Measuring SERVQUAL dimensions and their importance for customer-satisfaction using online reviews: a text mining approach

Received 11 June 2021
Revised 5 February 2022
Accepted 15 February 2022

Swagato Chatterjee
Vinod Gupta School of Management, Indian Institute of Technology Kharagpur, Kharagpur, India
Arpita Ghatak
Rajendra Mishra School of Engineering Entrepreneurship, Indian Institute of Technology Kharagpur, Kharagpur, India
Ratnadeep Nikte
Indian Institute of Technology Kharagpur, Kharagpur, India
Shivam Gupta
Department of Information Systems, Supply Chain Management and Decision Support, NEOMA Business School, Reims, France, and
Ajay Kumar
AIM Research Centre on Artificial Intelligence in Value Creation, EMLYON Business School, Ecully, France

Abstract

Purpose – The extant literature has utilized the SERVQUAL scale to measure service quality dimensions and their importance towards customer-satisfaction using close-ended survey-based questions and not open-ended questions and/or user-generated qualitative responses. On the other hand, while measuring customer-satisfaction drivers from user-generated content (UGC), extant studies have majorly used overall or aspect-wise evaluations and not evaluations specific to SERVQUAL dimensions. In this study, the authors try to bridge the gap.

Design/methodology/approach – The authors suggest a methodology consisting of text mining, machine learning and econometric techniques that can measure consumer evaluations of SERVQUAL dimensions. The authors used qualitative and quantitative UGC obtained from 27,052 online reviews on 362 airlines by reviewers of 158 nationalities for our analysis.

Findings – The authors established a unique method which combines qualitative and qualitative UGC to measure service quality. The authors have also uncovered the comparative importance of such dimensions in creating customer-satisfaction and recommendation in the context of the airline industry.

Originality/value – The paper is one of the pioneering studies that try to find measures of SERVQUAL dimensions from online consumer reviews and their influence on customer satisfaction.

Keywords SERVQUAL, Customer satisfaction, User generated content, Text mining, Sentiment analysis

Paper type Research paper

The paper is funded by "Sponsored Research and Industry Consultancy, Indian Institute of Technology Kharagpur" with the project id UAE.

Journal of Enterprise Information Management

1741-0398
DOI 10.1108/JEIM-06-2021-0252

1. Introduction

With the advent of the Internet, consumers' preference building and decision-making processes have seen a significant change over the last few decades. Online reviews shared by others about their experience profoundly influence consumer decisions. Reliable Internet connectivity and the massive increase in smartphone usage are also behind the change in the way consumers obtain information about goods and services. This also resulted in a steep rise in the availability of various products and services via the online medium (Xu *et al.*, 2013). Thus, relying on online data sources for service information has become more prevalent in current days.

As the antecedent of customer satisfaction and purchase intention, customer service plays a pivotal role in the success of organizations (Bairi and Manohar, 2011; Cenfetelli *et al.*, 2008; Delone and Mclean, 2004; Xu *et al.*, 2013; Wang, 2008), more so in the online environment. The success of IT-enabled customer services is forcing marketers to change the conventional views of keeping customers only as simple services' receivers (Chesbrough and Spohrer, 2006) and pushing them to give more emphasis on customer loyalty and customer satisfaction. Researchers have shown, with the help of technological innovation, companies can provide customized, high-quality services at affordable prices (Rust and Miu, 2006). Therefore, it is very important for the organizations to periodically evaluate the customer satisfaction, perceived service quality and loyalty intentions of the organizations. Scholars have proposed that perceived service quality measure is context-dependent and therefore separate scales should be made for various industries and situations (Liou *et al.*, 2011; Zeithaml *et al.*, 1990). Nevertheless, creating separate scales for each context may be very difficult. This leads to the need of formulation of a mechanism which can dynamically measure the dimensions of service quality in a context-dependent manner. The current study tries to fill this gap. Moreover, while trying to measure customers' satisfaction, organizations face difficulties in getting reliable and valid data. But now, customers are willingly and spontaneously sharing their service experience on social media and different review platforms (He *et al.*, 2019; Mudambi and Schuff, 2010). This user-generated content (UGC) helps organizations to access an enormous amount of reliable data, which in turn allows organizations to get reliable customer-satisfaction and service quality measurement (Weigend, 2009).

