Robotic service quality – Scale development and validation

Author

Prentice, C, Nguyen, M

Published

2021

Journal Title

Journal of Retailing and Consumer Services

Version

Accepted Manuscript (AM)

DOI

10.1016/j.jretconser.2021.102661

Robotic service quality – scale development and validation

Abstract

Dr Catherine Prentice *
Associate Professor of Marketing
Department of marketing
Griffith Institute for Tourism
Griffith Business School,
Griffith University
Staffordshire Business School
Staffordshire University
Australia
cathyjournalarticles@gmail.com

Mai Nguyen | PhD Candidate
Department of Marketing, Griffith Business School
Griffith University | Nathan campus | QLD 4111
E mai.journalarticles@gmail.com
mai.nguyenthituyet@griffithuni.edu.au

Robotic service quality – scale development and validation

Abstract

Viewing robots as service agents that provide services to customers for value exchange, the study developed a scale to measure robotic service quality. The scale underwent several stages of development including item generation, domain specification, scale refinement, and validity testing, including internal and external cross validation. A range of methods were used in this process. Data were collected from Australia, China, and Vietnam to test external validity. Four dimensions were identified to represent robotic service quality. Development of this scale has implications for artificial intelligence and service research. The scale can be used by practitioners to enhance customer experience and generate positive attitudinal and behavioural responses from customers.

Keywords: Robots, artificial intelligence, service quality, scale development, services marketing

INTRODUCTION

Robots with advanced AI technology and machine learning have been utilised in the service sector to facilitate business operations and customer service. Gartner (2019) suggested that 37 percent of organisations had adopted robots and AI in some form in 2019, increasing by 270 percent over the past four years. Robotic service today has been integrated into service organisations to enhance customer experience (Prentice et al., 2020). For example, Chatbots have been deployed to enable self-service procedures (Kumar & Balaramachandran, 2018). The concierge robot "Connie" (e.g. Hilton Worldwide) can interact with hotel guests and respond to individualised questions. The Henn-na Hotel in Japan uses robots for frontline and concierge services, as well as assisting check-in and check-out (Rajesh, 2015). These service robots function as service employees that deliver customer service as well as facilitate employees performance of service tasks (Cobos et al., 2016; Marinova et al., 2017; West et al., 2018).

Wirtz and colleagues (2018) discussed the role of service robots in the service industry and compared their features and capabilities with those of frontline employees on micro, meso, and macro levels. The authors indicated that service robots can replace human employees for some low-level service tasks, as well as collaborate with employees to achieve high-level service and facilitate service delivery. The quality of customer service is determined by customers' perceptions and evaluations (Parasuraman et al., 1994). Service quality assessment has been well documented although robotic service is relatively new to the literature. No research to date has incorporated robotic service within service quality assessment. This paper views service robots as service providers that deliver a service to customers to generate positive attitudes and behaviours towards the service firm for value creation. Consistent with this view, the aim of this study is to develop a scale to measure robotic services, referred to as robotic service quality. Given the prevalent application of robotic use in customer service, the development of this scale would help service firms to assess the role of robotic service in the organisation and the potential influence on customers' attitudes and behaviours towards the entity.

Service quality in the marketing literature refers to customers' perceptions and value-judgments of a product or service. Service quality has also been recognised as an imperative factor for competitive advantage and organisational profitability (Shi et al., 2014). These outcomes result from positive customer responses manifested in market share expansion, customer satisfaction, and loyalty (Ozment & Morash, 1994). The underlying logic is that improving service quality leads to customer satisfaction and customer loyalty, which relates to business profitability (Heskett et al., 1994). Consequently, measuring a firm's service quality is imperative to understand its impact on organisational outcomes.

The scale development in this study draws upon the conceptualisations of service and service quality. This initiative provides a better understanding of robotic service and its impact on organisational outcomes as service quality has been widely acknowledged as an antecedent of customer satisfaction and loyalty. The paper first discusses robotic service as a product that a service organisation provides to customers for value creation (e.g. customer satisfaction, purchase, and loyalty behaviours). Value is manifested in the level of robotic service quality assessed by customers. The existing service quality models will be briefly discussed to determine the protocols that robotic service quality measurement draw on. The scale development process will then be elaborated. The results and implications conclude this paper.

LITERATURE REVIEW

Robotic service

Robots can be described as intelligent physical devices with programmed autonomy, mobility, and sensor ability to perform certain human tasks, or intelligent machines acting like humans to perceive, learn, memorise, reason, and solve problems through machine learning, deep learning, or natural language processing (Russell et al., 2016). Robots can interpret and learn from external data for predefined goals and tasks (Kaplan & Haenlein, 2019). Robots in the service sector, referred to as service robots, can have a physical (e.g. Pepper) or virtual (e.g. Alexa) representation, with a humanoid (e.g. Sophia) or non-humanoid appearance (e.g. Roomba cleaning robot) (Wirtz et al., 2018). The physical humanoid and non-humanoid robots can function as frontline service employees to deliver customer service, and virtual robots can operate as e-service providers (Wirtz et al., 2018). Consistent with this analogy, the former are referred to as frontline service robots, and the latter as e-service robots.

