and Troshani, I. (2022) 'Developing a service quality scale for artificial intelligence service agents', European Journal of Marketing, doi: 10.1108/EJM-09-2020-0672.

Developing a service quality scale for artificial intelligence service agents

Abstract

Purpose – Service providers and consumers alike are increasingly adopting artificial intelligence service agents (AISA) for service. Yet, no service quality scale exists that can fully capture the key factors influencing AISA service quality. This study aims to address this shortcoming by developing a scale for measuring AISA service quality (AISAQUAL).

Design/methodology/approach – Based on extant service quality research and established scale development techniques, the study constructs, refines and validates a multidimensional AISAQUAL scale through a series of pilot and validation studies.

Findings – AISAQUAL contains 26 items across six dimensions: efficiency, security, availability, enjoyment, contact and anthropomorphism. The new scale demonstrates good psychometric properties and can be used to evaluate service quality across AISA, providing a means of examining the relationships between AISA service quality and satisfaction, perceived value as well as loyalty.

Research limitations/implications – Future research should validate AISAQUAL with other AISA types as they diffuse throughout the service sector. Moderating factors related to services, the customer and the AISA can be investigated to uncover the boundary conditions under which AISAQUAL is likely to influence service outcomes. Longitudinal studies can be carried out to assess how ongoing use of AISA can change service outcomes.

Practical implications – Service managers can use AISAQUAL to effectively monitor, diagnose and improve services provided by AISA, whilst enhancing their understanding of how AISA can deliver better service quality and customer loyalty outcomes.

Originality/value – Anthropomorphism is identified as a new service quality dimension. AISAQUAL facilitates theory development by providing a reliable scale to improve the current understanding of consumers' perspectives concerning AISA services.

Keywords Artificial intelligence service agents; service quality; scale development; customer service

Paper type Research paper

Introduction

Recent years have seen rapid advances in artificial intelligence (AI) technology development and the emergence of a plethora of applications powered by the technology (Huang and Rust, 2021, Rust, 2020, Davenport *et al.*, 2020). AI can be defined as technology (e.g., machine learning, big data, natural language processing and understanding) that enables software agents to "act intelligently" (Poole and Mackworth, 2010, p. 3). AI software agents are systems or machines that can complete tasks that typically require human intelligence (e.g., problem solving) in *rational* ways, achieving the best possible (expected) outcome, given the information available to them (Russell and Norvig, 2018, Poole and Mackworth, 2010).

There is a strong belief that AI is a key force in the expansion of the services industry and will have far-reaching and broad impacts on business (Huang and Rust, 2021, Rust and Huang, 2014, Noor *et al.*, 2021b, Fountaine *et al.*, 2021). Emerging evidence supports this belief (Colback, 2020, Chui *et al.*, 2020). AI offers business and service providers the potential to greatly boost revenue (e.g., by improving support for business and marketing decisions) and reduce operational costs (e.g., via automation) (Davenport *et al.*, 2020, Neuhofer *et al.*, 2020, Prentice *et al.*, 2020, Chui *et al.*, 2020). Strong market value growth forecasts for AI agents illustrate their potential (Makadia, 2020).

For example, the market value of chatbots and virtual assistants, two common types of AI agents used by business to provide service to consumers (Zarouali *et al.*, 2018, Hoy, 2018), is expected to increase at the compound annual growth rate of 33% between 2020 and 2025 (AMR, 2020). This business case is motivating innovative service providers to use AI agents to provide service to consumers over either part of or the entire customer service consumption journey (Prentice *et al.*, 2020, Oosthuizen *et al.*, 2020, Robinson *et al.*, 2020, De Keyser *et al.*, 2019, Paluch and Wirtz, 2020, Wirtz *et al.*, 2018).

In this study we are concerned with the broader *issue of measuring how consumers perceive the quality of service provided by AI service agents*. Following Zeithaml (1988), service quality can be defined as the overall excellence or superiority of the service performance by a service agent as perceived by consumers. Traditionally, the role of service agents has been performed by human service employees and/or technology-based agents, also known as self-service technologies (SSTs), which enable consumers to (co)produce service without the direct involvement of the employees (Meuter *et al.*, 2000, Bitner *et al.*, 2000, Grewal and Levy, 2009, Lin and Hsieh, 2011). AI service agents, henceforth AISA, constitute a new service agent type that is based on AI technology. Following Wirtz *et al.* (2018), we define AISA as "system-based autonomous and adaptable interfaces that interact, communicate and deliver service to an organization's customers" (p. 909). Examples of AISA include customer service chatbots, virtual assistants, social robots, and autonomous cars (Noor *et al.*, 2021a).

We argue that using AISA to provide service to consumers will change the nature of service encounters. A service encounter is the moment of interaction between a consumer and a firm (Bitner *et al.*, 2000) where the firm is traditionally represented by service agents. However, AISA are radically different to traditional, established service agents in that computer-based applications act in ways that emulate humans (Park *et al.*, 2021, Huang and Rust, 2021). The differences can contribute to different consumer experiences and consequentially dramatically alter consumers' perceptions of service quality (Lu *et al.*, 2020, Paluch and Wirtz, 2020, Huang and Rust, 2018, Verhagen *et al.*, 2014, Huang and Rust, 2021).

Service research in marketing has long established that "quality occurs during service delivery" (Zeithaml, Berry, and Parasuraman, 1988, p. 35). Service quality is thus widely seen as a key determinant of the service providers' long-term performance and success (Zeithaml *et al.*, 2002, Fassnacht and Koese, 2006, Parasuraman *et al.*, 1985, Zeithaml *et al.*, 1996a, Zeithaml *et al.*, 1988). Indeed, research looking at the role of consumers' perceptions of the quality of service offered electronically (e.g., via websites) found that such perceptions can affect a range of constructs, including customer satisfaction, loyalty, and attitudes towards ongoing use and recommendations, which in turn affect retailers' long term performance but were also found to be a key determinant of long term performance and success (Ladhari, 2010, Zeithaml *et al.*, 2002, Fassnacht and Koese, 2006, Cronin Jr and Taylor, 1994). Additionally, research from practice has shown that up to 42% of retail consumers show greater purchasing interest after experiencing good customer service, whilst 52% stopped purchasing due to a single disappointment in a customer service encounter (Zendesk, 2020).

There is limited understanding of how consumers' service encounters with AISA affect their perceptions of service quality and how the construct can be measured. This is problematic, particularly given the increasing number of service providers that are using or considering using AISA for service delivery (AMR, 2020). For example, Chui *et al.* (2020) report findings from a 2019 McKinsey survey indicating that 58% of surveyed businesses were using AI in at least one business function or unit, with product/service development and operations being the most commonly used functions, including AI-based enhancements, feature and operations optimization, and predictive service and interventions ranging between 19 and 24%. Similarly, Makadia (2020) cites a recent Gartner report forecasting that over half of established businesses are investing or considering investing in AI.

The business case for using AISA for service delivery is based on service providers' belief that consumers will accept and use AISA. However, evidence in the literature is mixed, challenging the foundation of this belief. For example, research discussing the use of AISA in service has argued that consumers should experience a range of benefits from the service provided by AISA, including greater efficiency and personalization (Huang and Rust, 2018, Huang and Rust, 2021, Verhagen *et al.*, 2014). These benefits are usually directly related to service quality which implies that they should improve the customers' perceptions of the service quality delivered by AISA. However, recent research that has empirically assessed the role of AISA in service shows mixed findings in terms of how consumers evaluate AI services. For instance, looking at the role of AI in the hospitality and hotel context, Prentice *et al.* (2020) find that, separately, both AI and employee service quality are significantly related to customer satisfaction. However, when the two constructs were considered jointly (i.e., regressed in the same equation), AI service quality had a negative effect on consumer satisfaction whilst employee service quality explains a significant part of the overall assessment of the quality of the service provided by both AI and service employees (Prentice *et al.*, 2020). We note that whilst Prentice et al. (2020) adopted 15 items from Makadia (2018) to represent a suite of AI-based services specific to the hotel context, there is no indication (see e.g., Makadia, 2018) how the items were generated and validated. More recently, research has also shown that how consumers perceive AI service quality also depends on the type of service. For instance, a recently published study by Park *et al.* (2021) shows that perceived AI service robots' usefulness is significant for credence services (e.g., at a hospital) but not significant for experience services (e.g., at a café). Overall, this evidence supports similar outcomes from AISA services that were anticipated by Wirtz *et al.* (2018) whilst also extending similar evidence by Elliot (2018) who also found that consumers prefer service employees to AISA in service encounters.

Taken together, this research suggests that there might be a *customer gap* between customers' AISA service needs and expectations and their perceptions of the quality of the services that are actually delivered by AISA (Prentice *et al.*, 2020, Bitner *et al.*, 2010). By implication, there is a need to improve the current understanding of consumers' perspective on AISA service quality (Prentice *et al.*, 2020, Lu *et al.*, 2020, Paluch and Wirtz, 2020, Huang and Rust, 2021).

Noor *et al.* (2021b) follow a qualitative approach to identifying twelve service quality dimensions representing AISA service quality. Ten of the proposed dimensions are adapted from key, established human- and technology-based service quality dimensions, with two dimensions reflecting AISA's unique underlying characteristics. Whilst Noor *et al.* (2021b) is a key study that makes important inroads into improving the understanding of AISA service quality, the proposed dimensions lack empirical, psychometric development and validation which limits their contribution to theory and practice.

We extend their work in this paper. Specifically, following on a strong tradition of established, proven scale-development processes and techniques (e.g., Churchill Jr., 1979, Netemeyer *et al.*, 2003), we build on Noor *et al.* (2021b)'s proposed qualitative dimensions to construct, refine and validate a multiple-item scale for measuring the AISA service quality of chatbots and virtual assistants, two types of AI service agents that are commonly used by service providers in the current market.

Our study culminates with an AISA service quality scale, henceforth AISAQUAL. For the purposes of this study, we follow Parasuraman *et al.* (2005) to define AISAQUAL as the extent to which AISA facilitate an overall perception of excellence or superiority as perceived by consumers. AISAQUAL is comprised of six dimensions, including 26 item measures. We thus contribute by providing a psychometrically validated, reliable service quality scale specifically developed for AISA. This scale can inform and help better explain theoretical findings and conclusions from emerging research in this area whilst also offering a new basis and opportunities for the development of new theoretical insights into the growing, novel domain of AI in service (Ranjan and Read, 2016). Specifically, the AISAQUAL scale can improve current understanding of the factors that *might* help reduce or bridge the *customer gap*, including service providers' understanding of customer expectations of AISA services, service design and standards, and service performance and delivery (e.g., Bitner *et al.*, 2010, Prentice *et al.*, 2020). The AISAQUAL scale also offers a robust tool for service providers to better understand the effective adoption of chatbots and virtual assistants as popular AISA in the competitive service sector and emergent issues and implications. Overall, AISAQUAL and related development work provides an important stepping stone towards addressing a critical shortcoming in the service quality literature in the area of AI in service whilst also forming a new platform for researching other types of AISA (Lu *et al.*, 2020, Bock *et al.*, 2020).

The paper is structured as follows. In the next section, we review related literature on service quality measurement and the recent research concerning AISA service quality. We then proceed to discuss the development of AISAQUAL, before concluding with the theoretical and managerial implications as well as directions for future research.

Related literature

Technology in service

Rapid technological advancements have seen the growing role of technology in service delivery (Bitner *et al.*, 2010). Service providers have used technology aimed at improving customer experiences. Broadly, technology has been used in two main ways (Bitner *et al.*, 2010). First, human service employees use technology to *support* interactions with consumers. In a support role, technology contributes by improving both the effectiveness and efficiency of the employees in service encounters with their customers. For example, service employees use technology to facilitate customer communication (e.g., email) and processing of historical transactional information to improve and personalize interactions with consumers (e.g., spreadsheets and database management systems). This form of technology use, where technology is instrumental to the fulfilment by service employees of service roles, is characterized by high interpersonal focus and contact between service employees and customers, and has been described as the *high (human) touch – low tech* paradigm (Verhagen *et al.*, 2014, Salomann *et al.*, 2007, Bitner *et al.*, 2000).

Second, service providers have also used technology to partially or completely *replace* human service employees. Technology-based service (i.e., via self-service technology) occurs when consumers use the technology independently, without the involvement of service employees, to provide service for themselves. Technology-based service helps consumers improve the effectiveness and efficiency of their *own* service encounter. Examples include bill payment portals, online shopping websites, self-service kiosks at airports and checkouts at department stores, and ATMs in banking. With greater infusion and use of technology, customers can access services conveniently anytime, anywhere, without the complications of interpersonal exchanges (e.g., bias, errors of service employees), but also in non-personal ways (Bitner *et al.*, 2000). This has been described as the *low (human) touch – high tech* paradigm (Verhagen *et al.*, 2014, Krehl, 2020, Bitner *et al.*, 2000).

Extant service quality and scale development research has predominantly focused on the high (human) touch – low tech and low (human) touch – high tech paradigms (Verhagen *et al.*, 2014). A plethora of scales have been developed to measure service quality for different types of traditional, established software agents including human employees and self-service technology operating in different contexts (e.g., Ladhari, 2010, Zeithaml *et al.*, 2002, Ladhari, 2009, Seth *et al.*, 2005).

Use of AISA in service encounters is associated with the emergence of a new *high (humanlike) touch – high tech* paradigm, where the AI technology provides *humanlike* interaction, that is, it offers scope for customers to interact independently with AISA which, in turn, emulate human interaction including undertaking rational approaches to problem solving, and relational communication (Russell and Norvig, 2018, Paschen *et al.*, 2020, Overgoor *et al.*, 2019, Verhagen *et al.*, 2014, Huang and Rust, 2021).