From the existing literature on online reviews, we have found that majority of the studies have focused on the influence of online reviews on customers' decisions. A few researchers have tried to explore different ways to use intelligence extracted from UGC to accumulate data for organizations, which creates a lacuna in the existing literature (He *et al.*, 2019; Mudambi and Schuff, 2010; Palese and Usai, 2018). We fill this gap in the literature by proposing a method which uses UGC to generate measurements of perceived service quality and its influence on customer satisfaction. Previous researchers have also mentioned that reliable service quality can be measured using customers' experience (Petter *et al.*, 2012). On the other hand, it isn't easy to extract valuable insight from a large volume of data (McAfee *et al.*, 2012).

In our study, we have shown how we can use elements of online textual reviews to measure customer satisfaction and acquire a meaningful insight. Initially, we have used text mining and sentiment analysis to extract relevant data from online reviews to measure service performance accurately. In this study, we have successfully obtained all the SERVQUAL dimensions (Parasuraman *et al.*, 1988) from the UGC. As we have measured these dimensions with regards to performance, our method is more relevant to SERVPERF than SERVEQUAL (Palese and Usai, 2018). Next, we have demonstrated that these dimensions do not have similar influences on customer recommendation and satisfaction. Unlike our work, prior studies have not focused on the impact of each dimension separately but looked into the combined service measure (Luo *et al.*, 2012). Moreover, only one study which focuses on individual measures did not use consumer sentiments expressed about the

dimensions to generate SERVQUAL/SERVPERF dimension level information (Palese and Usai, 2018). This makes our work unique. We have also checked the predictive power of the machine learning models with and without such dimensions. The paper contributes towards the existing literature on service quality, SERVQUAL, customer satisfaction and recommendation (Cenfetelli *et al.*, 2008; Chesbrough and Spohrer, 2006; Delone and Mclean, 2004; Rust and Miu, 2006; Xu *et al.*, 2013; Wang, 2008) by combining the usage of qualitative data with quantitative data in measuring service quality dimensions and uncovering their inter-connection with customer satisfaction and loyalty (Palese and Usai, 2018). The study strengthens the attitude-behavior paradigm of satisfaction and loyalty (Kumar *et al.*, 2013), the expectancy-value framework of attitude formation and the theory of planned behavior for behavior-formation (Ajzen and Fishbein, 2000), the accessibility-diagnosticity (AD) models (Chatterjee, 2019; Lynch, 2006; Siering *et al.*, 2018) etc. Finally, by providing a better methodology to predict customer satisfaction and recommendation by combining quantitative ratings and qualitative reviews and at the same time exploring the theoretical explanations behind the same, the paper contributes to explainable machine learning and artificial intelligence domain (Rai, 2020). Managerially, this is the first paper which suggests a mechanism to measure SERVQUAL/SERVPERF dimensions using the textual review data and not survey-based Likert scale data. The paper also helps the managers to understand which SERVQUAL/SERVPERF dimensions they should focus on or which dimension is more important for their customers and how such relative importance varies based on business context and business goals.

We study the above in airline satisfaction as the airline is a multi-attribute service context with a complex mix of service aspects leading to overall satisfaction and recommendations (Chatterjee, 2019). Moreover, consumer reviews play a pivotal role in customer purchase decisions and the airline's service design management. However, the below-mentioned methodology is unique and applicable beyond the context of airlines too.

In the following sections of this article, we showcase the literature gap, followed by the theoretical framework. Next, we detail the methodology and findings, followed by the discussions, theoretical contributions, managerial implications, limitations and future scopes.