Service in the marketing literature is defined as an act, deed, or performance that is offered as a value-added activity to benefit stakeholders (Wirtz and Lovelock, 2018). Services are generally categorised into four domains: people processing, possession processing, mental stimulus processing, and information processing (Wirtz and Lovelock, 2018). People processing services refer to those that directly serve customers, such as medical centres providing medical services and the examination of patients, requiring simultaneous participation from both the service provider and the customer. Robots can provide such services. For instance, robotic medical assistants have been used to monitor patients and alert the medical staff when needed. Possession processing services refer to services directed at customers' tangible possessions, for example, house cleaning. A robotic vacuum cleaner can perform this kind of service. Mental stimulus processing services refer to offerings that serve in a mental or educational capacity. Robots with big data analytics are used to create personalised learning experiences, tailored to students' individual abilities and needs (Rouhiainen, 2019). Social robots are used to create personalised learning experiences tailoring to individual students' abilities and needs (Belpaeme et al., 2018). Information processing services refer to information-based offerings, such as banking, accounting, and legal services. AI-powered chatbots provide round-the-clock information-based services, using advanced data analytics for fraudulent transactions and compliance improvement.

Robotic service is functioned through machine learning from the environment and past experience to improve decision making (Wu and Tegmark, 2019). It may also function through deep learning by mimicking the human brain to analyse data and draw conclusions manifested in image recognition, sound recognition, recommender systems, and natural language processing (Wu and Tegmark, 2019). However, the level of functional quality is manifested in technical outcomes by providing a consistent and timely service (West et al., 2018). For instance, robots in hotels are deployed to carry out concierge and room services and assist with check-in and check-out, order a taxi, answer customers' queries, and to refine responses based on customers' behaviour and feedback (Bartneck et al., 2009; Belanche et al., 2020; Solomon, 2016). Hilton Worldwide employs a robotic concierge named "Connie" to interact with customers just as a regular frontline employee. "Connie" has the capacity to personalise a customer's experience, by providing information and addressing general needs (Tavakoli & Mura, 2018). In the airline industry, KLM Airlines has adopted "Spencer" to answer travellers' queries and enhance their travel experience (West et al., 2018). The 1A-TA robot is used by travel agencies to identify travellers' needs and preferences for tailored service provision. In the retail industry, "Watson" is used by the North Face to support customers to choose the most suitable jackets by providing personalised recommendations based on an analysis of a large multivariance dataset. Robots provide quick and accurate personalised suggestions to enhance customer-brand interaction and also save on human labour costs within the organisation (Gursoy et al., 2019). These technical outcomes are reflective of service performance, acts, or efforts that are intentionally developed to increase value creation.

Robotic service quality

Any service that is offered for value exchange must be measured to assess its quality. In the marketing literature, the value of service is assessed based on the quality level and its impact on customers as any form of service should aim to generate positive customer responses and business profitability (Prentice, 2019; Zeithaml, 2000). In the services marketing literature, service quality is defined as the customer's perception or evaluation of a service organisation's overall excellence or superiority (Parasuraman et al., 1988). Empirical research has examined and confirmed a link between service quality, customer satisfaction, and loyalty (Helgesen, 2006; Kim, 2011; Woodside et al., 1989; Zeithaml, 2000). A high level of service quality leads to customer satisfaction and loyalty behaviours (Shi et al., 2014). Hence, assessing and evaluating service quality is imperative for an understanding of its impact on organisational outcomes.

Viewing robotic services as one type of customer service, this paper argues that the value must be reflective of the level of quality based on users' perceptions and evaluations. Various service quality measures and models have emerged over last three decades across different contexts, for instance, retail (Dabholkar et al., 1996; Siu & Tak-Hing Cheung, 2001), education (Brochado, 2009), health (Akter et al., 2013), casinos (Wong & Fong, 2012), and online services (Boshoff, 2007; Ding et al., 2011; Yang et al., 2004). Of these, models proposed by Gronroos (1982), Parasuraman et al. (1988), Parasuraman et al. (1985), and Brady and Cronin (2001) have attracted the most attention and have been extensively cited. Gronroos (date) assessed service quality from the technical and functional perspectives. The former refers to the outcome of the service performance or what the customer receives within the service encounter. The latter indicates the subjective perception of service delivery. Parasuraman et al.'s (1988) SERVQUAL describes service quality as the discrepancy between customers' expectations and perceptions of a service. Cronin and Taylor (1992) argued that adding customer expectation may be inefficient and unnecessary, and customers' perceptions alone are adequate to capture an assessment of the overall service quality (McDougall & Levesque, 1995). The rapid development of e-commerce has prompted numerous e-service quality scales in the literature. For instance, an e-service quality scale by Santos (2003), e-retailing service quality by Lee and Lin (2005), and Collier and Bienstock (2006), and an e-travel service quality scale by Ho and Lee (2007). These scales are reflective of the quality of service delivered in a virtual space.

Despite discrepancies in these measurements, these models consist of tangible and intangible services delivered through personal and impersonal encounters. A service encounter is dyadic with limited or narrow relational contact and communication (Gronroos, 1994). Personal encounters generally refer to interpersonal interactions between service employees who deliver the service and customers who receive the service (Inanov et al., 2018; Prentice, 2019). Impersonal encounters indicate, that in the absence of employee service, customers' interact with the physical or cyber/virtual/online settings of the service organisation (Chuah & Yu, 2021; Prentice, 2019; Yang et al., 2004). The physical settings include any tangible offerings, servicescape, ambience, or atmosphere. Online settings can be any virtual encounters.

Consistent with the foregoing discussion, a service quality measure should focus on customers' perceptions and evaluations of their encounters with service providers. Hence, robotic service quality should also be assessed by these encounters with customers and assessment must be drawn from customers perceptions and evaluations. The technology acceptance literature indicates that technology quality must be assessed from an information and systems perspective and evaluate how these factors affect users' beliefs, attitudes, and behaviours (Wixom & Todd, 2005; Xu et al., 2017). Consistent with this view, this study will approach service quality assessment from these two perspectives to develop a scale to measure robotic service quality. The following section describes the scale development procedure.