In the *high (humanlike) touch – high tech* paradigm, the AISA are radically different to traditional, established service agents in many ways (Lu *et al.*, 2020, Wirtz *et al.*, 2018). First, AI can efficiently process and use large volumes of structured and unstructured data from heterogeneous sources, including traditional CRM or ERP databases, email correspondence, social media, websites, location-based advertisements, and data types including text, audio, image and video (Dwivedi *et al.*, 2021, Mogaji and Erkan, 2019, Paschen *et al.*, 2020). Access and the power to process vast amounts of data enable AISA to solve service problems effectively and efficiently, often in ways superior to that of service employees (Bock *et al.*, 2020, Black and van Esch, 2020, Overgoor *et al.*, 2019, Beck and Libert, 2017).

Second, AISA can accurately, and efficiently profile consumers based on a broader range of criteria than traditional, established service agents can, which enhances service providers' understanding of customer preferences, consequentially creating the scope for effectively achieving personalization outcomes and managing relationships with consumers (Shankar, 2018). Third, AISA can also overcome the limitations of traditional human service agents, such as human judgement bias, fatigue, errors, and availability (i.e., by providing convenient service access to consumers anytime, anywhere) (Mogaji *et al.*, 2020, Wirtz *et al.*, 2018, Huang and Rust, 2018).

Fourth, AISA possess analytical and cognitive capabilities which make them capable of engaging consumers in a *humanlike* fashion (Wirtz *et al.*, 2018, Davenport *et al.*, 2020, Davenport and Ronanki, 2018). AISA can recognize human emotions, feelings and sentiments, but also provide socio-relational and -emotional responses resembling human emotions, such as empathy and compassion, when communicating with consumers, thereby addressing usability issues that are often associated with established technology-based service agents such as SSTs (Troshani *et al.*, 2020, Schniter *et al.*, 2020, Bitner *et al.*, 2002, Bitner *et al.*, 2000, Lu *et al.*, 2020, Wirtz *et al.*, 2018). Indeed, extant marketing literature shows that a functionalist approach to service delivery is necessary but not sufficient for achieving service quality (Bharadwaj *et al.*, 1993). Consumers place a premium on the socio-relational aspects that offer emotional and social value such as trust, engagement, and rapport (Stock and Merkle, 2018, Wirtz *et al.*, 2018). Early experimental research looking at virtual customer service agents, a form of AISA, found evidence that humanlike emotional responses exhibited by AISA such as friendliness, expertise, and smiling determine social presence and personalization which in turn influence consumers' satisfaction in the service encounter (Verhagen *et al.*, 2014).

Fifth, AISA have the ability to learn quickly from past interactions and historical information (Huang and Rust, 2018, Huang and Rust, 2021) which makes them capable of managing interactions in service encounters with greater scope and complexity relative to the specific, predefined interaction scenarios that are typically managed by service technologies such as SSTs (Lin and Hsieh, 2011, Orel and Kara, 2014, Considine and Cormican, 2016). For example, in their comparative analyses of SST and AISA, Wirtz *et al.* (2018) argue that SSTs follow rigid interaction scripts which are generally ineffective to recover from errors and service failure (e.g., Le *et al.*, 2020); by contrast, AISA have flexible interaction scripts and provide alternative solutions to consumers to recover from errors and failures.

For the purposes of this study, we define the *high (humanlike) touch – high tech paradigm* on the basis of the AI technology development and service use frameworks discussed in Wirtz *et al.* (2018) and Huang and Rust (2021). Specifically, the high (humanlike) touch – high tech paradigm includes types of AISA that can manage cognitive and analytical tasks that are associated with *low* socio-emotional and relational complexity (Wirtz *et al.*, 2018). Huang and Rust (2021) refer to these types of AISA as mechanical and thinking AI which are predominantly used for routine transaction-based services and utilitarian, data-based predictive services, respectively. We note that Huang and Rust (2021) also identify *feeling AI* which can manage feeling services, that is, cognitive and analytical service tasks associated with *high* socio-emotional and relational needs (e.g., high-risk healthcare service). Feeling AI, however, remains in the early stages of development with feeling services requiring the natural intelligence of human service agents who may rely on AISA for support (e.g., Wirtz *et al.*, 2018, Huang and Rust, 2021). Accordingly, for the purposes of this study, we exclude feeling AI (Huang and Rust, 2021) from the high (humanlike) touch – high tech paradigm.

In this study we use chatbots and virtual assistants as two types of AISA in the *high (humanlike) touch – high tech* paradigm. Both chatbots and virtual assistants are AISA that can autonomously support consumers or users. Chatbots are used by business to offer pre-defined services in response to customer queries (e.g., customer support) using mainly text-based, spoken language interactions (Whang and Im, 2021, Naveen, 2018). Virtual assistants provide general, customized, contextualized personal assistance services in response to user questions or commands (e.g., personal management) using natural voice interfaces (Youn and Jin, 2021, Kidd, 2019) (see Table I for additional detail).

[Insert Table I here]

In the sections that follow, we discuss the nature of the service quality construct, before providing an overview of service quality scale development research focusing on scale development for service encounters in the *high touch – low tech* and *low touch – high tech* paradigms. We conclude this section with a brief discussion of early AISA research, including the work of Noor *et al.* (2021b) and the proposed qualitative AISAQUAL dimensions. Taken together, these discussions form the basis for our development of the AISAQUAL scale focusing on chatbots and virtual assistants which will contribute to the emerging, but under-researched *high (humanlike) touch – high tech* paradigm.

Nature of service quality

Service quality is generally considered to be a long-term global judgement of service performance by consumers operationalized at an attitude level (Parasuraman *et al.*, 1994a, Cronin Jr and Taylor, 1994). There has been extensive discussion and rigorous validation of the service quality dimensions in various service settings. While some scholars have treated service quality dimensions as antecedents (e.g., Dabholkar, 1996, Dabholkar *et al.*, 2000), the majority conceptualize these dimensions as components of the multi-dimensional service quality construct (Brady and Cronin Jr, 2001).

Meanwhile the debate continues as to whether the construct is reflective as suggested in the majority of service quality scale development literature or whether it contains formative higher orders (e.g., Ladhari, 2009, Ladhari, 2010, Martínez and Martínez, 2010, Parasuraman *et al.*, 2005). Many scholars however, have advised caution in considering the formative specification (Hair *et al.*, 2018, Caro and Garcia, 2008).

In this study, we argue that measures of AISA service quality constitute more of a reflective view and are latent and not an index (Hair *et al.*, 2018). This is because the dimensions of the AISA service quality construct would be "expressions of the complexity" (Caro and Garcia, 2008, p. 716) of AISA service performance as perceived by consumers. Consistent with prior service quality measurement development studies, a series of service quality factors (e.g., availability, anthropomorphism) contribute to the overall AISA service quality perception and the change in any of the factors reflects the change in the overall service quality perception (Collier and Bienstock, 2009).

Another issue that has been debated in the service marketing literature concerns the empirical advantages of a performance-only assessment of service quality over the performance-expectations comparison approach (Cronin Jr and Taylor, 1992, Cronin Jr and Taylor, 1994, Parasuraman *et al.*, 1993, Parasuraman *et al.*, 1994b). In a longitudinal study, Dabholkar *et al.* (2000) concluded that a performance-only measurement of service quality is better than performance-expectation measurements, and is more suitable when the objective is to determine factors contributing to service quality. Similarly, in a meta-analysis study of 17 empirical service quality studies spanning 17 years, Carrillat *et al.* (2007) found no significant advantage to the use of performance-expectation indicators in determining overall service quality.

In addition, they highlighted the advantage of the performance-only method requiring half as many items as the performance-expectation approach. By contrast, the longer and more time-consuming disconfirmation measurement method is more suitable for a gap analysis (Dabholkar *et al.*, 2000), and offers more in-depth diagnostics of service quality (Carrillat *et al.*, 2007, Parasuraman *et al.*, 1993). As such, the perception-based measurement method was adopted for our study as it is most suitable for our purpose of developing AISA service quality scale dimensions (Dabholkar *et al.*, 2000).

Prior research has reached the consensus that service quality perceptions often affect service outcomes such as satisfaction, perceived value and loyalty (Zeithaml *et al.*, 1996b, Cronin Jr *et al.*, 2000, Prentice *et al.*, 2020). For example, in the context of self-service, it was found that self-service quality drives customer satisfaction and loyalty in e-retailing (Ding *et al.*, 2011). Similarly, customers' perception of self-service technology quality determines their behavioral intentions (Lin and Hsieh, 2011).

Service quality scales: research developments and limitations

Since its early conceptualization by Grönroos (1984), many service quality scales, predominantly grounded on the high (human) touch – low tech paradigm, have been proposed to help understand service quality (Seth *et al.*, 2005). Parasuraman *et al.* (1988)'s SERVQUAL model established a scale to measure service quality (Parasuraman *et al.*, 1991, Parasuraman *et al.*, 1988, Parasuraman *et al.*, 1994a) and has gained much popularity in the service marketing literature. This scale consists of five dimensions – tangibles, reliability, responsiveness, assurance and empathy – on which consumers base their assessment of the service quality of human service agents. The model has been validated and shown to be robust across many service industries, including education, healthcare, insurance, hotel, library, bank and retail services (Ladhari, 2009).

Since the development of SERVQUAL, technology has emerged and continued to play a growing, integral role in service delivery. The nature of the role of technology, however, began to substantially change from one where technology was used to support service employees to one where technology is used by service providers to replace them. Periodic leaps in technological innovation have resulted in the introduction of new technology-based SST service agents (Rust, 2020). These technologies include self-service machines, such as bank ATMs, vending machines (Fitzsimmons, 2003), and the internet with many online services (Yang *et al.*, 2004). Changes to the service environment due to technology-based service agents with unique interface designs and service delivery processes have affected the nature of service encounters including both the service and how service quality is perceived by consumers (Rust and Oliver, 1993), consequentially bringing about a shift to the low (human) touch – high tech service paradigm (Parasuraman, 2000, Bitner *et al.*, 2000).

There are different combinations of service quality dimensions applicable to different service environments, although some dimensions are more universal than others in terms of their perceived role by consumers in service quality. For example, responsiveness appears frequently as a key dimension for both human (Mittal and Lassar, 1996, Brady and Cronin Jr, 2001) and technology-based service agents (Ladhari, 2010). However, dimensions such as security and privacy are more salient for technology-based services such as websites (e.g., Yang *et al.*, 2004), mobile services (e.g., Huang *et al.*, 2015) and self-service technologies (e.g., Lin and Hsieh, 2011). Accordingly, scholars have developed different service quality scales for different contexts over the years including generic SSTQUAL scales for SSTs (Iqbal *et al.*, 2018, Considine and Cormican, 2016, Boon-itt, 2015, Lin and Hsieh, 2011, Ganguli and Roy, 2011). By contrast, other scales were designed more specifically to measure service quality for certain applications and environments, such as online shopping websites (e.g., E-S-QUAL) (Parasuraman *et al.*, 2005), technology-mediated transactions processing (eTransQual) (Bauer *et al.*, 2006), e-retailing self-service (e-SELFQUAL) (Ding *et al.*, 2011, Rita *et al.*, 2019), mobile service quality (MS-QUAL) (Huang *et al.*, 2015), and telematics applications (TeleServQ) (He *et al.*, 2017).

Taken together, service quality scale studies develop and validate scales that are anchored on *high (human) touch – low tech* or *low (human) touch – high tech* paradigms, featuring interactions of customers with human service personnel and with different technology-based SST service agents operating in different contexts and environments. This suggests that available scales are bounded by the nature of the context and technology, which makes their direct applicability to *high (humanlike) touch – high tech* settings questionable. For example, some scholars have attempted to assess the applicability of existing, established scales to AISA. For example, Morita *et al.* (2020) and Meyer-Waarden *et al.* (2020) found that the SERVQUAL scale was not suitable for measuring AISA service quality. AISA's inherent idiosyncrasies and differences from traditional service agents are likely to provide novel experiences and provoke unique reactions from consumers and consequentially also affect their perceptions of service quality, thereby affecting the relevance of existing service quality scales for AISA (Wirtz *et al.*, 2018, Bock *et al.*, 2020). This criticism is common of scale development studies that are developed for specific contexts, using specific applications (Ladhari, 2009, Ladhari, 2010).

Early AISA service quality research

There is limited research pertaining to AISA service quality. Two notable studies include Prentice *et al.* (2020) and Noor *et al.* (2021b). Focusing on the hospitality and hotel industry Prentice *et al.* (2020) adopt five constructs representing a suite of hotel services including services from concierge robots, digital assistance, voice-activated services, travel-experience enhancers and automatic data processing to measure AI service quality. The constructs and 15 items are sourced from Makadia (2018), without evidence of rigorous development and validation.

By contrast, Noor *et al.* (2021b) propose 12 dimensions which they argue represent the perceived service quality of AISA. Based on prior research into SERVQUAL dimensions development, they adopt a two-stage approach inclusive of a comprehensive review of service quality and information systems literatures followed by a qualitative validation stage. The second stage aimed to both validate identified dimensions and identify new dimensions that might be applicable to AISA service quality and included users of chatbots and virtual assistants, as common, popular types of AISA in the *high (humanlike) touch – high tech* paradigm. The 12 dimensions identified by Noor *et al.* (2021b) are summarized in Table II.

[Insert Table II here]

Of the 12 proposed dimensions representing the service quality of AISA, 10 are based on scales developed to capture service quality in high (human) touch – low tech and in low (human) touch – high tech settings. For instance, dimensions and items related to "assurance" (Parasuraman *et al.*, 1988, Burgers *et al.*, 2000) and "privacy" (Huang *et al.*, 2015) were adapted to describe the degree of security that AISA is perceived to provide to consumers.

However, other dimensions such as attitude (of service employees) from Brady and Cronin Jr (2001) were excluded as unsuitable.

Two of the proposed dimensions, namely proactiveness and anthropomorphism, were new to the service quality literature. Specifically, *proactiveness* refers to the extent AISA can anticipate or predict consumers' future needs, and prompt service beyond what is explicitly required by a customer (Noor *et al.*, 2021b). Proactiveness is grounded in the customer service literature on the construct of proactive behavior of employees which is characterized by personal initiative criteria (Frese and Fay, 2001) including being "self-started, long-term-oriented, and persistent behavior that is organizationally function and goal directed" (Rank *et al*, 2007, p. 366). AISA can be said to be proactive when it manifests the capability of anticipating and going beyond a straightforward reaction to commands, that is, alerting consumers about tasks they may have overlooked or of which they may not be aware due to limited familiarity with the service context (Rank *et al.*, 2007), including assisting consumers with advice concerned with alternative courses of action (Tan and Chou, 2008).