2. Literature review

2.1 Service quality, customer satisfaction and loyalty

Service quality plays a crucial role behind customers' purchase decisions (Anderson and Zeithaml, 1984) as it is positively related to customer satisfaction and strengthens customer loyalty (Szwarc, 2005), which leads to an increase in profits (Szwarc, 2005; Hamouda, 2019). Researchers have pointed out that serving better than competitors can create a firm ground for customer loyalty (Curry and Gao, 2012). Researchers have also mentioned the perceived price and the tangible environment as the necessary components which impact customer satisfaction and customer loyalty (Han and Ryu, 2009). In his paper, Liou *et al.* (2011) demonstrated that it is not possible to converge to a universal definition to define service quality as service quality is context-dependent. Hence, it is necessary to know what consumers perceive as high-quality service for a particular industry (Zeithaml *et al.*, 1990). For instance, in the case of the airline industry, researchers have investigated the connection between service quality and customer satisfaction and loyalty (Chen and Hu, 2013; Curry and Gao, 2012; Hamouda, 2019; Namukasa, 2013; Ostrowski *et al.*, 1993). Chen and Chang (2005) have divided the entire airline service delivery into two levels: ground services and in-flight services. Ostrowski *et al.* (1993) have suggested to consider the carrier image, along with these two above mentioned aspects. On the other hand, Gourdin (1988) divided the airline service quality into three dimensions: safety, timeliness and price. However, airline passengers have reported the following as significant important issues of airline service quality-safety, punctuality, frequency, cabin service, the fine for ticket alteration and in-flight comfort in

seating arrangement ability (Chen and Chang, 2005; Gilbert and Wong, 2003; Liou and Tzeng, 2007; Liou *et al.*, 2011; Martin *et al.*, 2011; Mason, 2001). The above discussion suggests that the measurement of service quality needs to be more dynamic and contextual. However, the above cannot be achieved if a pre-defined scale is used to measure service quality instead of contextualization of the scale. Moreover, it is impossible to create separate scales for all possible industries. This creates a gap and develops a need of formulation of a methodology which can produce context-dependent dynamic measures of service quality. Through our paper we are trying to do the same.

The positive/negative disconfirmation of the perceived product or service quality from the expected quality leads to a feeling of pleasure/disappointment leading to customer satisfaction (Oliver, 1981; Tse and Wilton, 1988). Therefore, service quality and satisfaction are found to be related. On the other hand, the behavioral outcome of satisfaction is customer loyalty, which plays a crucial role in case of services industry such as airlines (Chatterjee, 2019). Customer loyalty can be defined as multiple behavioral responses of the consumers, which includes purchasing repeatedly, patronization and positive word-of-mouth (Ndubisi *et al.*, 2007). Services industries have considered all these forms of loyalty to be very important for their strategic performance, and therefore, they keep on running loyalty programs to ensure the above. For instance, the frequent flier program (FFP), introduced by American Airlines, is one such loyalty program that rewards purchasing-repeatedly/frequent-travel by free upgrades, discounts, lounge services, additional luggage allowances, etc (Martin *et al.*, 2011). With the increase in the importance of strategic alliances, the FFP benefits have now been extended to eCommerce, hotels, restaurants, retail chains, etc. Such programs and benefits lead consumers to choose the preferred service organizations (Seelhosrt and Liu, 2015). However, loyalty is not always bought by such programs. The extant literature has time and again established that service experiences have been a major antecedent of loyalty with satisfaction acting as a mediator (Anderson and Jacobsen, 2000; Calisir *et al.*, 2016; Namukasa, 2013). Thus, finding the impact of individual service components on loyalty and satisfaction becomes a crucial area of research. The existing literature focusing on this area has used survey-based methods predominantly (Anderson and Jacobsen, 2000; Calisir *et al.*, 2016; Namukasa, 2013). Such a method has its own bias. Moreover, the data are often quantitative and collected through closed-form questions. However, in the advent of the Internet, UGC becomes an important and rich source of information (Chatterjee, 2019). However, little study has been done on the inter-connection of service quality dimensions, loyalty and customer satisfaction using textual data from UGC on the Internet (Palese and Usai, 2018). Our study fills this literature gap.