METHOD

Both qualitative and quantitative methods were employed to determine the items and dimensionality of robotic service. In accordance with conventional scale development protocols, three stages of development were undertaken across three countries (Australia, Vietnam, and China): 1) item generation, 2) scale refinement and validation, and 3) cross-validation. Details of three stages are now discussed.

Stage 1: Item generation and initial purification

Domain specification

Prior to identifying the dimensions and items for the assessment of robotic service quality, the domain of robotic service was specified (Churchill, 1979). First, relevant literature (including blogs, magazines, news, and industry reports) on artificial intelligence and robots were reviewed. Second, service robots were categorised according to physical and virtual forms. The former refers to frontline robots, including humanoid and non-humanoid robots, and can be symbolised as frontline service employees. The latter refers to virtual robots, symbolising employees offering remote or online services.

After reviewing existing service quality measures, SERVQUAL (Parasuraman et al., 1991) and e-SERVQUAL (Zeithaml et al., 2002) were drawn upon to determine the domains of robotic service quality. The SERVQUAL model has five core dimensions: reliability, assurance, empathy, responsiveness, and tangibility. The first four dimensions correspond to employee service with an emphasis on service promptness, accuracy, consistency, and employee friendliness and caring. The last dimension relates to the physical setting of the service, including the employees' appearance and other tangible components. This model was drawn upon to understand the encounter between robots and customers. Zeithaml et al.'s e-SERVQUAL has four core dimensions (efficiency, reliability, fulfilment, and privacy) and was adapted to determine virtual robotic service. As robots provide information-based services, Wixom and Todd's (2005) integrated user satisfaction and technology acceptance scale was adapted to represent the informational and technological aspects of robotic service. After scrutinising these scales and perusing web information relating to robots and artificial intelligence, three domains were determined to represent robotic service quality: automated service, humanised service, and informational service. Automated service represents mechanical robotic service such as mobile check-in/out in the hotel. Humanised service represents intuitive and empathetic services delivered by humanoid and non-humanoid robots, for instance, concierge robots, robot bartenders, robot chefs, porter robots, and vacuum cleaning robots. Informational service represents analytical robots that provide accurate and up to date information (such as chatbots). This process generated 54 items that represent robotic service.

Item generation

Two artificial intelligence experts and three operation managers from three hotels that have used robotic service were consulted to ensure the appropriateness and relevance of the items. As a result, 32 items remained. As service quality is formed on customers' perceptions and evaluations, 30 hotel customers who had used robotic services were invited to assess the importance of these items using a seven-point scale ranging from 1 "not important" to 7 "very important". Any item with an average score less than 5 or "somewhat important" were removed from the study (Brakus et al., 2009). This process resulted in 18 items for scale refinement and validation.

Stage 2: Scale refinement and validation

Sample and data collection

To refine these items, an online survey was conducted with Australian consumers who had experienced robotic service. The Qualtrics platform was used for data collection. The survey was designed to ensure that respondents were unable to skip questions. The target respondents must have used a robotic service within the last 12 months. To ensure the appropriateness of the sample, screening questions were developed to help prospective respondents understand the research purpose and filter those who did not fulfil the study criteria. Virtual snowball sampling was utilised for this study. This method relies on the virtual and social networks of participants and has the advantage of accessing hidden or hard-to-reach populations, hence, likely increasing the sample size and representativeness (Baltar and Brunet, 2012). Incentives (e.g. gift vouchers) were provided to prospective respondents to encourage participation (an anonymous link generated by Qualtrics) through their social media networks (e.g. Facebook). A total of 380 usable questionnaires were obtained. Of the respondents, 49.2% were male, 48.2% were between the ages of 26 to 45, and 59.7% held a bachelor's degree or higher. Table 1 presents the demographic information of the respondents.

Insert Table 1 here

Findings

Following the suggestion of Hair et al. (2014), Schumacker and Lomax (2004), and Costello and Osborne (2005), the dataset was randomly split into two subsets, one for exploratory factor analysis (EFA) and the other for confirmatory factor analysis (CFA) (see Table 2). The remaining items were factor analysed through promax oblique rotation in SPSS 26. This analysis revealed four factors after removing 6 items with low factor or cross loadings (see Appendix for original factor loadings of the eighteen items tested). The four-factor solution accounted for 83.57 of variation. The Kaiser-Mayer-Olkin value was .90, revealing satisfactory sampling adequacy. The primary loadings of all the remaining items were above .5 and the secondary loadings were less than .3. The coefficient alpha for each factor was above .7, indicating adequate internal consistency (Nunnally, 1994).

Based on the contents of the remaining items, four factors were labelled as automisation, personalisation, precision, and efficiency. Automisation had the highest eigenvalue (7.21), explaining 57.30% of the variance in robotic service quality, whereas precision has the lowest eigenvalue (1.06) and 5.23% of the variance. CFA was performed in AMOS 26 to confirm dimensionality and assess validity. The average variance extracted for all factors were above .50, indicating adequate convergence (Fornell & Larcker, 1981). The composite reliability for each factor was higher than the cut-off level of .70 as shown in Table 2.

Insert Table 2 here.

As suggested by Brakus et al. (2009), three competing models were analysed to examine the factor structure for this scale: 1) the one-factor model with all items loading on a single factor; 2) the first-order factor with the four dimensions as independent factors; and 3) the second-order factor with robotic service quality as the higher order. The results in Table 3 show that the second-order factor structure had the best model fit indices: χ^2= 86.19, def=46, χ^2 /df = 1.87, CFI = .99, RMSEA = .05, SRMR = .02. Consistent with the Akaike's information criterion (AIC) (Hu & Bentler, 1995), Model 3 was the most accurate and parsimonious with the lowest AIC value (see Table 3).