Anthropomorphism is another key characteristic distinguishing AISA from other non-AI technology-based service agents. It describes the attribution of human capacities to non-human agents (Troshani *et al.*, 2020, Bartneck *et al.*, 2009, Moussawi, 2016, Epley *et al.*, 2007). Prior research suggests that a consumer who interacts with AISA exhibiting anthropomorphic capacities, including humanlike features, may develop perceptions of a social presence that reduce privacy concerns (Benlian *et al.*, 2020) and increase trust in the use of AISA (Qiu and Benbasat, 2009, Troshani *et al.*, 2020) and potentially enhance service quality perceptions (Noor *et al.*, 2021b). Consumers may also experience negative feelings when they realize the service they have received, also known as "counterfeit service" was from AISA rather than human service employees (Robinson *et al.*, 2020). Additionally, increasing anthropomorphism beyond a certain point is likely to lead to feelings of discomfort and uneasiness (Duffy, 2003, Noor *et al.*, 2021b, Troshani *et al.*, 2020, Kim *et al.*, 2019).

The current study builds on the work of Noor *et al.* (2021b). To the best of our knowledge, Noor *et al.* (2021b) are the only available study making important inroads into AISA scale development research by qualitative validation. Accordingly, we adopt their proposed dimension (Table II) to develop AISAQUAL. We discuss the scale development process that we adopted in the sections that follow.

Scale development

Consistent with the psychometric procedure for scale development advocated by marketing scholars, our scale development process consists of domain definition, item generation, scale development, and scale validation (Churchill Jr., 1979, Netemeyer *et al.*, 2003). The first phase involves defining the domain and phenomenon to be measured based on extant literature followed by the generation of a pool of items through literature reviews, interviews, and domain experts' input, in conjunction with an assessment of content validity of the items. The scale development step requires selecting and categorizing appropriate items to establish desirable reliability and validity. The final step involves scale validation in which we evaluate the scale to ensure desirable psychometric properties. The process culminates with six AISAQUAL dimensions comprised of 26 items which, whilst grounded in extant related scale validation and emerging AI literature, are also customized specifically for the AISA context. We frame AISAs in the novel *high (humanlike) touch – high tech paradigm* and the domain in which they operate is the broader area of service delivery where they trigger consumer service quality perceptions. So the scale development process capitalizes on the rich base of service quality literature and adaptation for AISA. A summary of the process we followed in the study is shown in Figure 1. Each step is elaborated in the sections that follow.

[Insert Figure 1 here]

Step 1: phenomenon and domain definition

Following Zeithaml (1988), we define AISAQUAL as the extent to which AISA facilitate an overall perception of excellence or superiority as perceived by consumers. It is a global assessment and attitude consisting of consumer judgements of AISA service performance. We adopt the dimensions proposed by Noor *et al.* (2021b) for which we identify measure items at the perceptual level to effectively capture the abstract nature of service quality comparisons which consumers make across categories (Zeithaml, 1988). These service quality perceptions can be formed through AISA usage, and in turn affect various outcomes such as customer satisfaction and loyalty (Cronin Jr *et al.*, 2000).

Following Noor *et al.* (2021b), we also use chatbots and virtual assistants in this study as common forms of AISA used by service providers. We argue that this choice is appropriate due to the popularity and wide availability of these types of AISA, which is important in facilitating the recruitment of participants with experience in the use of these types of AISA, therefore providing access to meaningful responses for the subsequent stages of scale development and validation.

Step 2: item generation

To generate the item pool for Noor *et al.* (2021b)'s 12 conceptual dimensions, we adopted key established scales from the literature. As shown in Figure 2, measure items from service quality scales in human service contexts were found to generally capture half of the 12 conceptual AISAQUAL dimensions (i.e., reliability, responsiveness, availability, aesthetics, personalization and security), whereas item measures from technology-based service quality scales were able to tap into all dimensions except proactiveness and anthropomorphism.

[Insert Figure 2 here]

As proactiveness and anthropomorphism represented two new dimensions (Noor *et al.*, 2021b), a logical way to analyze the quality of AISA service is to examine what proactiveness and anthropomorphism entail which are mostly discussed in information and system quality literature. Thus, we turned to non-service quality scale studies which contained similar constructs that may capture these dimensions. Specifically, item measures for proactiveness were adapted from Rank *et al.* (2007), whereas those for anthropomorphism were from Bartneck *et al.* (2009), Han and Yang (2018) and Moussawi (2016). Further, based on (Noor *et al.*, 2021b)'s qualitative evidence, additional measure items were introduced to better capture the intended dimensions (Netemeyer *et al.*, 2003). For example, one of the items "the AISA uses its own 'judgment' to complete a task" was added from (Noor *et al.*, 2021b)'s findings. After screening for irrelevant, redundant, ambiguous and double-barrelled statements, an initial battery of 85 items was produced.

To assess the item statements, we recruited a panel of six senior expert academics with an extensive publication track record in service and in emerging AI in business research. These academics are active researchers based in top-ranked universities in Australia (3), New Zealand (1), Singapore (1), and the USA (1). To maintain confidentiality, we do not disclose the identity of the academics.

Through an online questionnaire sent to each expert via Qualtrics, item statements were assessed for their representativeness (i.e., content validity) and appearance to be relevant (i.e., face validity) to the target construct (Netemeyer *et al.*, 2003). The following scale was used:

1="not representative", 2="somewhat representative", 3="clearly representative" (Bearden *et al.*, 2001). Only items which scored 2 (i.e., "somewhat representative") or 3 (i.e., "clearly representative") by at least 80 percent of the panel were retained in the item pool (Lin and Hsieh, 2011). We also reviewed the qualitative suggestions that experts made to improve item wording for inclusion in the item pool. In addition, we assessed the remaining items for redundancy and added measures to ensure a sufficient item pool per dimension in the subsequent scale refinement process (Netemeyer *et al.*, 2003). This content and face validation process reduced the number of items from 85 to 75.

Step 3: scale refinement

The next phase of scale development involved the testing of the preliminary 75-item AISAQUAL scale. A self-administered questionnaire was constructed and consisted of two sections. The first section contained demographic and AISA usage questions. The second section contained statements for 75 items to measure respondents' perceptions of AISA performance. Items for the second section were measured using a seven-point Likert scale anchored from 1 = strongly disagree to 7 = strongly agree, which is an established practice for scale reliability and validity (Netemeyer *et al.*, 2003).

Surveys were distributed by the online panel company Qualtrics using purposive sampling to users of chatbots and virtual assistants. A condition of participation in the survey was to have used these service agents in the three months prior to the survey. We asked respondents to choose the AISA type (i.e., chatbots or virtual assistants) with which they were most familiar. To clarify our definitional differences between chatbots and virtual assistants and establish a common language to ensure applicability of responses, examples of each type of AISA were provided in the survey introduction, including illustrating images, basic definitions, and textual descriptions of common uses. A US sample was deliberately used since the US represents one of the top 10 countries with a significant number of AISA users (PwC, 2018), and is expected to continue to occupy the largest global market share of chatbots and virtual assistants (AMR, 2020).

The sample consisted of 211 respondents with an almost even gender split within each category of chatbot (male=50.5%, female=49.5%) and virtual assistant (male=50.9%, female=49.1%) users. The sample size of 211 complied with requirements of approximately 200 for an initial test stage of a new scale (Clark and Watson, 1995, Parasuraman *et al.*, 1988). Almost two-thirds of respondents (67.3%) were aged 25 to 44. Table III summarizes the profiles of respondents for the scale refinement phase.

[Insert Table III here]

Exploratory Factor Analysis

As recommended by Churchill Jr. (1979), to better prepare the core items for the exploratory factor analysis, we first categorized the 75 item measures into the 12 a priori conceptual dimensions of AISA service quality before examining the reliability Cronbach Alpha score for each dimension. We then inspected items with low individual reliabilities (<.50) (e.g., Bagozzi and Yi, 1988) and removed one from the security dimension and the other from anthropomorphism.

Next, a principal components analysis with oblimin rotation (Kaiser normalization) was conducted on the remaining 73 items to empirically identify the underlying dimensions. Consistent with extant service quality scales (e.g., Parasuraman *et al.*, 1988, Parasuraman *et al.*, 2005, Ding *et al.*, 2011), we used the oblique rotation – oblimin – to allow for correlations between factors in order to obtain interpretable components. We note that several factor correlations after oblique rotation exceeded the suggested threshold of 0.32 by Tabachnick *et al.* (2007).

The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy (.95 > .50) and Bartlett's sphericity test (p < .001) were significant and indicated the suitability of using exploratory factor analysis for our data (Field, 2013). To determine the number of components, the Kaiser criterion of eigenvalues greater than one was used (Netemeyer *et al.*, 2003). Next, items were dropped using a minimum loading criterion of .40 (Ford *et al.*, 1986) or those exhibiting cross-loadings over .40 on two or more components (Hair *et al.*, 2018). Remaining items were again subjected to factor analysis.

The above process was done iteratively using SPSS 25 till all items and dimensions satisfied the required minimum thresholds. After six extractions, 34 items remained and loaded distinctly onto six dimensions D1 to D6 as shown in Table IV. The six components accounted for 66.9% of the variance (Hair *et al.*, 2018) and indicated good internal consistency among items with reliability coefficients ranging from .85 to .90 (Nunnally, 1978).

[Insert Table IV here]

As shown in Table IV, dimensions D2, D4 and D5 remained reflective of the three original dimensions of security, enjoyment and contact respectively. An exception was the security item SEC1 "A clear privacy policy is accessible before I use the AISA" which loaded with contact items in D5. Conceptually, consumers may relate the availability of such a privacy policy as originating from contact with a human service professional. Removing SEC1 would also marginally lower the coefficient alpha of D5. Thus, at this stage, SEC1 was retained with the other contact items in D5.

The remaining nine dimensions from the original 12 were collapsed into three. D1 – containing several items for reliability, responsiveness, aesthetics and control – was found to be similar in its conceptual item composition to the dimension of efficiency in E-S-QUAL (Parasuraman *et al.*, 2005) and functionality in SSTQUAL (Lin and Hsieh, 2011). D3 – mostly containing items for availability – included an original responsiveness item RES3 which can be conceptually related to the perceived availability of AISA to launch quickly when required. In relation to the loading of an aesthetics item AES1 in D3, conceptually an innovative interface design may signal the competency of AISA as being capable of service anytime. Similar to SEC1, the removal of RES3 and AES1 would lower the coefficient alpha of D3. Thus, we kept these items for further empirical scrutiny in the subsequent CFA stage. Finally, D6 contained items related to anthropomorphism, personalization and proactiveness. This composition of human and intelligence performance traits expected by AISA consumers is significant given that AISA is a non-human service agent and past scales involving technology service agents have not featured a service quality dimension of a similar nature.

Confirmatory Factor Analysis

A confirmatory factor analysis (CFA) of the 34-item, six-dimension model was next conducted to further verify the model. A sample of 275 chatbot and virtual assistant users was used to conduct this analysis. Just over half of the respondents are female (52%). The majority of the respondents are from the age group of 25-34 (33.0%) and 35-44 (33.4%). The profile of the sample can be found in Table III.

Using indices recommended by Bagozzi and Yi (1988), initial CFA results indicated a significant chi-square value (χ^2 = 1077.43, p < .001); with a Root Mean Square Error of Approximation (RMSEA) = .09 (recommended RMSEA ≤ .07), Tucker and Lewis Index (TLI)

= .91 (recommended TLI ≥ .90), Comparative Fit Index (CFI) = .87 (recommended CFI ≥ .90) and the Standardized Root Mean Square Residual (SRMR) = .10 (recommended SRMR ≤ .08).

To improve the model fit, we first looked at the item-to-factor loadings and removed items with loading values below .70 (Hair *et al.*, 2018). Two iterations removed items CTL3, SEC1, PRO5 and RES2 which had item-to-factor loading values of .69, .65, .52 and .69, respectively. Next, we inspected the standardized residual covariance matrices. Although the standardized residual values were less than 2.5, there were observable patterns of fairly large standardized residual loadings across several variables that were worthy of closer inspection (Hair *et al.*, 2018). Accordingly, we removed residuals greater than 2 (Anderson and Gerbing, 1988, Bagozzi and Yi, 1988). This iterative process deleted items AES1, REL7, SEC5 and RES3, and resulted in an acceptable model fit (Bagozzi and Yi, 1988). The final confirmatory model contained six factors and 26 items (see Table IV for items in bold) with values of χ^2 = 835.01, $p < .001$; RMSEA = .07 (≤ .07), TLI = .91 (≥ .90), CFI = .90 (≥ .90) and SRMR = .08 (≤ .08). Based on the content of the items in each dimension, six labels and definitions were chosen below. Of these, we kept five from the initial 12 conceptual dimensions with efficiency combining several attributes.

1. *Efficiency*: AISA's ease of use and speed.
2. *Security*: Perceived safety of the AISA from intrusion of privacy, fraud and loss of personal information.
3. *Availability:* the extent to which AISA is ready for use anywhere, anytime.
4. *Enjoyment*: Extent to which using the AISA is perceived to be enjoyable regardless of consequential performance.
5. *Contact*: Access to human assistance is available via AISA.
6. *Anthropomorphism*: AISA display human-like characteristics, motivations, intentions, or emotions.

In comparison with existing human service agents' service quality scales, the dimensions of security and availability are found in both human- as well as technology-based service quality scales. The dimensions of efficiency, enjoyment and contact are not found in human-based service quality scales but instead in extant technology-based service quality scales. The remaining anthropomorphism dimension represents a new service quality dimension unique to AISA service quality.