2.2 Customer reviews as a source of information

Past studies have extensively focused on how online reviews influence sales (Ghose and Ipeirotis, 2006; Hsu, 2021; Hu *et al.*, 2008; Zhu and Zhang, 2010), trust with the business (Ba and Pavlou, 2002; Pavlou and Gefen, 2004; Pavlou and Dimoka, 2006) and even the helpfulness on the reviews (Hsu, 2021; Ghose and Ipeirotis, 2010; Mudambi and Schuff, 2010). A few studies also focus on how reviews posted by peers influence online reviews (Kumar and Benbasat, 2006; Senecal and Nantel, 2004). On the other hand, the existence of customer reviews in an eCommerce website improves consumer participation and their perception about the usefulness of the website (Kumar and Benbasat, 2006). Moreover, online reviews also impact customer traffic/footfall, consumer-community building and consumer purchase decision-making process (Dabholkar, 2006; Jiang and Benbasat, 2004; Gharib *et al.*, 2020; Kohli *et al.*, 2004). However, we are focusing on the influence of online reviews on the reviewers' own evaluations and recommendation behavior.

One important aspect of online reviews which makes them different from classical word-of-mouth is the IT-mediation, i.e. the effect of information technology. This allows

information exchange in a large quantity and in superficial level, unlike face-to-face word-of-mouth exchange where the information-sharing is deep and more personal. Therefore, in the case of online reviews, readers often using their own simplified strategies to get the overall information from multiple reviews (Ghose and Ipeirotis, 2006). Moreover, often online reviews are found to be bimodal, making average rating a not so good measure of central tendency of the reviews (Hu *et al.*, 2006; Jabr and Zheng, 2014). On the other hand, while a lot of studies have focused on online reviews, very few of them have used online reviews as a mechanism to measure the customer service quality for an organization (Palese and USai, 2018). Therefore, both customers and organizations receive an incomplete and biased picture of service quality when they only focus on quantitative ratings readily available in the reviews. This creates a problem in the decision making of future customers. At the same time, it creates a hindrance for service managers to get an idea on which service aspect is important and needs immediate attention, which the extant literature cannot solve. In our paper, we propose a solution to the above problem by using rich UGC in the form of textual consumer reviews.

In line with our study, a few scholars have explored the online reviews using sentiment analysis leading to an understanding about the antecedent of the users' general evaluation and review writing (Duan *et al.*, 2013). In another attempt, topic modeling has been done to find the SERVQUAL/SERVPERF dimensions on overall satisfaction (Palese and Usai, 2018). Our study expands this previous effort by utilizing the text of online reviews to measure the dimensions of perceived service quality and analyze their influence on customer satisfaction.

3. Theoretical framework

3.1 Service quality, customer satisfaction and recommendation

Quality assessment is a crucial interdisciplinary area of research involving operations research, information systems and marketing. Initially, scholars concentrated on measuring the quality of tangible goods. Later on, the unique characteristics of service like intangibility and heterogeneity, motivated researchers to focus on developing methods to measure service quality (Grönroos, 1984; Parasuraman *et al.*, 1985).

The extant literature on service quality provides a diverse range of definitions. One set of researchers have seen service quality as a combination of technical and functional quality. Technical quality talks about everything that customers are actually getting from service providers. In addition, the functional quality – how the service has been delivered (Grönroos, 1982). According to another group of researchers, service is co-created among providers and the receivers. Moreover, it has three dimensions (Lehtinen and Lehtinen, 1991): corporate quality (organization's image), physical quality (tangible features of the service) and interactive quality (communication among customers and contact-personnel). The SERVQUAL model (Parasuraman *et al.*, 1985) focuses on the gap between the expectations of the customer before receiving any service and their perceptions of the actual service delivered. Initially, it had ten dimensions of service quality; after tuning it, researchers have reduced SERVQUAL (Parasuraman *et al.*, 1988) into five dimensions. These dimensions are-reliability (accuracy of the service); responsiveness (willingness to provide prompt service and quick responding to customer questions or complaints); tangibles (the impression of the physical facility, personnel, communication materials on the buyer and equipment); assurance (the knowledge and behavior of workers which generates trust and credibility for customers); and empathy (provision of care, individualized attention that organization provides to its buyers). Extensive research has concentrated on testing SERVQUAL and on the development of scales to estimate service quality (Ladhari, 2009; Tian *et al.*, 2020). After multiple validations across different industries, SERVQUAL continues to be the most practiced instrument to evaluate service quality for academia and industry (Asubonteng *et al.*, 1996; Ladhari, 2009; Tian *et al.*, 2020). However, over the years, it has received some

criticism, especially by Cronin and Taylor (1992). They have developed the SERVPERF model to estimate customers' perceptions of the organization's quality of service. We were not intended to argue which model is more suitable (Lee *et al.*, 2000). Our target was to notify that SERVPERF and SERVQUAL are based on the same dimensions.