Insert Table 3 here.

The nomological validity of robotic service quality was examined by assessing the inter-factor correlations among the four factors. The results show significant correlations ranging from .51 to .76 (see Table 4). The criterion validity was tested by examining the relationship between AI service quality as the second-order construct and customer satisfaction, since the two are related as shown in the literature (Prentice, 2013; 2014; 2019). A customer satisfaction measure was adopted from Wixom and Todd (2005). The standardised path coefficient between robotic service quality and customer satisfaction was .69 ($p<.001$).

Insert Table 4 here

Internal cross-validation

The analyses show content validity, convergent validity, discriminant validity, nomological validity, and criterion validity (Hair et al., 2014). To cross-validate the scale, full metric and scalar invariance was assessed. The factor loadings were freely estimated for the two subsets of data in Model 1 and then constrained to be equal between the two groups in Model 2. The results in Table 5 show that the measurement model was invariant ($\Delta\chi^2$ (11)=12.03, p=.36; ΔCFI=.00), thus, full metric invariance was established.

For full scalar invariance, the dataset was divided into two groups by gender (see Wong and Fong, 2012). The full metric invariance by gender was examined. The results in Model 3 and 4 in Table 5 reveal that the factor loadings of the two groups were invariant ($\Delta\chi^2$ (11)=18.92, p=.06; ΔCFI=.00). A latent mean analysis was performed by constraining the intercepts of the structural equation of the observed variables on the latent factors which were equivalent across the two groups. The results in Model 5 suggest that the model fit of the latent mean model was not significantly different from the full metric invariant model ($\Delta\chi^2$ (10)=12.50, p=.25; ΔCFI=.00), thus, full scalar invariance was achieved. Results of the two invariances tests confirmed a strong factorial invariance of the robotic service quality scale.

Insert Table 5 here

External cross-validation

The literature contends that a new scale should be cross validated to be replicable and generalisable. Additional data was collected from Vietnam (350 usable responses) and China (423 usable responses) employing a virtual snowball sampling method. Table 1 shows the demographics of the two datasets.

Confirmatory factor analysis was performed to assess the measurement model for robotic service quality. The results show good model fit indices for both countries (Vietnam: χ^2 /df = 1.80, CFI = .97, RMSEA = .07, SRMR = .03; China: χ^2 /df = 3.39, CFI = .95, RMSEA = .08, SRMR = .04). To assess full metric invariance, datasets from Australia and Vietnam, and from Australia and China were used to perform a group invariance test by constraining the factor-loading estimates to be equal between the two groups. The results in Model 6 and 7 in Table 5 reveal that the model was invariant across the different sample cohorts between Australia and Vietnam ($\Delta \chi^2$ (11) =13.55, p=.26; ΔCFI=.00). The results in Model 8 and 9 reveal that the model was invariant between the Australian and Chinese data ($\Delta \chi^2$ (11) =15.71, p=.15; ΔCFI=.00). These results support the replicability of the scale.

Cross validation among the three datasets was conducted. Respondents (350) were randomly selected from data collected in China and Australia to validate the robotic service quality scale across countries. The relationship between robotic service quality as the second-order construct and user satisfaction were tested across the three datasets, Australia, Vietnam, and China (see Table 6). The standardised path coefficients between robotic service quality and customer satisfaction were significant across the three datasets, supporting the external validity of robotic service quality measurement.

Insert Table 6 here

DISCUSSION

This study draws upon service quality conceptualisation and developed a scale to measure robotic service quality. Service quality is formed as a result of customers' service encounter experience with the service provider. The service encounter, within the service quality literature, primarily refers to customers' interaction with service employees and the online or physical settings associated with the service organisation. Existing service quality measures fail to capture the services provided by robots that have been widely used in service industries. Robots may have physical and virtual forms. The former includes humanoid robots that look and behave like humans (e.g. Connie) and non-humanoid robots (e.g. mobile robots). The latter refers to online automated services. Service organisations utilise robotic services to improve customer experience and generate positive customer response. This study views robotic services as a service product and developed a robotic service quality scale. The interaction between customers and humanoid robots is referred to as a personal encounter, interactions between customers and other forms of robots are impersonal encounters.

Given the services provided by robots is largely information- and technology- based, the items that were generated to develop the scale were drawn from the service quality and technology acceptance literature to reflect customers' perception and evaluation of robotic service quality. Through a series of scale development processes, four dimensions were identified to represent robotic service quality: automisation, personalisation, precision, and efficiency. These dimensions are indicative of mechanical, analytical, intuitive, and empathetic artificial intelligence proposed by Huang and Rust (2018).

Automisation refers to the level of automation and performance by robots. This dimension mainly captures customers' experience with AI system quality. Sluggish operating systems underlying the robotic services evokes a negative customer response. Advanced technology today has changed customers' expectations of technology-related services. Personalisation indicates how robots can be flexible, like service employees, to meet customer's requests and demands. Service robots are designed to represent and behave like humans. Customers expect their interactions with these robots to be similar to those with service employees. Robots represent these employees and provide the requested information and function like employees. Customers may have different needs and wants from the service provider. Whilst employees' adaptability is expected, they are nevertheless limited to fixed working hours and oft times, inadequate training. Human physical and emotional constraints do not affect machine operated robots. These robots are available 24/7 and are capable of storing infinite amounts of information that can be adapted to meet customers' requests.