Step 4: scale validation

Additional empirical research was conducted to confirm the reliability and validity of the 26-item AISAQUAL scale. A self-administered questionnaire includes the 26 AISAQUAL items and service outcome variables. The new sample consisted of 304 respondents which was larger than the pilot sample size (n=211) for the scale refinement phase and satisfied the requirements for scale validation (Clark and Watson, 1995). The sample profile is similar to the sample used in the scale refinement phase (MacKenzie *et al.*, 2011) and consists of US residents who had used chatbots and virtual assistants in the three months prior to the survey. There was an almost even gender split within each category of chatbot (male=50.3%, female=49.7%) and virtual assistant (male=49.7%, female=50.3%) users. Almost an equal portion used their AISA on a daily (31.3%) or weekly (31.9%) basis. In addition, the majority of respondents (47.4%) had used their AISA for two to three years. In terms of usage context, most respondents used their chatbots for services related to the retail trade (19.9%). This differs from those in the scale refinement phase who predominantly interacted with chatbots for electricity, gas and waste services (20.0%). For virtual assistants, similar to respondents from our scale refinement phase, Google Assistant, Alexa and Siri were the most popular. This is also representative of the overall US customer adoption for virtual assistants (Olson and Kemery, 2019). Table V summarizes the profiles of respondents for the scale validation phase.

[Insert Table V here]

The descriptive statistics of the dimensions have been provided in Table VI. In addition, we conducted *t*-tests to detect if the mean scores of the dimensions differ between virtual assistants and chatbots. The results suggest invariance of AISAQUAL dimensions across the two types of AISA. To assess the proposed scale's construct validity, we deployed several methods. First, we conducted a CFA (χ^2 = 562.94, $p < .001$; RMSEA = .06, TLI = .94, CFI = .95 and SRMR = .04). All coefficient alphas are above the .70 level (Hair *et al.*, 2018) except for item EFF1 "The AISA works correctly at first attempt" which had a loading of .69. Upon inspection, EFF1 contributed to the content validity of its latent variable and its removal did not result in an increase in the composite reliability (CR) score of the efficiency dimension (see Table VI) (Hair *et al.*, 2011). Its item loading also fell within the acceptable range of between .50 to .90 (Bagozzi and Yi, 1988) and was not below the absolute threshold value of .50 (Hair *et al.*, 2018). Thus, EFF1 displayed sufficient indicator reliability and was retained in the scale.

To assess convergent validity, the composite reliability (CR) of all six dimensions of AISAQUAL was found to be between .80 and .90 which is above the recommended value of .70 (Bagozzi and Yi, 1988). In addition, the average variance extracted (AVE) of the six factors also ranged from .57 to .68 which is above the recommended level of .50 (Bagozzi and Yi, 1988), indicating high levels of convergence among the items in measuring their respective constructs.

[Insert Table VI here]

Following the recommendations of Voorhees *et al.* (2016), the Hetero-Trait Mono-Trait (HTMT) Ratio of the correlations was used to test for discriminant validity (Henseler *et al.*, 2015). All ratios were found to meet the conservative cut-off of .85 (Hair *et al.*, 2018) except for those between Efficiency and Enjoyment and between Efficiency and Anthropomorphism (see Table VII). Their correlation ratio of .86 was within the acceptable threshold of .90 (Zheng *et al.*, 2020) and supported for conceptually similar constructs (Hair *et al.*, 2018). In addition, we examined the discriminant validity by comparing the correlation between the dimensions and the AVE of each dimension (Fornell and Larcker, 1981). As shown in Table VIII, the square roots of AVE ranges between .75 and .82, exceeding all the correlations squared coefficients. This shows that AISAQUAL has satisfactory discriminant validity (Boudreau *et al.*, 2001).

[Insert Table VII here]

[Insert Table VIII here]

To assess the nomological validity of AISAQUAL, we test the relationship between the six dimensions of AISAQUAL and three theoretically related variables of customer satisfaction, perceived value and customer loyalty. This method is used to demonstrate the proposed scale's practical value and indicates its ability to explain and predict other dependent variables (Arnold and Reynolds, 2003). Service quality has been shown to affect customer satisfaction (Ding *et al.*, 2011, Caruana, 2002) and perceived value (Parasuraman *et al.*, 2005), and that all three constructs work together to affect the behavioral outcome of customer loyalty (Cronin Jr *et al.*, 2000, Oh, 1999). To assess these relationships in the AISA context, five loyalty intention items were adapted from Zeithaml *et al.* (1996a), three customer satisfaction items from Bodet (2008), and two perceived value items from Tam (2004) and Cronin Jr *et al.* (2000). AISAQUAL was modelled as an exogenous variable by aggregating its six dimensions into six indicators using the average score of items per dimension (Ding *et al.*, 2011, Lin and Hsieh, 2011).

The structural model illustrated in Figure 3 shows a good model fit (χ^2 = 260.87, p < .001; RMSEA = .07, TLI = .96, CFI = .97 and SRMR = .03). All loadings of these paths were found to be significant at p < .001. The effect of AISAQUAL was strongest on perceived value (β=.82, p < .001) followed by satisfaction (β=.34, p < .001) and loyalty intentions (β=.17, p < .001). These path-strength patterns echo findings from extant research in both services marketing (e.g., Cronin Jr *et al.*, 2000) and IS (e.g., Kuo *et al.*, 2009). In addition, the results indicate that all extracted service quality dimensions have significant effects on customer satisfaction, perceived value and customer loyalty. AISA service quality accounts for 41% of the variance in customer satisfaction, 38% of the variance in perceived value, and 23% of the variance in customer loyalty, indicating good external validity. This pattern of evidence suggests that AISAQUAL demonstrates nomological validity.

[Insert Figure 3 here]

In addition to our theoretical reasoning regarding AISAQUAL's reflective model structure, we empirically tested our model specification using a confirmatory tetrad analysis (CTA-PLS) (Hair *et al.*, 2019). The CTA-PLS has been used in recent marketing studies (e.g., Nath, 2020) to assess if the difference between pairs of covariances among construct indicators (i.e. tetrads) is significantly different from zero which would indicate a formative construct. A reflective model would result in tetrads having a value of zero. Using the recommendations outlined by Bollen and Ting (2000) and Gudergan *et al.* (2008), we find that all tetrad results were non-significant (i.e. confidence intervals include zero), providing empirical support for our view that AISAQUAL is a reflective construct.

Finally, we conducted a multi-group analysis (Henseler, 2012, Henseler *et al.*, 2009) to check if the proposed scale varies across the two types of AISA. Results for the difference between group-specific path coefficients are non-significant at the 5 percent probability of error level (see Table IX), suggesting that AISAQUAL demonstrates sufficient invariance across chatbots and virtual assistants.

[Insert Table IX here]

Concluding discussion and implications

Theoretical implications

The growing research interest in services enabled by AISA and the continued use of AISA in service sectors makes urgent the development of a suitable AISA service quality scale (Lu *et al.*, 2020, Bock *et al.*, 2020). Our study is a direct response to this need. Specifically, our development of AISAQUAL appears to be the first AISA service quality scale for AI-based applications in the emergent *high (humanlike) touch – high tech* paradigm, validated via rigorous, established scale validation processes using two popular, common AISA, namely chatbots and virtual assistants. Consisting of six dimensions and 26 item measures, AISAQUAL fills an important gap and extends current understanding of consumer service quality evaluations for different AISA service environments using validated and generalizable scale instruments.

Our findings suggest that anthropomorphism is a key dimension driving AISA service quality (β=.89, p < .001), thereby providing support for the emerging research underscoring the importance of further considering anthropomorphism in the improvement of user experiences with AISA (Wirtz *et al.*, 2018, Benlian *et al.*, 2020, Troshani *et al.*, 2020, Sheehan *et al.*, 2020).

Our findings also suggest that the hedonic element of AISA, i.e., enjoyment, is important to consumers in the evaluation of service quality (β=.87, p < .001). This is consistent with the findings of Lin and Hsieh (2011) and supports the role enjoyment plays in service quality evaluations of AISA beyond factors that help fulfil customers' utilitarian needs (Davis *et al.*, 1992). Importantly, we also identified efficiency as an important dimension of AISAQUAL (β=.91, p < .001) which is consistent with and further strengthens the view that AISA provide utilitarian value to consumers (Meyer-Waarden *et al.*, 2020, Noor *et al.*, 2021b). The presence of both hedonic and utilitarian dimensions also support the strategic importance of AISA in optimising quality and relationships in its service of consumers (Huang and Rust, 2021).

Taken together, the AISAQUAL dimensions capture what consumers perceive to be important when using services provided by AISA. Prentice *et al.* (2020) have argued that the disconnect, i.e., *customer gap*, between consumers' AISA service needs and expectations and their perceptions of the service quality actually delivered by AISA can explain the findings in research that consumers often prefer human service employees to AISA. Our AISAQUAL scale extends current understanding of the key factors that *might* contribute towards the *customer gap* in the context of AISA-based services.

Additionally, these findings add to the emerging research stream of AISA and validate relationships between service quality and customer satisfaction, perceived value and loyalty intentions in the AISA context. The findings of the current study suggest how the service quality of AISA can be determined, and as such will facilitate further theory development through the use of AISAQUAL (Ranjan and Read, 2016). This is important as it extends the link between service quality perception and positive service outcomes in the AISA context.

Managerial implications

The growing popularity of novel AI-based applications such as AISA (Davenport and Ronanki, 2018) increases the onus on service providers to effectively design AISA that customers will use for service. Our contribution of the AISAQUAL scale has practical and managerial relevance. It helps to improve service providers' understanding of consumer expectations when using AISA. The factors may also be useful to inform further efforts of service providers to improve AISA design, putting them in a stronger position to ensure that the manner in which the AISA they use to provide service are consistent with consumers' service quality expectations (Prentice *et al.*, 2020).

For example, given the significance of the anthropomorphism and enjoyment dimensions, companies should take these factors into consideration in the design phase of their AISA interface with their customers. One way to do this is to test different designs and interaction modes (e.g., speech *versus* gesture) of the AISA with their target audience at multiple stages (Kepuska and Bohouta, 2018). Our study also highlights the prevailing relevance of contact with human service agents in the eyes of consumers. Specifically, managers should ensure that human service agents are an available option for consumers during their AISA interaction (Shell and Buell, 2019). On AISA security and governance (Shepardson, 2020), companies need to foster greater trust and transparency with users (Bandara *et al.*, 2020) by being forthcoming with regards to their data privacy and protection policies as they continue to access increasing personal data through AISA. Finally, developers can increase the availability of their AISA within the wider service ecosystem as device interconnectivity matures through the Internet of Things (IoT) (Huang and Rust, 2017).

Consistent with prior service quality research, AISAQUAL validates the importance of service quality for AISA consumers as this can lead to perceived value, customer satisfaction and loyalty. AISAQUAL can thus serve as a diagnostic tool to improve current AISA service performance. As the AISAQUAL measurement is parsimonious, it helps service managers to better understand customer perceptions and address service quality concerns in a systematic way.

Limitations and future research

Our study contributes to the rich service quality literature by developing a robust service quality scale with good psychometric properties to accommodate the new AISA service environment. However, as with any scale development study, several caveats should be noted which also represent opportunities for further research.

First, AISAQUAL is developed as both a first and second order reflective construct based on underlying theoretical considerations (MacKenzie *et al.*, 2011) and empirical testing. While we have established this position in our study, future research can explore the implications of an alternative formative model (Collier and Bienstock, 2009, Theodosiou *et al.*, 2019) which requires additional reflective indicators to be tested against the proposed formative constructs (Jarvis *et al.*, 2003).

Second, at the time of this study, the AISA types most widely used by consumers are chatbots and virtual assistants. Accordingly, these software applications were used as suitable representative technology types for the construction of our AISAQUAL scale. An inspection of the measure items of the six dimensions of AISAQUAL suggests that the item measures might also be applicable to other AISA types including social robots (Noor *et al.*, 2021a). We recommend that the AISAQUAL be adapted and tested for these other AISA types, which have more physical features than chatbots and virtual assistants, as they become available in the market.

Third, as previously indicated, anthropomorphism is a key dimension in our AISAQUAL scale. Our evidence clearly suggests the important role that anthropomorphism plays in consumers' AISA service quality perceptions and evaluations. Whilst existing literature has associated anthropomorphism of AI with the *uncanny valley phenomenon*, we did not find evidence to indicate existence of this phenomenon in the context of AISA. AISAQUAL indicates that consumers perceive anthropomorphism as a favorable dimension rather than one which may cause uneasiness when interacting with AISA. Our findings both raise questions concerning the extent of applicability of uncanny valley to AISA applications, and also underscore the need for further research to enhance the context- and application-specific understanding of anthropomorphism.

Fourth, the focus of this study was to produce AISAQUAL, a service quality scale for improving current understanding of the key factors that influence consumers' perception of AISA service quality. The scale does not differentiate between customers' AISA service expectations and perceptions. Accordingly, further research is needed to identify the key factors that influence the AISA customer gap (Bitner *et al.*, 2010) including specific measurements of what consumers expect for services provided by AISA and their perceptions pertaining to services that are delivered by AISA.

Fifth, key steps in the scale development and validation process were carried out by using respondents that were recruited via the Qualtrics platform. This platform has become increasingly popular in recent years for respondent recruitment in research. Further research may further validate our findings across other types of platforms.

Finally, the effects of AISAQUAL on the outcome variables used in this study was based on a cross-sectional view. Additionally, given our focus on scale development, we have not considered moderating factors. As AISA are expected to serve consumers in the long run, understanding possible shifts in attitudes over time (Hussain *et al.*, 2019) and whether these may lead to more positive or negative outcomes is worth further investigation. Hence, we recommend future longitudinal studies assess how the ongoing service performance of AISA can change consumer outcomes including consideration of moderating factors in relation to the AISA application (e.g., the form and type), the service (e.g. types, criticality), and the customer (e.g., demographics, cultural values, usage situation, orientation, and technology savviness and proneness).