Instead, we intended to use those same dimensions to examine their pertinence in customer-generated review data available on social portals. Our contribution in this paper is to algorithmically extract the service quality dimensions from the reviews that customers voluntarily have written on the web; instead of measuring the dimensions from the survey data. For our study, we have chosen the SERVQUAL, as numerous researchers extensively used it as an investigation instrument (Davies *et al.*, 1999).

Our first objective was to demonstrate that applying bag-of-words techniques if we can extract the dimensions of SERVQUAL/SERVPERF model directly from the online customer reviews. Second, we wanted to study the association between the SERVQUAL/SERVPERF dimensions with customer satisfaction and recommendations in the airline context. Figure 1 explains the theoretical model for our paper. The hypotheses we are testing are

H1. Tangibles, as measured from textual reviews, have positive relationships with (a) customer satisfaction and (b) customer loyalty.

H2. Assurance, as measured from textual reviews, has positive relationships with (a) customer satisfaction and (b) customer loyalty.

H3. Empathy, as measured from textual reviews, has positive relationships with (a) customer satisfaction and (b) customer loyalty.

H4. Reliability, as measured from textual reviews, has positive relationships with (a) customer satisfaction and (b) customer loyalty.

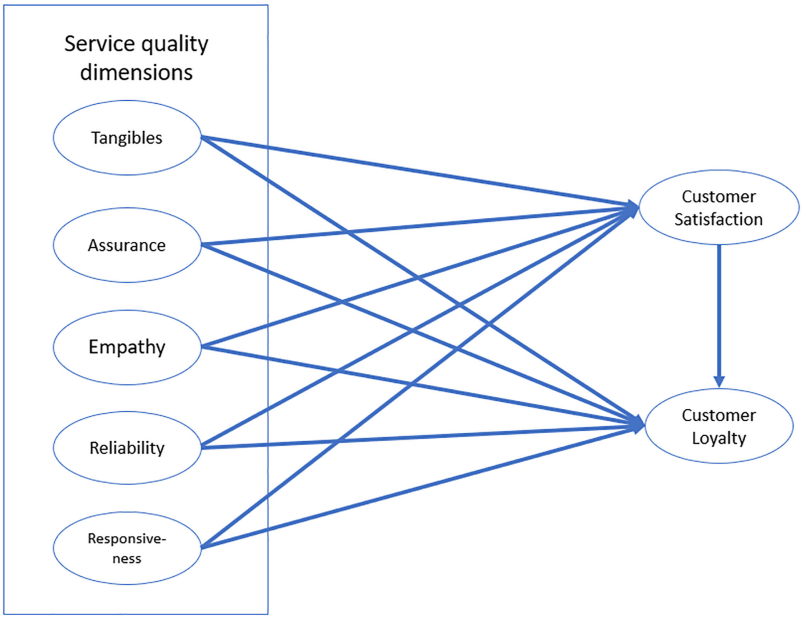


Figure 1.
Theoretical model

H5. Responsiveness, as measured from textual reviews, has positive relationships with (a) customer satisfaction and (b) customer loyalty.

H6. Customer satisfaction mediates the relationships between (a) tangibles, (b) assurance, (c) empathy, (d) reliability, (e) responsiveness and customer loyalty

4. Empirical study

The overall methodology is explained in Figure 2.