Precision is reflective of the accuracy and exactness of the information provided by robots. Such information is programmed through big data analytics and machine learning without unintended human intervention. Information generated by robots must be achieved in a timely manner and outdated information can be misleading. Customers expect information from robotic system to be updated, objective, thorough, and accurate. Precise information can guide customers to make informed decisions on consumption and purchase activities. Efficiency indicates timely and responsive service from robots. The ability to respond to customer requests dependably and reliably affects customer experience, attitudes, and subsequent behaviours.

IMPLICATIONS

Theoretical implications

This study developed a scale to measure robotic service quality. Establishment of this scale enriches the service quality and marketing literature. Numerous scales have been developed to measure service quality. Some are modelled to measure generic service quality (e.g. SERVQUAL, SERVPERF), others are more context based, for instance, e-travel service quality (Ho & Lee, 2007), retail service quality (Dabholkar et al., 1996), bank service quality (Karatepe et al., 2005), e-service quality (Li & Suomi, 2009), web information service quality (Yang et al., 2005), and healthcare service quality (Dagger et al., 2007). Given the pervasive use of robots for customer service in various service industries, it is imperative to understand how robotic service quality may affect customer attitudes and behaviours.

Service robots can deliver service similarly or beyond that of service employees. Therefore, robotic service quality can complement the other service quality scales in industries where robots are used. For instance, robots in HSBC can greet customers, guide them on the selection of appropriate banking services and summon an employee to help with transactions. The robot “Pepper” (Figure 1) can also pose for selfies and tell jokes. Robotic service quality can be used in conjunction with Karetepe et al’s (2005) service quality of retail banks to measure bank service quality and understand its impact on customer satisfaction and behaviour.

Insert Figure 1 here

Robotic service quality can also be integrated with Parasuraman et al.’s SERVQUAL, which consists of five dimensions (tangibility, reliability, assurance, responsiveness, and empathy). Each has relevance to service employees (Prentice, 2013). This scale has been widely used in the hospitality industry (Gabbie & O'Neill, 1996; Raspor, 2010; Saleh & Ryan, 1991; Shi et al., 2014). Robots in hotels (see Figure 2) are able to interact with guests and provide the requested information. In addition to measuring customer perceptions of hotel service quality on the five SERVQUAL dimensions, robotic service quality can be assessed to gain a wholistic picture of hotel service quality and its influence on customer experience (see Figure 2).

Insert Figure 2 here

Given that some robotic services are online information-based services, robotic service quality can also be integrated to e-service quality to present a broader perspective of the quality of web-based services. Existing e-service quality scales (Dabholkar et al., 1996; Madu & Madu, 2002; Zeithaml et al., 2002) fail to capture the service provided by chatbots that are available to answer queries and provide information around the clock. Chatbot services are ubiquitous and the topic is becoming increasingly popular in the literature (Dale, 2016; Feine et al., 2019; Jones & Jones, 2019; Lasek & Jessa, 2013). Understanding the service quality of chatbot services is conducive to improving customer experience. Robotic service quality addresses this void in the chatbot literature.

The study also adds a new measure to services and relationship marketing to manage customer relationships. Traditional service marketing strategies are primarily focused on the 7Ps (product, promotion, price, place, people, process, and physical evidence) and manifested in defensive (or relationship marketing) and aggressive approaches (Prentice, 2019). Service quality is considered a relationship marketing technique to elicit customer satisfaction and loyalty. Robotic service quality can be incorporated into the services and relationship marketing literature to better understand customer attitudes and behaviours.

Practical implications

Robotic service quality as developed in this study can be applied to relevant industries and organisations that extensively use service robots to provide customer service. Robots are mostly designed and developed by IT experts. Most users may be layman and not understand the underlying principles and operating systems. This scale helps service providers and customers appreciate the importance of robotic service without understanding the complex IT construction. Service providers can use this scale to assess return on investment as the purchase of robots can be expensive. Service providers may also gain a better understanding of the quality of these tools and their impact on customer service. Robotic service quality can help providers address the advantages and disadvantages of robotic service. From customer perspective, robotic service quality can be a useful tool for them to communicate with service providers to address their appreciation and concerns. This scale can also assist customers to provide feedback to the service provider. For example, some information provided by chatbots can be limited and unhelpful. Without proper measures, customers are unable to raise the issue with the service provider. This scale serves this purpose so that providers can improve robotic service quality to enhance customer experience and positive experiences.

LIMITATIONS

Despite endeavours to the contrary, limitations within this study are acknowledged. First, the sample frame was not representative of every service industry. Thus, the generalisability of this scale is limited. The sample frame could be extended to better represent the targeted population. Second, robotic service varies across different service contexts and industries. Development of this scale should differentiate between services to gain greater insights. Third, cross-validation in different countries should have taken the cultural effect into account. Finally, this study did not include a B2B context in the scale development process. In practice, robots have been used in these contexts to improve overall service quality for business customers. Future research should address these limitations.

Reference

Akter, S., D'Ambra, J., & Ray, P. (2013). Development and validation of an instrument to measure user perceived service quality of mHealth. *Information & Management, 50*(4), 181-195. doi: https://doi.org/10.1016/j.im.2013.03.001

Baltar, F., Brunet, I., 2012. Social research 2.0: virtual snowball sampling method using Facebook. Internet Res. 22 (1), 57–74. https://doi.org/10.1108/ 10662241211199960.

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

Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science robotics*, *3*(21), 1-9.