References

AMR (2020) *Intelligent Virtual Assistant (IVA) Market to grow at 33% CAGR during forecast period (2020-2025) - Insights on Growth Drivers, Size and Share Analysis, Key Trends, Leading Players, and Business Opportunities: Adroit Market Research. Available: https://www.globenewswire.com/news-release/2020/02/24/1988963/0/en/Intelligent-Virtual-Assistant-IVA-Market-to-grow-at-33-CAGR-during-forecast-period-2020-2025-Insights-on-Growth-Drivers-Size-and-Share-Analysis-Key-Trends-Leading-Players-and-Busin.html. Accessed April 19th, 2021.* Dallas, TX: Adroit Market Research.

Anderson, J.C. and Gerbing, D.W. (1988), "Structural equation modeling in practice: A review and recommended two-step approach". *Psychological Bulletin,* Vol. 103 No. 3, pp. 411.

Arnold, M.J. and Reynolds, K.E. (2003), "Hedonic shopping motivations". *Journal of Retailing,* Vol. 79 No. 2, pp. 77-95.

Bagozzi, R.P. and Yi, Y. (1988), "On the evaluation of structural equation models". *Journal of the academy of marketing science,* Vol. 16 No. 1, pp. 74-94.

Bandara, R., Fernando, M. and Akter, S. (2020), "Managing consumer privacy concerns and defensive behaviours in the digital marketplace". *European Journal of Marketing,* Vol. 55 No. 1, pp. 219-246.

Bartneck, C. *et al.* (2009), "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots". *International journal of social robotics,* Vol. 1 No. 1, pp. 71-81.

Bauer, H.H., Falk, T. and Hammerschmidt, M. (2006), "eTransQual: A transaction process-based approach for capturing service quality in online shopping". *Journal of Business Research,* Vol. 59 No. 7, pp. 866-875.

Bearden, W.O., Hardesty, D.M. and Rose, R.L. (2001), "Consumer self-confidence: Refinements in conceptualisation and measurement". *Journal of Consumer Research,* Vol. 28 No. 1, pp. 121-134.

Beck, M. and Libert, B. (2017), "The rise of AI makes emotional intelligence more important. Available: https://hbr.org/2017/02/the-rise-of-ai-makes-emotional-intelligence-more-important. Accessed April 20th, 2021". *Harvard Business Review,* Vol. February No.

Benlian, A., Klumpe, J. and Hinz, O. (2020), "Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: A multimethod investigation". *Information Systems Journal,* Vol. 30 No. 6, pp. 1010-1042.

Bharadwaj, S.G., Varadarajan, P.R. and Faly, J. (1993), "Sustainable competitive advantage in service industries: a conceptual model and research propositions". *Journal of Marketing,* Vol. 57 No. 4, pp. 83-99.

Bitner, M.J., Brown, S.W. and Meuter, M.L. (2000), "Technology infusion in service encounters". *Journal of the Academy of marketing Science,* Vol. 28 No. 1, pp. 138-149.

Bitner, M.J., Ostrom, A.L. and Meuter, M.L. (2002), "Implementing successful self-service technologies". *Academy of Management Executive,* Vol. 16 No. 4, pp. 96-109.

Bitner, M.J., Zeithaml, V.A. and Gremler, D.D. (2010) 'Technology's impact on the gaps model of service quality'. In: P., M., Kieliszewski C. and J., S. (eds.) *Handbook of*

Service Science: Research and Innovations in the Service Economy. Boston, MA: Springer, pp. 197-218.

Black, J.S. and van Esch, P. (2020), "AI-enabled recruiting: what is it and how should a manager use it". *Business Horizons,* Vol. 63 No. 2, pp. 215-226.

Bock, D.E., Wolter, J.S. and Ferrell, O. (2020), "Artificial intelligence: disrupting what we know about services". *Journal of Services Marketing,* DOI 10.1108/JSM-1101-2019-0047.

Bodet, G. (2008), "Customer satisfaction and loyalty in service: Two concepts, four constructs, several relationships". *Journal of Retailing and Consumer Services,* Vol. 15 No. 3, pp. 156-162.

Bollen, K.A. and Ting, K.F. (2000), "A Tetrad Test for Causal Indicators". *Psychological Methods,* Vol. 5 pp. 3-22.

Boon-itt, S. (2015), "Managing self-service technology service quality to enhance e-satisfaction". *International Journal of Quality and Service Sciences,* Vol. 7 No. 4, pp. 373-391.

Boudreau, M.-C., Gefen, D. and Straub, D.W. (2001), "Validation in information systems research: A state-of-the-art assessment". *MIS Quarterly,* Vol. 25 No. 1, pp. 1-16.

Brady, M.K. and Cronin Jr, J.J. (2001), "Some new thoughts on conceptualizing perceived service quality: a hierarchical approach". *Journal of marketing,* Vol. 65 No. 3, pp. 34-49.

Burgers, A. *et al.* (2000), "Customer expectation dimensions of voice-to-voice service encounters: a scale-development study". *International Journal of Service Industry Management,* Vol. 11 No. 2, pp. 142-161.

Caro, L.M. and Garcia, J.A.M. (2008), "Developing a multidimensional and hierarchical service quality model for the travel agency industry". *Tourism Management,* Vol. 29 No. 4, pp. 706-720.

Carrillat, F.A., Jaramillo, F. and Mulki, J.P. (2007), "The validity of the SERVQUAL and SERVPERF scales: A meta-analytic view of 17 years of research across five continents". *International Journal of Service Industry Management,* Vol. 18 No. 5, pp. 472-490.

Caruana, A. (2002), "Service loyalty: The effects of service quality and the mediating role of customer satisfaction". *European Journal of Marketing,* Vol. 36 No. 7/8, pp. 811-828.

Chi, C. (2021) *12 of the Best AI Chatbots for 2021. Available: https://blog.hubspot.com/marketing/best-ai-chatbot. Accessed April 12th, 2021.*

Chui, M. *et al.* (2020) *The state of AI in 2020. Available: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020#. Accessed April 21st, 2021.* San Francisco: McKinsey & Company.

Churchill Jr., G.A. (1979), "A paradigm for developing better measures of marketing constructs". *Journal of Marketing Research,* Vol. 16 No. 1, pp. 64-73.

Clark, L.A. and Watson, D. (1995), "Constructing validity: Basic issues in objective scale development". *Psychological Assessment,* Vol. 7 No. 3, pp. 309-319.

Colback, L. (2020) *The impact of AI on business and society. Available: https://www.ft.com/content/e082b01d-fbd6-4ea5-a0d2-05bc5ad7176c. Accessed: April 14th, 2021.* Financial Times.

Collier, J.E. and Bienstock, C.C. (2009), "Model misspecification: contrasting formative and reflective indicators for a model of e-service quality". *Journal of Marketing Theory and Practice,* Vol. 17 No. 3, pp. 283-293.

Considine, E. and Cormican, K. (2016), "Self-service technology adoption: an analysis of customer to technology interactions". *Procedia Computer Science,* Vol. 100 pp. 103-109.

Cronin Jr, J.J., Brady, M.K. and Hult, G.T.M. (2000), "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments". *Journal of Retailing,* Vol. 76 No. 2, pp. 193-218.
Cronin Jr, J.J. and Taylor, S.A. (1992), "Measuring service quality: a reexamination and extension". *Journal of marketing,* Vol. 56 No. 3, pp. 55-68.
Cronin Jr, J.J. and Taylor, S.A. (1994), "SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality". *Journal of marketing,* Vol. 58 No. 1, pp. 125-131.
Dabholkar, P. (1996), "Consumer evaluations of new technology-based self-service options: An investigation of alternative modes of service quality". *International Journal of Research in Marketing,* Vol. 13 No. 1, pp. 29-51.
Dabholkar, P.A., Shepherd, C.D. and Thorpe, D.I. (2000), "A comprehensive framework for service quality: an investigation of critical conceptual and measurement issues through a longitudinal study". *Journal of Retailing,* Vol. 76 No. 2, pp. 139-173.
Davenport, T. *et al.* (2020), "How artificial intelligence will change the future of marketing". *Journal of the Academy of Marketing Science,* Vol. 48 No. 1, pp. 24-42.
Davenport, T. and Ronanki, R. (2018), "Artificial intelligence for the real world". *Harvard Business Review,* Vol. 96 No. 1, pp. 108-116.
Davis, F.D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance in information technology". *MIS Quarterly,* Vol. 13 No. 3, pp. 319-339.
Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1992), "Extrinsic and intrinsic motivation to use computers in the workplace". *Journal of Applied Social Psychology,* Vol. 22 No. 14, pp. 1111-1132.
De Keyser, A. *et al.* (2019), "Frontline service technology infusion: conceptual archetypes and future research directions". *Journal of Service Management,* Vol. No.
Ding, D.X., Hu, P.J.-H. and Sheng, O.R.L. (2011), "e-SELFQUAL: A scale for measuring online self-service quality". *Journal of Business Research,* Vol. 64 No. 5, pp. 508-515.
Duffy, B.R. (2003), "Anthropomorphism and the social robot". *Robotics and autonomous systems,* Vol. 42 No. 3-4, pp. 177-190.
Dwivedi, Y.K. *et al.* (2021), "Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy". *International Journal of Information Management,* Vol. 57 No. Article ID 101994.
Elliot, C. (2018) *Chatbots Are Killing Customer Service. Here's Why. Available:* https://www.forbes.com/sites/christopherelliott/2018/08/27/chatbots-are-killing-customer-service-heres-why/?sh=27678f2a13c5. *Accessed April 16th, 2021.*   Forbes.
Epley, N., Waytz, A. and Cacioppo, J.T. (2007), "On seeing human: a three-factor theory of anthropomorphism". *Psychological review,* Vol. 114 No. 4, pp. 864.
Fassnacht, M. and Koese, I. (2006), "Quality of electronic services: conceptualizing and testing a hierarchical model". *Journal of Service Research,* Vol. 9 No. 1, pp. 19-37.
Field, A. (2013) *Discovering statistics using IBM SPSS statistics.* London: Sage.
Fitzsimmons, J.A. (2003), "Is self-service the future of services?". *Managing Service Quality: An International Journal,* Vol. 13 No. 6, pp. 443-444.
Ford, J.K., MacCallum, R.C. and Tait, M. (1986), "The application of exploratory factor analysis in applied psychology: A critical review and analysis". *Personnel Psychology,* Vol. 39 No. 2, pp. 291-314.
Fornell, C. and Larcker, D.F. (1981), "*Structural equation models with unobservable variables and measurement error: Algebra and statistics".* Sage Publications Sage CA: Los Angeles, CA.
Fountaine, T., McCarthy, B. and Saleh, T. (2021) *Reimagining your business for AI. Available:* https://www.mckinsey.com/business-functions/mckinsey-analytics/our-

insights/reimagining-your-business-for-ai#. *Accessed April 17, 2021.* McKinsey & Company.
Frese, M. and Fay, D. (2001), "4. Personal initiative: An active performance concept for work in the 21st century". *Research in organizational behavior,* Vol. 23 No. 133-187.
Ganguli, S. and Roy, S. (2011), "Generic technology-based service quality dimensions in banking: impact on customer satisfaction and loyalty.". *International Journal of Bank Marketing,* Vol. 29 No. 2, pp. 168-189.
Grewal, D. and Levy, M. (2009), "Emerging issues in retailing research". *Journal of retailing,* Vol. 85 No. 4, pp. 522-526.
Grönroos, C. (1984), "A service quality model and its marketing implications". *European Journal of Marketing,* Vol. 18 No. 4, pp. 36-44.
Gudergan, S.P. *et al.* (2008), "Confirmatory tetrad analysis in PLS path modeling". *Journal of Business Research,* Vol. 61 pp. 1238-1249.
Hair, J.F. *et al.* (2018) *Multivariate Data Analysis.* Cengage.
Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011), "PLS-SEM: Indeed a silver bullet". *Journal of Marketing Theory and Practice,* Vol. 19 No. 2, pp. 139-152.
Hair, J.F. *et al.* (2019), "When to use and how to report the results of PLS-SEM". *European Business Review,* Vol. 31 pp. 2-24.
Han, S. and Yang, H. (2018), "Understanding adoption of intelligent personal assistants". *Industrial Management & Data Systems,* Vol. No.
He, Z. *et al.* (2017), "Measuring service quality in telematics service: development and validation of multidimensional TeleServQ scale". *Total Quality Management & Business Excellence,* Vol. 28 No. 9-10, pp. 1166-1182.
Henseler, J. (2012) 'PLS-MGA: A non-parametric approach to partial least squares-based multi-group analysis'. *Challenges at the interface of data analysis, computer science, and optimization.* Springer, pp. 495-501.
Henseler, J., Ringle, C.M. and Sarstedt, M. (2015), "A new criterion for assessing discriminant validity in variance-based structural equation modeling". *Journal of the Academy of Marketing Science,* Vol. 43 No. 1, pp. 115-135.
Henseler, J., Ringle, C.M. and Sinkovics, R.R. (2009), "The use of partial least squares path modeling in international marketing". *Advances in International Marketing,* Vol. 20 No. 277-319.
Hoy, M.B. (2018), "Alexa, Siri, Cortana, and more: an introduction to voice assistants". *Medical reference services quarterly,* Vol. 37 No. 1, pp. 81-88.
Huang, E.Y., Lin, S.-W. and Fan, Y.-C. (2015), "MS-QUAL: Mobile service quality measurement". *Electronic Commerce Research and Applications,* Vol. 14 No. 2, pp. 126-142.
Huang, M.-H. and Rust, R.T. (2017), "Technology-driven service strategy". *Journal of the Academy of Marketing Science,* Vol. 45 No. 6, pp. 906-924.
Huang, M.-H. and Rust, R.T. (2018), "Artificial intelligence in service". *Journal of Service Research,* Vol. 21 No. 2, pp. 155-172.
Huang, M.-H. and Rust, R.T. (2021), "Engaged to a Robot? The Role of AI in Service". *Journal of Service Research,* Vol. 24 No. 1, pp. 30-41.
Hussain, K. *et al.* (2019), "The dynamic outcomes of service quality: a longitudinal investigation". *Journal of Service Theory and Practice,* Vol. 29 No. 4, pp. 513-536.
Iqbal, M.S., Hassan, M.U. and Habibah, U. (2018), "Impact of self-service technology (SST) service quality on customer loyalty and behavioral intention: The mediating role of customer satisfaction". *Cogent Business & Management,* Vol. 5 No. 1, pp. 1423770.