4.1 Data

We have extracted the data from airlinequality.com, an airline-review website. The data had 27,052 online reviews on 362 airlines by reviewers of 158 nationalities written in between 2005 and 2015. The dataset had the customer satisfaction score of the reviewers measured on a scale ranging from one to ten and recommendation behavior as a dichotomous variable. It also had textual reviews written in English language and aspect-wise quantitative ratings on the comfortable seating arrangement, behavior of the cabin-staff, services related to food/beverages, entertainment available in the flight, Wi-Fi-connectivity, ground service and value for money. As Wi-Fi and ground service columns had very large numbers of missing data, 83 and 88%, respectively, we did not consider these two variables for our analysis. Such large quality of missing data can be because the reviewers were not obliged to respond for all the quantitative aspects and most must have found Wi-Fi and ground service less important. The rest of the data did not have any missing value.

To generate information from the textual reviews, first, we have done the pre-processing by removal of quantitative figures, unnecessary words, blank-spaces and punctuations. The word cloud of the cleaned corpus can be found in Figure 3. Some of the important and most frequent words in the corpus were "flight", "food", "service", "time", "seat", "staff", "business", "economy", "class", "crew", "cabin", "entertainment" etc. Text mining techniques suggest that highly frequent words have less information as often they are repetitive in nature. However, in case of review data, the most common words often express the aspects based on which

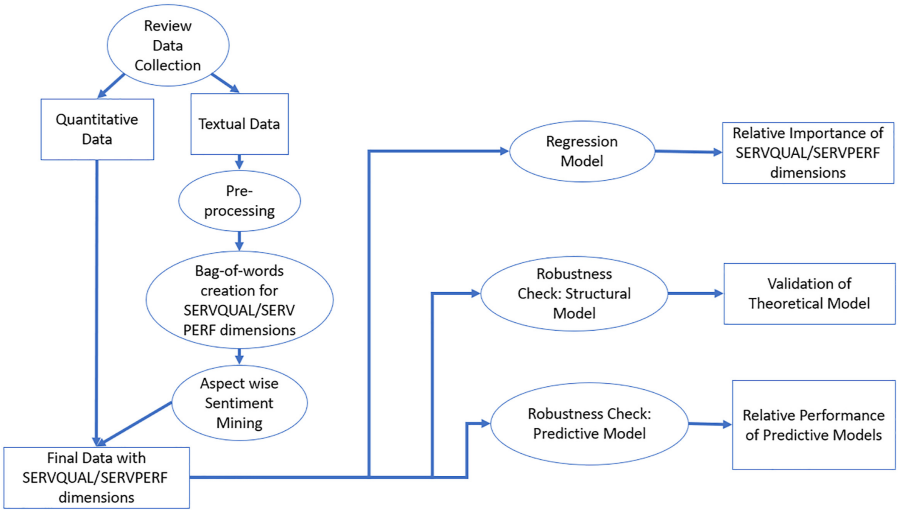


Figure 2.
Flow diagram from methodology

Figure 3.
Word cloud of the review text before SERVQUAL dimension generation

customers analyze the firm (Siering *et al.*, 2018). In this dataset too we find the similar outcome. Therefore, we used frequency-based filtering of important words to find the service aspects, as discussed later.

With this cleaned corpus, we generated the sentiment about the SERVQUAL/SERVPERF dimensions by following a two-step approach, as suggested in the literature (Siering *et al.*, 2018). Initially, authors have prepared the bag-of-words that can illustrate the dimensions following the method proposed by Chatterjee (2019). We considered all words which occurred in at least 5% of all the reviews. Using the R package for Ripple Down Rules-based Part-Of-Speech Tagging (RDRPOS) (Nguyen *et al.*, 2016), we first considered the nouns of the review text. Nouns primarily define the SERVQUAL/SERVPERF aspects on which evaluation is conducted while writing a review. On the other hand, certain adjectives and verbs are also found to be close to SERVQUAL/SERVPERF dimensions. We then divided the words into bag-of-words to explain the SERVQUAL/SERVPERF dimensions. We asked the views of 2 academic experts in the service-management domain and also of 5 online reviewers to validate whether the bag-of-words are explaining the dimensions. This led to the dropping of some work, and the final list is expressed in Table 1.