Belanche, D., Casaló, L. V., & Flavián, C. (2020). Frontline robots in tourism and hospitality: service enhancement or cost reduction?. *Electronic Markets*, 1-16.

Boshoff, C. (2007). A psychometric assessment of E-S-QUAL: A scale to measure electronic service quality. *Journal of Electronic Commerce Research*, *8*(1), 101.

Brady, M. K., & Cronin, J. J. (2001). Some new thoughts on conceptualizing perceived service quality: A hierarchical approach. *Journal of marketing*, *65*(3), 34-49. doi: https://doi.org/10.1509/jmkg.65.3.34.18334

Brakus, J. J., Schmitt, B. H., & Zarantonello, L. (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of marketing*, *73*(3), 52-68. doi: https://doi.org/10.1509/jmkg.73.3.52

Brochado, A. (2009). Comparing alternative instruments to measure service quality in higher education. *Quality Assurance in Education, 17*(2), 174-190. doi: https://doi.org/10.1108/09684880910951381

Churchill, G. A. (1979). A Paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(1), 64-73. doi: https://doi.org/10.2307/3150876

Cobos, L. M., Mejia, C., Ozturk, A. B., & Wang, Y. (2016). A technology adoption and implementation process in an independent hotel chain. *International Journal of Hospitality Management, 57*, 93-105. doi: https://doi.org/10.1016/j.ijhm.2016.06.005

Collier, J. E., & Bienstock, C. C. (2006). Measuring service quality in e-retailing. *Journal of service research*, *8*(3), 260-275.

Cronin, J. J., & Taylor, S. A. (1992). Measuring service quality: A reexamination and extension. *Journal of marketing, 56*(3), 55-68. doi: https://doi.org/10.1177/002224299205600304

Dabholkar, P. A., Thorpe, D. I., & Rentz, J. O. (1996). A measure of service quality for retail stores: Scale development and validation. *Journal of the Academy of Marketing Science*, *24*(1), 3-16. doi: https://doi.org/10.1007/BF02893933

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

Dale, R. (2016). The return of the chatbots. *Natural Language Engineering, 22*(5), 811-817. doi: https://doi.org/10.1017/S1351324916000243

Ding, D. X., Hu, P. J.-H., & Sheng, O. R. L. (2011). e-SELFQUAL: A scale for measuring online self-service quality. *Journal of Business Research, 64*(5), 508-515. doi: https://doi.org/10.1016/j.jbusres.2010.04.007

Farooq, R. (2016). Role of structural equation modeling in scale development. *Journal of advances in management research, 13*(1), 75-91. doi: https://doi.org/10.1108/jamr-05-2015-0037

Feine, J., Morana, S., & Gnewuch, U. (2019). *Measuring service encounter satisfaction with customer service chatbots using sentiment analysis*. Paper presented at the in Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI2019), Siegen, Germany.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. doi: https://doi.org/10.2307/3151312

Gabbie, O., & O'Neill, M. A. (1996). SERVQUAL and the Northern Ireland hotel sector: a comparative analysis - part 1. *Managing Service Quality: An International Journal, 6*(6), 25-32. doi: https://doi.org/10.1108/09604529610149194

Gartner (2019). Gartner survey shows 37 percent of organizations have implemented AI in some form. Retrieved from https://www.gartner.com/en/newsroom/press-releases/2019-01-21-gartner-survey-shows-37-percent-of-organizations-have

Gronroos, C. (1982). *Strategic Management and Marketing in the Service Sector:* . London: Chartwell-Bratt.

Gursoy, D., Chi, O. H., Lu, L., & Nunkoo, R. (2019). Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management, 49*, 157-169. doi: https://doi.org/10.1016/j.ijinfomgt.2019.03.008

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (Seventh, Pearson new international ed.). Harlow, Essex: Pearson Education Limited.

Helgesen, Ø. (2006). Are loyal customers profitable? Customer satisfaction, customer (action) loyalty and customer profitability at the individual level. *Journal of Marketing Management, 22*(3-4), 245-266. doi: https://doi.org/10.1362/026725706776861226

Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E., & Schlesinger, L. A. (1994). Puting the service-profit chain to work*72*(2), 164-174. Retrieved from https://hbr.org/2008/07/putting-the-service-profit-chain-to-work

Ho, C.-I., & Lee, Y.-L. (2007). The development of an e-travel service quality scale. *Tourism management, 28*(6), 1434-1449. doi: https://doi.org/10.1016/j.tourman.2006.12.002

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage.

Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155-172.

Ivanov, S., Webster, C. and Berezina, K. (2018). Adoption of robots and service automation by tourism and hospitality companies. Revista Turismo & Desenvolvimento, 27/28, 1501-1517.

Jones, B., & Jones, R. (2019). Public service chatbots: Automating conversation with BBC news. *Digital Journalism, 7*(8), 1032-1053. doi: https://doi.org/10.1080/21670811.2019.1609371

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons, 62*(1), 15-25. doi: https://doi.org/10.1016/j.bushor.2018.08.004

Karatepe, O. M., Yavas, U., & Babakus, E. (2005). Measuring service quality of banks: Scale development and validation. *Journal of Retailing and Consumer Services, 12*(5), 373-383. doi: https://doi.org/10.1016/j.jretconser.2005.01.001

Kim, H. J. (2011). Service orientation, service quality, customer satisfaction, and customer loyalty: Testing a structural model. *Journal of Hospitality Marketing & Management, 20*(6), 619-637. doi: https://doi.org/10.1080/19368623.2011.577698