Jarvis, C.B., MacKenzie, S.B. and Podsakoff, P.M. (2003), "A critical review of construct indicators and measurement model misspecification in marketing and consumer research". *Journal of Consumer Research,* Vol. 30 No. 2, pp. 199-218.

Kepuska, V. and Bohouta, G. (2018) Published. 'Next-generation of virtual personal assistants (microsoft cortana, apple siri, amazon alexa and google home)'. *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC),* 2018. IEEE, pp.99-103.

Kidd, C. (2019) *Chatbot vs Virtual Agent: What's The Difference? Available: https://www.bmc.com/blogs/chatbot-vs-virtual-agent/ Accessed April 12, 2021.* BMC.

Kim, S.Y., Schmitt, B.H. and Thalmann, N.M. (2019), "Eliza in the uncanny valley: Anthropomorphizing consumer robots increases their perceived warmth but decreases liking". *Marketing Letters,* Vol. 30 No. 1, pp. 1-12.

Krehl, E.-H. (2020) 'How Technology is changing Employees' Roles in the Service Encounter–A Skill-based Analysis'. *Automatisierung und Personalisierung von Dienstleistungen.* Springer, pp. 227-248.

Kuo, Y.-F., Wu, C.-M. and Deng, W.-J. (2009), "The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services". *Computers in Human Behavior,* Vol. 25 No. 4, pp. 887-896.

Ladhari, R. (2009), "A review of twenty years of SERVQUAL research". *International journal of quality and service sciences,* Vol. 1 pp. 172-198.

Ladhari, R. (2010), "Developing e-service quality scales: a literature review". *Journal of Retailing and Consumer services,* Vol. 17 pp. 464-477.

Le, N.T., Rao Hill, S. and Troshani, I. (2020), "Perceived control and perceived risk in self-service technology recovery". *Journal of Computer Information Systems,* https://doi.org/10.1080/08874417.08872020.01756533.

Lin, J.-S.C. and Hsieh, P.-L. (2011), "Assessing the self-service technology encounters: development and validation of SSTQUAL scale". *Journal of retailing,* Vol. 87 No. 2, pp. 194-206.

Lu, V.N. *et al.* (2020), "Service robots, customers and service employees: what can we learn from the academic literature and where are the gaps?". *Journal of Service Theory and Practice,* Vol. 30 No. 3, pp. 361-391.

MacKenzie, S.B., Podsakoff, P.M. and Podsakoff, N.P. (2011), "Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques". *MIS Quarterly,* Vol. No. pp. 293-334.

Makadia, M. (2020) *Can Artificial Intelligence Improve Your Customer Service? Available: https://www.business2community.com/customer-experience/can-artificial-intelligence-improve-your-customer-service-02336788. Accessed April 19th, 2021.* Business 2 Community.

Martínez, J.A. and Martínez, L. (2010), "Some insights on conceptualizing and measuring service quality". *Journal of Retailing and Consumer Services,* Vol. 17 No. 1, pp. 29-42.

Meuter, M.L. *et al.* (2000), "Self-service technologies: understanding customer satisfaction with technology-based service encounters". *Journal of marketing,* Vol. 64 No. 3, pp. 50-64.

Meyer-Waarden, L. *et al.* (2020), "How service quality influences customer acceptance and usage of robots". *Journal of Service Management Research,* Vol. 4 No. 1, pp. 35-51.

Mittal, B. and Lassar, W.M. (1996), "The role of personalization in service encounters". *Journal of Retailing,* Vol. 72 No. 1, pp. 95-109.

Mogaji, E. and Erkan, I. (2019), "Insight into consumer experience on UK train transportation services". *Travel Behaviour and Society,* Vol. 14 No. 21-33.

Mogaji, E., Soetan, T.O. and Kieu, T.A. (2020), "The implications of artificial intelligence on the digital marketing of financial services to vulnerable customers". *Australasian Marketing Journal (AMJ),* Vol. No.

Morita, T. *et al.* (2020), "Evaluation of a multi-robot cafe based on service quality dimensions". *The Review of Socionetwork Strategies,* Vol. 14 No. 1, pp. 55-76.

Moussawi, S. (2016) *Investigating personal intelligent agents in everyday life through a behavioral lens.* PhD Thesis City University of New York, Ann Arbor.

Nath, P. (2020), "Taking measure: the link between metrics and marketing's exploitative and explorative capabilities". *European Journal of Marketing,* Vol. 54 No. 1549-1580, pp.

Naveen, J. (2018) *Yes, Chatbots And Virtual Assistants Are Different! Available: https://www.forbes.com/sites/cognitiveworld/2018/12/23/yes-chatbots-and-virtual-assistants-are-different/?sh=35542e2a6d7d. Accessed April 12th, 2021.* Forbes.

Netemeyer, R.G., Bearden, W.O. and Sharma, S. (2003) *Scaling procedures: Issues and applications.* London: Sage Publications.

Neuhofer, B., Magnus, B. and Celuch, K. (2020), "The impact of artificial intelligence on event experiences: a scenario technique approach". *Electronic Markets,* Vol. No. 1-17.

Noor, N., Rao Hill, S. and Troshani, I. (2021a), "Artificial Intelligence Service Agents: Role of Parasocial Relationship". *Journal of Computer Information Systems,* Vol. https://doi.org/10.1080/08874417.2021.1962213 No. 1-15.

Noor, N., Rao Hill, S. and Troshani, I. (2021b), "Recasting service quality for AI-based service". *Australasian Marketing Journal,* Vol. No. https://doi.org/10.1177/18393349211005056.

Nunnally, J.C. (1978) *Psychometric Theory.* New York: McGraw-Hill.

Oh, H. (1999), "Service quality, customer satisfaction, and customer value: A holistic perspective". *International Journal of Hospitality Management,* Vol. 18 No. 1, pp. 67-82.

Olson, C. and Kemery, K. (2019), "Voice report: From answers to action: customer adoption of voice technology and digital assistants". *Microsoft Search and Market Intelligence, Tech. Rep,* Vol. No.

Oosthuizen, K. *et al.* (2020), "Artificial intelligence in retail: The AI-enabled value chain". *Australasian Marketing Journal,* Vol. No. https://doi.org/10.1016/j.ausmj.2020.1007.1007.

Orel, F.D. and Kara, A. (2014), "Supermarket self-checkout service quality, customer satisfaction, and loyalty: Empirical evidence from an emerging market". *Journal of Retailing and Consumer services,* Vol. 21 No. 2, pp. 118-129.

Overgoor, G. *et al.* (2019), "Letting the computers take over: Using AI to solve marketing problems". *California Management Review,* Vol. 61 No. 4, pp. 156-185.

Paluch, S. and Wirtz, J. (2020), "Artificial intelligence and robots in the service encounter". *Journal of Service Management Research,* Vol. 4 No. 1, pp. 3-8.

Parasuraman, A. (2000), "Technology readiness index (TRI): a multiple-item scale to measure readiness to embrace new technologies". *Journal of Service Research,* Vol. 2 No. 4, pp. 307-320.

Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1991), "Understanding customer expectation of service". *Sloan Management Review,* Vol. 32 No. 3, pp. 39-48.

Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1993), "More on improving service quality measurement". *Journal of Retailing,* Vol. 69 No. 1, pp. 140-147.

Parasuraman, A., Zeithaml, V.A. and Berry, L. (1988), "SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality". *1988,* Vol. 64 No. 1, pp. 12-40.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985), "A conceptual model of service quality and its implications for future research". *Journal of Marketing,* Vol. 49 No. 1, pp. 41-50.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994a), "Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria". *Journal of retailing,* Vol. 70 No. 3, pp. 201-230.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994b), "Reassessment of expectations as a comparison standard in measuring service quality: implications for further research". *Journal of marketing,* Vol. 58 No. 1, pp. 111-124.

Parasuraman, A., Zeithaml, V.A. and Malhotra, A. (2005), "ES-QUAL: A multiple-item scale for assessing electronic service quality". *Journal of service research,* Vol. 7 No. 3, pp. 213-233.

Park, S.S., Tung, C.D. and Lee, H. (2021), "The adoption of AI service robots: a comparison between credence and experience service settings". *Psychology & Marketing,* Vol. 38 No. 691-703, pp.

Paschen, J., Wilson, M. and Ferreira, J.J. (2020), "Collaborative intelligence: how human and artificial intelligence create value along the B2B sales funnel". *Business Horizons,* Vol. 63 No. 3, pp. 403-414.

PATResearch (2021) *Top 22 intelligent personal assistants or automated personal assistants. Available: https://www.predictiveanalyticstoday.com/top-intelligent-personal-assistants-automated-personal-assistants/. Accessed April 12, 2021.* Predictive Analytics Today.

Poole, D.L. and Mackworth, A.K. (2010) *Artificial Intelligence: foundations of computational agents.* Camridge: Cambridge University Press.

Prentice, C., Lopes, S.D. and Wang, X. (2020), "The impact of artificial intelligence and employee service quality on customer satisfaction and loyalty". *Journal of Hospitality Marketing & Management,* Vol. 29 No. 7, pp. 739-756.

PwC (2018) *Artificial intelligence: Touchpoints with consumersAvailable https://www.pwc.com/co/es/industrias/Retail/insights/gcis-ai_vf.pdf. Accessed April 24th 2021.* PwC.

Qiu, L. and Benbasat, I. (2009), "Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems". *Journal of management information systems,* Vol. 25 No. 4, pp. 145-182.

Ranjan, K.R. and Read, S. (2016), "Value co-creation: concept and measurement". *Journal of the Academy of Marketing Science,* Vol. 44 No. 3, pp. 290-315.

Rank, J. *et al.* (2007), "Proactive customer service performance: Relationships with individual, task, and leadership variables". *Human Performance,* Vol. 20 No. 4, pp. 363-390.

Rita, P., Oliveira, T. and Farisa, A. (2019), "The impact of e-service quality and customer satisfaction on customer behavior in online shopping". *Heliyon,* Vol. 5 No. 10, pp. e02690.

Robinson, S. *et al.* (2020), "Frontline encounters of the AI kind: An evolved service encounter framework". *Journal of Business Research,* Vol. 116 No. 366-376.

Russell, S. and Norvig, P. (2018) *Artificial intelligence: a modern approach.* Boston, MA: Pearson.

Rust, R.T. (2020), "The future of marketing". *International Journal of Research in Marketing,* Vol. 37 No. 1, pp. 15-26.

Rust, R.T. and Huang, M.-H. (2014), "The service revolution and the transformation of marketing science". *Marketing Science,* Vol. 33 No. 2, pp. 206-221.

Rust, R.T. and Oliver, R.L. (1993) *Service quality: New directions in theory and practice.* London: Sage Publications.

Salomann, H. *et al.* (2007), "Self-service Revisited: How to Balance High-tech and High-touch in Customer Relationships". *European Management Journal,* Vol. 25 No. 4, pp. 310-319.
Schniter, E., Shields, T.W. and Sznycer, D. (2020), "Trust in humans and robots: Economically similar but emotionally different". *Journal of Economic Psychology,* Vol. 78 No. https://doi.org/10.1016/j.joep.2020.102253.
Seth, N., Deshmukh, S. and Vrat, P. (2005), "Service quality models: a review". *International journal of quality & reliability management,* Vol. 22 No. 9, pp. 913-949.
Shankar, V. (2018), "How artificial intelligence (AI) is reshaping retailing". *Journal of Retailing,* Vol. 94 No. 4, pp. vi-xi.
Sheehan, B., Jin, H.S. and Gottlieb, U. (2020), "Customer service chatbots: Anthropomorphism and adoption". *Journal of Business Research,* Vol. 115 No. 14-24.
Shell, M.A. and Buell, R.W. (2019), "How Necessary Is the Human Touch?". *Harvard Business Review,* Vol. 97 No. 4, pp. 28-28.
Shepardson, D. (2020) *White House proposes regulatory principles to govern AI use. Available: https://www.reuters.com/article/us-tech-ces-ai-white-house-idUSKBN1Z60GL. Accessed April 24, 2021.* Reuters.
Shewan, D. (2021) *10 of the Most Innovative Chatbots on the Web. Available: https://www.wordstream.com/blog/ws/2017/10/04/chatbots Accessed April 12th, 2021.*
Stock, R.M. and Merkle, M. (2018) Published. 'Can humanoid service robots perform better than service employees? A comparison of innovative behavior cues. Available: https://scholarspace.manoa.hawaii.edu/bitstream/10125/50020/1/paper0133.pdf'. *Proceedings of the 51st Hawaii International Conference on System Sciences,* 2018 Waikoloa Village, HI, January 3-6. pp.1056-1065.
Tabachnick, B.G., Fidell, L.S. and Ullman, J.B. (2007) *Using multivariate statistics.* Pearson Boston, MA.
Tam, J.L. (2004), "Customer satisfaction, service quality and perceived value: an integrative model". *Journal of Marketing Management*, Vol. 20 No. 7-8, pp. 897-917.
Tan, F.B. and Chou, J.P. (2008), "The relationship between mobile service quality, perceived technology compatibility, and users' perceived playfulness in the context of mobile information and entertainment services". *Intl. Journal of Human–Computer Interaction,* Vol. 24 No. 7, pp. 649-671.
Theodosiou, M. *et al.* (2019), "A comparison of formative versus reflective approaches for the measurement of electronic service quality". *Journal of Interactive Marketing,* Vol. 47 No. 53-67.
Troshani, I. *et al.* (2020), "Do We Trust in AI? Role of Anthropomorphism and Intelligence". *Journal of Computer Information Systems,* Vol. No.
Verhagen, T. *et al.* (2014), "Virtual customer service agents: Using social presence and personalization to shape online service encounters". *Journal of Computer-Mediated Communication,* Vol. 19 No. 3, pp. 529-545.
Voorhees, C.M. *et al.* (2016), "Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies". *Journal of the academy of marketing science,* Vol. 44 No. 1, pp. 119-134.
Whang, C. and Im, H. (2021), "" I Like Your Suggestion!" the role of humanlikeness and parasocial relationship on the website versus voice shopper's perception of recommendations". *Psychology & Marketing,* Vol. 38 No. 4, pp. 581-595.
Wirtz, J. *et al.* (2018), "Brave new world: service robots in the frontline". *Journal of Service Management*, Vol. 29 No. 5, pp. 776-808.

Yang, Z., Jun, M. and Peterson, R.T. (2004), "Measuring customer perceived online service quality". *International Journal of Operations & Production Management*, Vol. 24 No. 11, pp. 1149-1174.

Youn, S. and Jin, S.V. (2021), "In AI we trust?" The effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging "feeling economy". *Computers in Human Behavior,* Vol. 119 No. 106721.

Zarouali, B. *et al.* (2018), "Predicting consumer responses to a chatbot on Facebook". *Cyberpsychology, Behavior, and Social Networking,* Vol. 21 No. 8, pp. 491-497.

Zeithaml, V.A. (1988), "Consumer perceptions of price, quality and value: a means-end model and synthesis of evidence". *Journal of Marketing,* Vol. 52 No. 1, pp. 2-22.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1988), "Communication and Control Processes in the Delivery of Service Quality". *Journal of Marketing,* Vol. 52 No. 2, pp. 35-48.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996a), "The behavioral consequences of service quality". *The Journal of Marketing,* Vol. No. 31-46.

Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996b), "The behavioural consequences of service quality". *Journal of Marketing,* Vol. 60 No. 1, pp. 31-46.

Zeithaml, V.A., Parasuraman, A. and Malhotra, A. (2002), "Service quality delivery through web sites: a critical review of extant knowledge". *Journal of the Academy of Marketing Science,* Vol. 30 No. 4, pp. 362-375.

Zendesk (2020) *The business impact of customer service on customer lifetime value. Available: https://www.zendesk.com/blog/customer-service-and-lifetime-customer-value/. Published April 8th, 2013. Updated October 6th, 2020. Accessed April 19th, 2021.* Zendesk.

Zheng, X. *et al.* (2020), "Role of technology attraction and parasocial interaction in social shopping websites". *International Journal of Information Management,* Vol. 51 No. 102043.

APPENDIX: Measures of Constructs

Respondents were asked to rate the following statements using a seven-point Likert scale anchored from 1 = strongly disagree to 7 = strongly agree. Items appeared in random order in the survey.

<u>AISAQUAL</u>

Efficiency

EFF1 The AISA works correctly at first attempt.
EFF2 I can get my task done with the AISA in a short time.
EFF3 The AISA interface design provides information clearly.
EFF4 The AISA adequately meets my requirements.

Security

SEC1 There is no risk of loss associated with disclosing personal information to the AISA.
SEC2 I feel secure in providing sensitive information to the AISA.
SEC3 I believe that information that the AISA has about me is protected.
SEC4 I trust that my personal information with the AISA will not be misused.

Availability

AVA1 The AISA is always available.
AVA2 The AISA is never too busy to respond to my requests.
AVA3 The AISA is always accessible.

Enjoyment

ENJ1 Using the AISA is fun.
ENJ2 Using the AISA is enjoyable.
ENJ3 Using the AISA is interesting.
ENJ4 Using the AISA is entertaining.

Contact

CON1 Human assistants are available to contact via the AISA.
CON2 Follow-up services with human assistants are available to me when necessary.
CON3 I can speak to a human assistant via the AISA.
CON4 Human assistance is easy to access via the AISA.
CON5 The AISA provides detailed contact information when I need human assistance.

Anthropomorphism

ANT1 The AISA has humanlike features.
ANT2 The AISA has personality.
ANT3 The AISA gradually gets to know me.
ANT4 The AISA is able to behave like a human.
ANT5 The AISA responds in ways that are personalized.
ANT6 The AISA is able to communicate like a human.

<u>Satisfaction</u>

SAT1 I am satisfied with my decision to use the AISA.
SAT2 I think that I did the right thing by using the AISA.
SAT3 My choice to use the AISA was a wise one.

<u>Perceived Value</u>

VAL1 Overall, the AISA gives me good value.
VAL2 The time I spent on the AISA was worthwhile.

<u>Loyalty Intentions</u>

LOY1 I will say positive things about the AISA to other people.
LOY2 I will recommend the AISA to someone who seeks my advice.
LOY3 I will encourage friends and others to use the AISA.
LOY4 I will consider the AISA to be my first choice for future tasks.
LOY5 I will use the AISA more in the coming months.

Table I.
Chatbots and Virtual assistants

	Chatbot	Virtual assistant
Common functionality and uses	Chatbots are rule-based AI applications that are typically used in business to facilitate customer interaction. Chatbot are also known as 'conversational agents' as typical interaction with customers is in the form for a structured dialogue or chat via spoken language text-based conversational interface, although there are also chatbots that support audio and images (Shewan, 2021). Typically, consumers ask specific questions, while the chatbot provides live pre-defined responses. Companies have been using chatbots to engage with consumers in a range of roles such as consumer service and support such as order customization, scheduling (e.g., product deliveries, or travel bookings) or refunds.	Virtual assistants are software agents that act as personal assistant to users for a wide range of tasks in daily activities, both work-related and personal, such as personal communications management such as email and SMS, productivity management such as scheduling of meetings, or performing general tasks such as playing music, setting alarms. Virtual assistants are designed to understand natural language voice commands and questions including implied meaning, user emotions, language slangs, and dialects. Virtual assistants can learn from prior interactions with their users, consequently improve ability to contextualize and customize interactions overtime.
Usage channel	Websites, support portals, messaging channels and mobile applications.	Mobile phones, laptop computers, smart speakers and interactive devices.
Common examples	IBM's Watson Assistant, SalesForce's Eistein (Chi, 2021).	Apple's Siri, Amazon's Alexa, Microsoft's Cortana (PATResearch, 2021)

Table II.
Dimensions representing service quality of AISA

Dimension	Definition	Key supporting service quality scale studies
Reliability	Ability of the AISA to perform the service dependably and accurately	Parasuraman *et al.* (1988)
Responsiveness	Prompt response of the AISA to customer requests and the speed in resolving customer problems	Yang *et al.* (2004)
Availability	Ability of the AISA to be ready for use anytime, anywhere	(Lin and Hsieh, 2011), Parasuraman *et al.* (2005)
Aesthetics	Appeal and clarity associated with the AISA interface design	Dabholkar (1996)
Personalization	Ability of the AISA to meet the customer's individual preferences	He *et al.* (2017)
Security	Perceived safety of the AISA from intrusion, fraud and loss of personal information and privacy	He *et al.* (2017)
Control	Degree of control that the customer feels over the process or outcome of the service encounter with the AISA	Dabholkar (1996)
Ease of Use	Degree to which using the AISA would be free of effort	Davis (1989)
Enjoyment	Extent to which using the AISA is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated	Davis *et al.* (1992)
Contact	Access to human assistance	Parasuraman *et al.* (2005)
Proactiveness	Extent to which AISA is able to predict and anticipate customers' future needs and acts beyond explicitly prescribed commands.	Noor *et al.* (2021b)
Anthropomorphism	Extent to which AISA shows human-like characteristics, motivations, intentions, or emotions.	Noor *et al.* (2021b)

Table III.
Profile of respondents for scale refinement phase

Category	EFA (N=211)	CFA (N=275)
Gender		
Male	107 (50.7%)	132 (48%)
Female	104 (49.3%)	143 (52%)
Total	211	275
Age		
18-24	21 (10.0%)	26 (9.5%)
25-34	70 (33.2%)	91 (33.0%)
35-44	72 (34.1%)	92 (33.4%)
45-54	30 (14.2%)	49 (17.8%)
55-64	11 (5.2%)	13 (4.7%)
65 and above	7 (3.3%)	4 (1.5%)
Total	211	275
Chatbot usage context		
Accommodation and food services	8 (7.6%)	11 (7.9%)
Administrative and support services	7 (6.7%)	9 (6.4%)
Arts and recreation services	5 (4.8%)	7 (5%)
Education and training	4 (3.8%)	5 (3.6%)
Electricity, gas, water and waste services	21 (20%)	26 (18.6%)
Financial and insurance services	15 (14.3%)	18 (12.9%)
Health care and social assistance	5 (4.8%)	8 (5.7%)
Information media and telecommunications	16 (15.2%)	17 (12.1%)
Professional, scientific and technical services	7 (6.7%)	9 (6.4%)
Public administration and safety	3 (2.9%)	6 (4.3%)
Rental, hiring and real estate services	3 (2.9%)	4 (2.9%)
Retail trade	4 (3.8%)	9 (6.4%)
Transport, postal and warehousing	7 (6.7%)	12 (8.6%)
Total	105	140
Virtual assistant		
Alexa	31 (29.2%)	40 (29.6%)
Bixby	2 (1.9%)	3 (2.2%)
Google Assistant	42 (39.6%)	43 (31.9%)
Google Home Mini	4 (3.8%)	8 (5.9%)
Siri	27 (25.5%)	41 (30.4%)
Total	106	135

Table IV.
EFA and CFA results and final AISAQUAL scale

Dimension after EFA	Item[a]	Original Identifier[c]	EFA Loading	CFA[b] Loading	Final Label
D1 (Cronbach's α = .84)	The AISA provides the service as expected.	REL7	.74		Efficiency
	The AISA works correctly at first attempt.[a]	**REL5**	.73	**.72**	
	I can get my task done with the AISA in a short time.[a]	**RES1**	.66	**.74**	
	The AISA can perform the task quickly.	RES2	.65		
	The AISA interface design provides information clearly.[a]	**AES5**	.63	**.73**	
	I know how long it takes to complete the task with the AISA.	CTL3	.60		
	The AISA adequately meets my requirements.[a]	**REL3**	.52	**.82**	
D2 (Cronbach's α = .88)	I trust that my personal information with the AISA is safe.	SEC5	.83		Security
	There is no risk of loss associated with disclosing personal information to the AISA.[a]	**SEC9**	.73	**.77**	
	I feel secure in providing sensitive information to the AISA.[a]	**SEC3**	.71	**.76**	
	I believe that information that the AISA has about me is protected.[a]	**SEC7**	.65	**.88**	
	I trust that my personal information with the AISA will not be misused.[a]	**SEC6**	.59	**.80**	
D3 (Cronbach's α = .80)	**The AISA is always available.[a]**	**AVA2**	.84	**.71**	Availability
	The AISA is never too busy to respond to my requests.[a]	**AVA4**	.71	**.87**	
	The AISA launches right away.	RES3	.70		
	The AISA is always accessible.[a]	**AVA1**	.58	**.79**	
	The AISA interface design is innovative.	AES1	.46		
D4 (Cronbach's α = .89)	**Using the AISA is fun.[a]**	**ENJ5**	.67	**.74**	Enjoyment
	Using the AISA is enjoyable.[a]	**ENJ3**	.52	**.87**	
	Using the AISA is interesting.[a]	**ENJ6**	.46	**.79**	
	Using the AISA is entertaining.[a]	**ENJ4**	.43	**.84**	
D5 (Cronbach's α = .88)	**Human assistants are available to contact via the AISA.[a]**	**CTC2**	.81	**.78**	Contact
	Follow-up services with human assistants are available to me when necessary.[a]	**CTC5**	.71	**.71**	
	I can speak to a human assistant via the AISA.[a]	**CTC3**	.65	**.78**	
	Human assistance is easy to access via the AISA.[a]	**CTC1**	.64	**.82**	
	The AISA provides detailed contact information when I need human assistance.[a]	**CTC4**	.60	**.80**	
	A clear privacy policy is accessible before I use the AISA.	SEC1	.45		
D6 (Cronbach's α = .90)	**The AISA has humanlike features.[a]**	**ANT1**	.85	**.76**	Anthropomorphism
	The AISA has personality.[a]	**ANT5**	.77	**.74**	
	The AISA gradually gets to know me.[a]	**PER9**	.72	**.75**	
	The AISA is able to behave like a human.[a]	**ANT4**	.64	**.84**	
	The AISA responds in ways that are personalized.[a]	**PER8**	.62	**.84**	
	The AISA is able to communicate like a human.[a]	**ANT3**	.61	**.78**	
	The AISA uses its own 'judgment' to complete a task.	PRO5	.47		

[a] final AISAQUAL items are shown in bold
[b] $\chi^2 = 835.01$, $p < .001$; RMSEA = .07, TLI = .91, CFI = .90, SRMR = .08.
[c] REL=Reliability, RES=Responsiveness, AVA=Availability, AES=Aesthetics, PER=Personalization, SEC=Security, CTL=Control, EAS=Ease of Use, ENJ=Enjoyment, CTC=Contact, PRO=Proactiveness, ANT=Anthropomorphism

Table V.