Kumar, K. N., & Balaramachandran, P. R. (2018). Robotic process automation - A study of the impact on customer experience in retail banking industry. *Journal of Internet Banking and Commerce, 23*(3), 1-27.
Kuo, C.-M., Chen, L.-C. and Tseng, C.-Y. (2017). Investigating an innovative service with hospitality robots. International Journal of Contemporary Hospitality Management, 29(5), 1305-1321.
Lasek, M., & Jessa, S. (2013). Chatbots for customer service on hotels' websites. *Information Systems in Management, 2*(2), 146-158.
Li, H., & Suomi, R. (2009). A proposed scale for measuring e-service quality. *International Journal of u-and e-Service, Science and Technology, 2*(1), 1-10.
Lee, G. G., & Lin, H. F. (2005). Customer perceptions of e- service quality in online shopping. *International Journal of Retail & Distribution Management*.
Madu, C. N., & Madu, A. A. (2002). Dimensions of e- quality. *International Journal of Quality & Reliability Management*.
Marinova, D., de Ruyter, K., Huang, M.-H., Meuter, M. L., & Challagalla, G. (2017). Getting smart: Learning from technology-Empowered frontline interactions. *Journal of Service Research, 20*(1), 29-42. doi: https://doi.org/10.1177/1094670516679273
McDougall, G. H. G., & Levesque, T. J. (1995). A revised view of service quality dimensions: An empirical investigation. *Journal of Professional Services Marketing, 11*(1), 189-210. doi: https://doi.org/10.1300/J090v11n01_13
Nunnally, J. C. (1994). *Psychometric theory (3rd ed.)* (3rd ed.). New York: McGraw-Hill.
Ozment, J., & Morash, E. A. (1994). The augmented service offering for perceived and actual service quality. *Journal of the Academy of Marketing Science, 22*(4), 352-363. doi: https://doi.org/10.1177/0092070394224004
Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing, 67*(4), 420-450.
Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of marketing*, *49*(4), 41-50. doi: https://doi.org/10.1177/002224298504900403
Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual: A Multiple-Item Scale For Measuring Consumer Perc. *Journal of Retailing*, *64*(1), 12.
Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1994). Alternative scales for measuring service quality: a comparative assessment based on psychometric and diagnostic criteria. *Journal of retailing*, *70*(3), 201-230.
Prentice, C. (2019). *Emotional intelligence and marketing*: World Scientific.
Prentice, C., Dominique Lopes, S., & Wang, X. (2020). The impact of artificial intelligence and employee service quality on customer satisfaction and loyalty. *Journal of Hospitality Marketing & Management*, *29*(7), 739-756.
Raspor, S. (2010). Measuring perceived service quality using SERVQUAL: A case study of the Croatian hotel industry. *Management (18544223), 5*(3).
Rajesh, M. (2015). Inside Japan's first robot-staffed hotel. Retrieved from https://www.theguardian.com/travel/2015/aug/14/japan-henn-na-hotel-staffed-by-robots
Rouhiainen, L. (2019). How AI and data could personalize higher education. Retrieved from https://hbr.org/2019/10/how-ai-and-data-could-personalize-higher-education
Russell, S. J., Norvig, P., & ProQuest. (2016). *Artificial intelligence: a modern approach* (Global;3rd; ed.). Harlow, Essex, England: Pearson Education Limited.
Saleh, F., & Ryan, C. (1991). Analysing service quality in the hospitality Industry using the SERVQUAL model. *The Service Industries Journal, 11*(3), 324-345. doi: https://doi.org/10.1080/02642069100000049

Santos, J. (2003). E- service quality: a model of virtual service quality dimensions. *Managing Service Quality: An International Journal*.
Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, N.J: Lawrence Erlbaum Associates.
Shek, D. T. L., & Yu, L. (2014). Use of structural equation modeling in human development research. *International Journal on Disability and Human Development, 13*(2), 157-167. doi: https://doi.org/10.1515/ijdhd-2014-0302
Shi, Y., Prentice, C., & He, W. (2014). Linking service quality, customer satisfaction and loyalty in casinos, does membership matter? *International Journal of Hospitality Management, 40*, 81-91. doi: https://doi.org/10.1016/j.ijhm.2014.03.013
Siu, N. Y. M., & Tak-Hing Cheung, J. (2001). A measure of retail service quality. *Marketing Intelligence & Planning, 19*(2), 88-96. doi: https://doi.org/10.1108/02634500110385327
Tavakoli, R., & Mura, P. (2018). Netnography in tourism – Beyond Web 2.0. *Annals of Tourism Research, 73*, 190-192. doi: https://doi.org/10.1016/j.annals.2018.06.002
West, A., Clifford, J., & Atkinson, D. (2018, 2018). *"Alexa, build me a brand" An investigation into the impact of artificial intelligence on branding*, London.
Wirtz, J., & Lovelock, C. (2018). *Developing service products and brands*: WS Professional.
Wixom, B. H., & Todd, P. A. (2005). A theoretical integration of user satisfaction and technology acceptance. *Information Systems Research, 16*(1), 85-102. doi: https://doi.org/10.1287/isre.1050.0042
Wong, I. A., & Fong, V. H. I. (2012). Development and validation of the casino service quality scale: CASERV. *International Journal of Hospitality Management, 31*(1), 209-217. doi: https://doi.org/10.1016/j.ijhm.2011.04.005
Woodside, A. G., Frey, L. L., & Daly, R. T. (1989). Linking sort/ice anlity, customer satisfaction, and behavioral intention. *Journal of health care marketing, 9*(4), 5-17.
Yang, Z., Cai, S., Zhou, Z., & Zhou, N. (2005). Development and validation of an instrument to measure user perceived service quality of information presenting Web portals. *Information and Management, 42*(4), 575-589. doi: https://doi.org/10.1016/S0378-7206(04)00073-4
Yang, Z., Jun, M., & Peterson, R. T. (2004). Measuring customer perceived online service quality. *International Journal of Operations & Production Management, 24*(11), 1149-1174. doi: https://doi.org/10.1108/01443570410563278
Zeithaml, V. A. (2000). Service quality, profitability, and the economic worth of customers: What we know and what we need to learn. *Journal of the Academy of Marketing Science, 28*(1), 67-85. doi: https://doi.org/10.1177/0092070300281007
Zeithaml, V. A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science, 30*(4), 362-375. doi: https://doi.org/10.1177/009207002236911