Profile of respondents for scale validation phase

Category	Frequency	Percentage
Gender		
Male	152	50.0
Female	152	50.0
Total	304	100.0
Age		
18-24	54	17.8
25-34	79	26.0
35-44	72	23.7
45-54	44	14.5
55-64	35	11.5
65 and above	20	6.6
Total	304	100.0
Highest education		
Less than high school	2	0.7
High school	50	16.4
Vocational training	13	4.3
Some college	86	28.3
Bachelor’s degree	95	31.3
Postgraduate degree	58	19.1
Total	304	100.0
Work industry		
Accommodation and food services	10	3.3
Administrative and support services	20	6.6
Arts and recreation services	16	5.3
Construction	8	2.6
Education and training	26	8.6
Electricity, gas, water and waste services	8	2.6
Financial and insurance services	23	7.6
Health care and social assistance	34	11.2
Information media and telecommunications	18	5.9
Manufacturing	7	2.3
Professional, scientific and technical services	24	7.9
Public administration and safety	9	3.0
Rental, hiring and real estate services	5	1.6
Retail trade	33	10.9
Transport, postal and warehousing	19	6.3
Other Industries	7	2.3
Retired	19	6.3
Unemployed	18	5.9
Total	304	100.0
Personal annual income (USD)		
Less than $25,000	62	20.4
$25,000 to $49,999	75	24.7
$50,000 to $74,999	60	19.7
$75,000 to $99,999	45	14.8
$100,000 and more	62	20.4
Total	304	100.0

Category	Frequency	Percentage
AISA usage frequency		
Daily	95	31.3
Weekly	97	31.9
Every 2-3 weeks	43	14.1
Monthly	33	10.9
Every 2-3 months	14	4.6
Every 4-6 months	10	3.3
Once a year	12	3.9
Total	304	100.0
AISA usage experience		
Less than 1 year	85	28.0
2-3 years	144	47.4
4-5 years	51	16.8
6-7 years	13	4.3
8 years and above	11	3.6
Total	304	100.0
Chatbot usage context		
Accommodation and food services	10	6.6
Administrative and support services	19	12.6
Arts and recreation services	6	4.0
Construction	2	1.3
Education and training	11	7.3
Electricity, gas, water and waste services	5	3.3
Financial and insurance services	19	12.6
Health care and social assistance	12	7.9
Information media and telecommunications	18	11.9
Professional, scientific and technical services	5	3.3
Public administration and safety	2	1.3
Rental, hiring and real estate services	4	2.6
Retail trade	30	19.9
Transport, postal and warehousing	4	2.6
Others	4	2.6
Total	151	100.0
Virtual assistant		
Alexa	60	39.2
Bixby	5	3.3
Google Assistant	36	23.5
Google Home Mini	18	11.8
Siri	34	22.2
Total	153	100.0

Table VI.

CFA results of AISAQUAL for scale validation phase

Dimension	Item	Mean	S.D.	CFA Loading	CR	AVE
Efficiency (Cronbach's α = .84)	The AISA works correctly at first attempt.	4.77	1.38	.69	.84	.57
	I can get my task done with the AISA in a short time.	5.12	1.37	.79		
	The AISA interface design provides information clearly.	5.20	1.27	.72		
	The AISA adequately meets my requirements.	5.10	1.40	.82		
Security (Cronbach's α = .89)	There is no risk of loss associated with disclosing personal information to the AISA.	3.99	1.69	.78	.89	.67
	I feel secure in providing sensitive information to the AISA.	4.16	1.71	.81		
	I believe that information that the AISA has about me is protected.	4.46	1.63	.86		
	I trust that my personal information with the AISA will not be misused.	4.43	1.65	.81		
Availability (Cronbach's α = .80)	The AISA is always available.	5.42	1.37	.75	.80	.58
	The AISA is never too busy to respond to my requests.	5.38	1.53	.74		
	The AISA is always accessible.	5.40	1.38	.80		
Enjoyment (Cronbach's α = .89)	Using the AISA is fun.	5.02	1.60	.85	.89	.68
	Using the AISA is enjoyable.	5.05	1.48	.80		
	Using the AISA is interesting.	5.17	1.44	.79		
	Using the AISA is entertaining.	4.91	1.61	.84		
Contact (Cronbach's α = .87)	Human assistants are available to contact via the AISA.	4.59	1.50	.76	.87	.58
	Follow-up services with human assistants are available to me when necessary.	4.76	1.45	.76		
	I can speak to a human assistant via the AISA.	4.53	1.56	.76		
	Human assistance is easy to access via the AISA.	4.79	1.59	.81		
	The AISA provides detailed contact information when I need human assistance.	4.77	1.49	.71		
Anthropomorphism (Cronbach's α = .90)	The AISA has humanlike features.	4.56	1.64	.74	.90	.60
	The AISA has personality.	4.46	1.60	.79		
	The AISA gradually gets to know me.	4.59	1.55	.74		
	The AISA is able to behave like a human.	4.40	1.65	.77		
	The AISA responds in ways that are personalized.	4.74	1.47	.80		
	The AISA is able to communicate like a human.	4.72	1.55	.83		

$\chi^2 = 562.94$, $p < .001$; RMSEA = .06, TLI = .94, CFI = .95 and SRMR = .04.

Table VII.
HTMT analysis of AISAQUAL

Constructs	EFF	SEC	AVA	ENJ	CON	ANT
Efficiency (EFF)						
Security (SEC)	.75					
Availability (AVA)	.78	.50				
Enjoyment (ENJ)	.86	.61	.68			
Contact (CON)	.80	.63	.53	.66		
Anthropomorphism (ANT)	.86	.79	.62	.84	.73	

Table VIII.
Squared correlations between AISAQUAL dimensions.

Constructs	EFF	SEC	AVA	ENJ	CON	ANT
Efficiency (EFF)	.75					
Security (SEC)	.65	.82				
Availability (AVA)	.64	.42	.76			
Enjoyment (ENJ)	.74	.53	.58	.82		
Contact (CON)	.68	.55	.44	.58	.76	
Anthropomorphism (ANT)	.74	.71	.53	.76	.65	.77

Notes: average variance extracted appears on the diagonal. All correlations are significant at p< .01.

Table IX.
Multigroup comparison test results of AISAQUAL

Path	Chatbot Users (151)		Virtual Assistant Users (153)		Multigroup Analysis		
	β	t	β	t	β Difference	*p* Difference	Result
AISAQUAL -> Loyalty Intentions	.17	2.62	.16	2.90	.01	.95	Rejected
AISAQUAL -> Perceived Value	.83	29.98	.81	23.85	.01	.81	Rejected
AISAQUAL -> Satisfaction	.37	4.81	.34	5.45	.03	.74	Rejected
Perceived Value -> Loyalty Intentions	.33	3.76	.12	1.18	.22	.11	Rejected
Perceived Value -> Satisfaction	.55	7.22	.62	9.73	-.07	.51	Rejected
Satisfaction -> Loyalty Intentions	.47	5.70	.69	6.15	-.23	.11	Rejected

Note: Multigroup analysis based on 5000 bootstrap. Results are based on two tail test at 5% probability of error level.

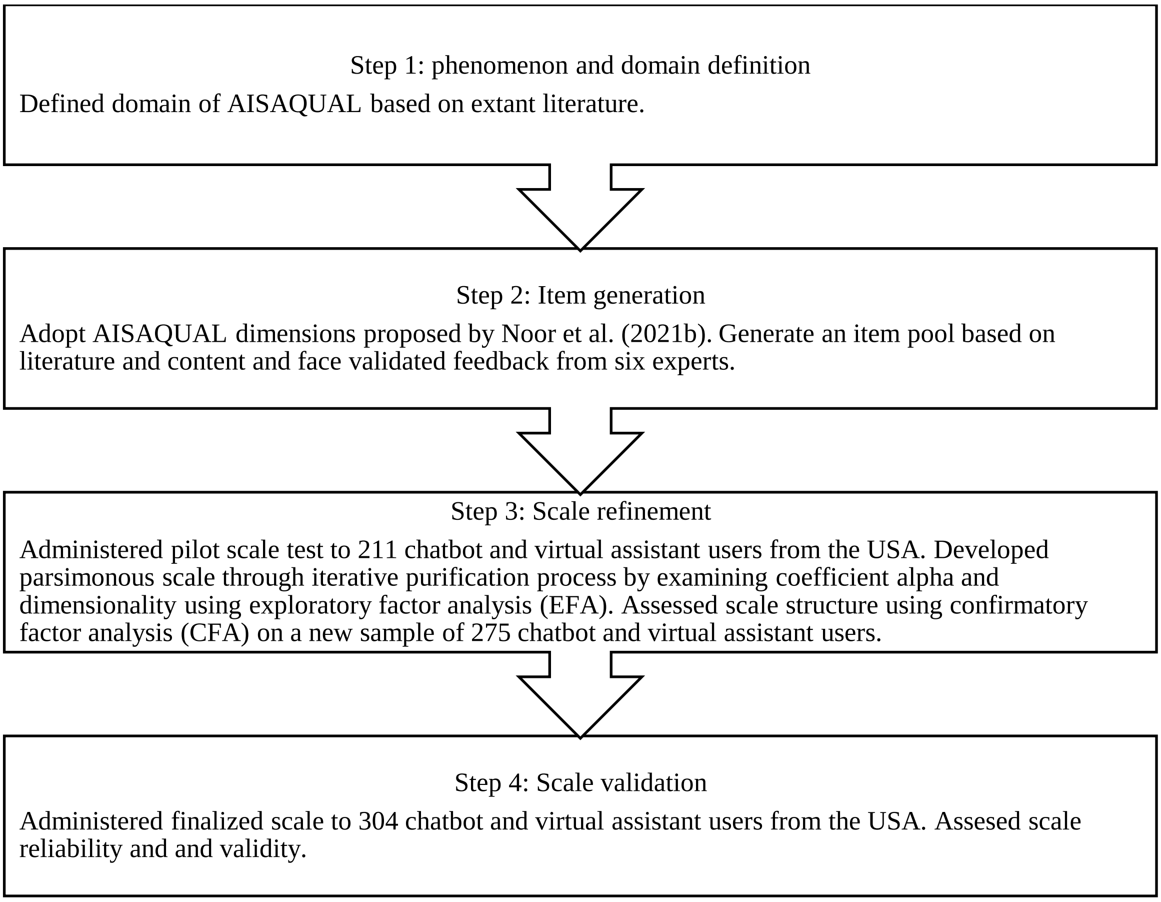

Figure 1. Research approach

MEASURE ITEMS FROM HUMAN-BASED SERVICE QUALITY SCALES	
Service Context	**Representative Study**
General service environments	Parasuraman *et al.* (1988)
Person-processing services involving interpersonal interactions	Mittal and Lassar (1996)
Retail services (goods and services)	Dabholkar *et al.* (1996)
Call center representatives	Burgers *et al.* (2000)
Service encounters involving interaction, physical environment and outcome evaluations	Brady and Cronin Jr (2001)

CONCEPTUAL DIMENSIONS OF AISAQUAL
Reliability
Responsiveness
Availability
Aesthetics
Personalization
Security
Control
Ease of Use
Enjoyment
Contact
Proactiveness
Anthropomorphism

MEASURE ITEMS FROM TECHNOLOGY-BASED SERVICE QUALITY SCALES	
Representative Study	**Service Context**
Dabholkar (1996)	General self-service technologies involving cognitive or affective assessments
Yang *et al.* (2004)	Online services involving variety of service processes (e.g. online banking)
Yang *et al.* (2005)	Web portals
Parasuraman *et al.* (2005)	E-commerce
Parasuraman *et al.* (2005)	Service recovery in e-commerce
Loiacono *et al.* (2007)	General websites used for information gathering, transacting, entertainment
Ding *et al.* (2011)	E-retailing self-services
Lin and Hsieh (2011)	Self-service technologies across various service industries
Huang *et al.* (2015)	Mobile commerce
He *et al.* (2017)	Telematics

Figure 2. Extant service quality scales used to form initial AISAQUAL item battery

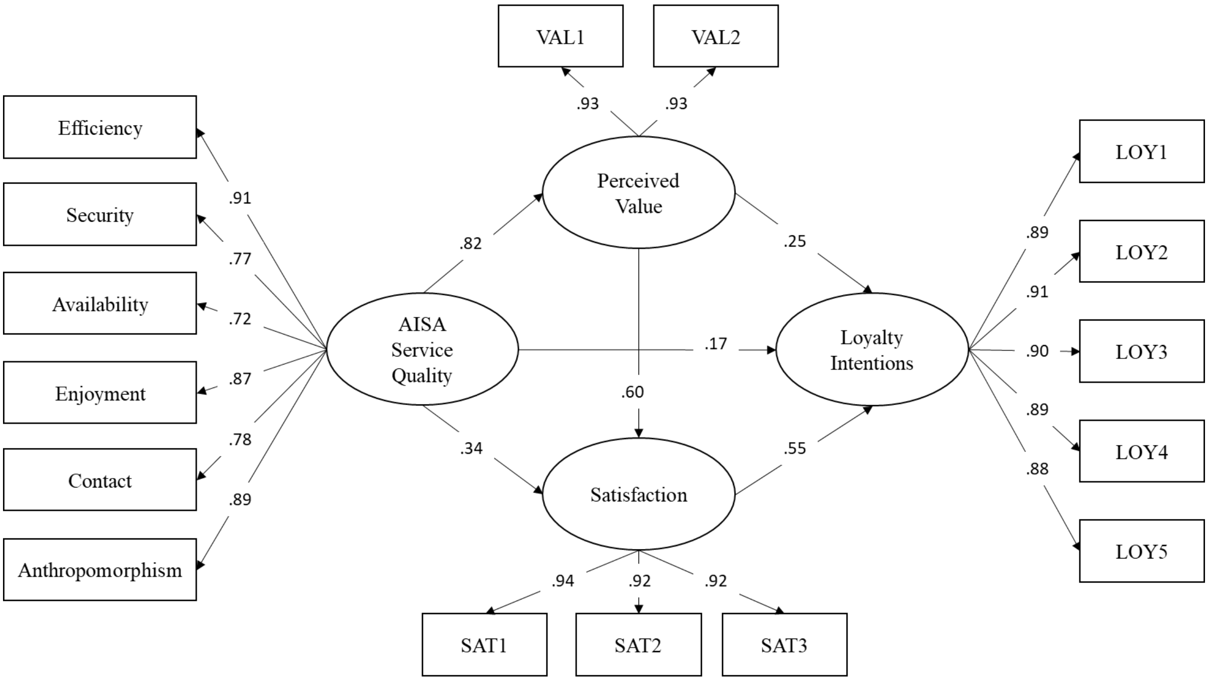

Figure 3. Model for AISAQUAL nomological validity assessment. All parameter estimates are significant at the .001 level.