Table 1. Demographic profile (N = 380)

Characteristics		Australia	Vietnam	China
Gender	Male	187 (49.2)	219 (63%)	233 (55%)
	Female	193 (50.8)	129 (37%)	190 (45%)
Age	18-35	142 (37.4)	192 (55%)	332 (79%)
	35-46	183 (48.2)	139 (40%)	87 (20%)
	>55	55 (14.4)	19 (5%)	4 (1%)
Education	< High school	5 (1.3)	10 (3%)	17 (4%)
	Vocational school	66 (17.4)	53 (15%)	61 (14%)
	Bachelor	76 (20)	105 (30%)	300 (71%)
	Master or higher	141 (37.1)	155 (44%)	43 (10%)
	Others	86 (24.2)	27 (8%)	2 (1%)

Note: Values in the brackets are percentage

Table 2. Item descriptions and measurement model results for robotic service quality

Scale items	Pattern coefficient	Variance explained	Cronbach's alpha	CR	AVE
Automation	β = 7.21	57.30%	.86	.85	.65
Robots operates reliably	.77				
Robots performs effectively	.85				
Robots functions dependably	.79				
CFI = 1.00; TLI= 1.00; IFI = 1.00; SRMR = .00					
Personalisation	β = 1.18	6.13%	.89	.88	.72
Robotic service is adaptive to meet a variety of my needs	.85				
Robotic service is flexibly adjusted to meet my new demands	.87				
Robotic service is versatile in addressing my needs	.70				
CFI = 1.00; TLI= 1.00; IFI = 1.00; SRMR = .00					
Efficiency	β = 2.34	11.08%	.91	.91	.77
Robots are very responsive to my requests	.90				
Robots provide service in a timely manner	.82				
Robots solve my problems effectively	.91				
CFI = 1.00; TLI= 1.00; IFI = 1.00; SRMR = .00					
Precision	β = 1.06	5.23%	.85	.83	.63
Information from robotic system is accurate	.83				
Information from robotic system is reliable	.85				
Information from robotic system is always up to date	.68				
CFI = 1.00; TLI= 1.00; IFI = 1.00; SRMR = .00					

β= eigenvalues; CR = composite reliability; AVE = average variance extracted

Table 3. Confirmatory factor analysis model fit comparisons

Model	Description	χ^2	df	χ^2 /df	p	CFI	RMSEA	SRMR	AIC
1	One factor	550.77	54	10.20	.00	.86	.16	.05	598.77
2	First-order factor	125.79	38	3.31	.00	.98	.08	.03	181.79
3	Second-order factor	86.19	46	1.87	.00	.99	.05	.02	150.19

Table 4. Correlation coefficients, means, and standard deviations of robotic service quality and customer satisfaction

	Means	SD	1	2	3	4
1.Automisation	4.93	1.10				
2.Personalisation	4.94	1.12	.72**			
3.Efficiency	4.10	1.53	.59**	.51**		
4. Precision	4.95	1.12	.76**	.69**	.57**	
5.Cust. satisfaction	4.81	1.27	.66**	.60**	.61**	.60**

**p<.01

Table 5. Test of measurement model invariance

Model	Description	Compare with	χ^2	df	$\Delta\ \chi^2$ (Δdf)	p	CFI	ΔCFI	RMSEA	SRMR
1	Baseline model (G1/2)	-	154.20	90	-	-	.98	-	.05	.03
2	Factor pattern (G1/2)	Model 1	169.66	101	12.03(11)	.36	.98	.00	.04	.04
3	Baseline model (M/F)	-	157.19	90	-	-	.98	-	.04	.03
4	Factor pattern (M/F)	Model 3	176.11	101	18.92(11)	.06	.98	.00	.04	.03
5	Scalar invariance (M/F)	Model 4	188.61	111	12.50(10)	.25	.98	.00	.04	.04
6	Based model (Aus/VN)	-	300.50	90	-	-	.96	-	.06	.04
7	Factor pattern (Aus/VN)	Model 6	314.05	101	13.55(11)	.26	.96	.00	.04	.04
8	Based model (Aus/CN)	-	240.12	90	-	-	.97	-	.05	.04
9	Factor pattern (Aus/CN)	Model 8	255.83	101	15.71(11)	.15	.97	.00	.04	.05

Note: M=male, F=female, Aus=Australia, VN- Vietnam, CN-China.

Table 6. External cross-validation

		Customer satisfaction		
		Australia	Vietnam	China
Robotic service quality	Australia	.78***	.31*	.85***
	Vietnam	.50***	.57***	.44***
	China	.70***	.15*	.68***

*p<.05, **p<.01, ***p<.001

Figure 1: SoftBank Robotics' humanoid robot, Pepper in HSBC at the Fifth Ave branch on Monday, June 25, 2018 in New York. Mark Von Holden | AP Images for HSBC

Figure 2: Connie in Hilton interacting with hotel guests.