Table 1.
SERVQUAL/ SERVPERF dimensions and the bag of words used

SEQVQUAL dimensions	Bag of words
Tangibles	Cramped, appeared, reclined, seats, flights, class, airplane, plane, airlines, aircraft, airport, leg and legroom
Assurance	Flying, expected, experienced, food, meal, luggage, baggage, breakfast and snacks
Empathy	Boarding, offered, check, boarded, managed, smiling, enjoyed, considering, refused, treated, handled, presented, entertainment, leg, lounge, water, legroom and comfort
Responsiveness	Delayed, arrived, canceled, waiting, departed, scheduled, informed, time, hours, attendants, system and delay
Reliability	Recommend, landing, landed, recommended, experienced and safe

Once we have created the bag-of-words, we separated the reviews into various sentences and inspected if the words explaining a dimension is occurring in a sentence or not. We found the dimension-wise sentiment by calculating the global sentiment of such sentences, which consists of any one of the words explaining the dimension. Lexicon based methods were used for sentiment mining as it is a well-accepted method in the marketing, operations research, information systems and data science literature (Siering *et al.*, 2018; Mostafa, 2013; Dang *et al.*, 2010; Taboada *et al.*, 2011). Other methods, such as statistical-learning-based sentiment mining, may work better in a specific context; however, its generalizability remains a challenge (Taboada *et al.*, 2011). We adopted the EmoLex or NRC Word-Emotion Association Lexicon (Mohammad and Turney, 2013). NRC furnishes users with the count of positive and negative words in a review. The following equation is used to measure sentiment:

Sentiment = (posf − negf)/(posf + negf) where posf = positive words frequency and negf = negative-words' frequency.

Further, we have checked the reliability of our method by randomly choosing 100 reviews and asking three independent experts to rate the reviews on a 5-point Likert scale assessing if the review expresses positive sentiment about a SERVQUAL/SERVPERF dimension. The correlation of our sentiment scores with the average sentiment score of the 3 experts (tangibles = 0.75, empathy = 0.61, assurance = 0.63, responsiveness = 0.68, reliability = 0.65) suggest that the measures are reliable. We have provided the variable summary in Table 2. Figures 4 and 5 give the histograms of the SERVQUAL dimensions and quantitative ratings.

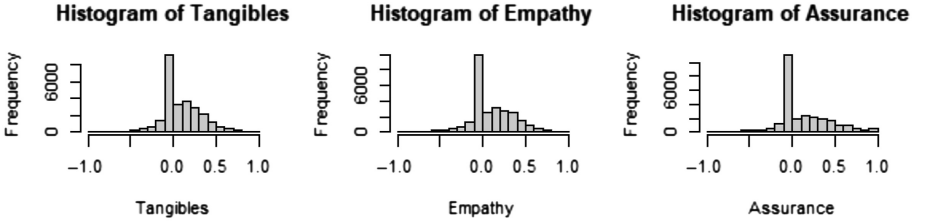
4.2 Explanatory model

We have used an ordered logit regression model to explain the effects of attribute wise ratings and SERVQUAL/SERVPERF dimensions of customer satisfaction. We have chosen ordered logit regression as the customer satisfaction is not normally distributed (refer to Figure 6). Further, we have used logistic regression to explain recommendation behavior. The results of the analysis are given in Model 1 and Model 1.1 in Table 3. We have checked for multicollinearity using the variance inflation factor and took a cut-off of 3; we did not find any multicollinearity issue.

The results suggest that the attribute wise ratings have positive and significant relationships with both customer satisfaction and recommendation behavior. Moreover, we found core attributes to be the major antecedents of recommendation behavior and customer satisfaction as their relative importance to explain the outcome variables are higher than ancillary aspects. For instance, the core aspects such as value for money (CSATβ = 1.14, $p < 0.001$; RBVβ = 1.24, $p < 0.001$), comfort in seating arrangement (CSATβ = 0.57, $p < 0.001$;

	Average	Standard deviation
Overall rating	6.03	3.24
Rating for seat comfort	3.23	1.36
Rating for cabin staff	3.50	1.47
Rating for food and beverages	2.98	1.53
Rating for inflight entertainment	2.50	1.70
Rating for value for money	3.39	1.46
Recommendations	0.60	0.49
Tangibles	0.10	0.24
Assurance	0.14	0.30
Empathy	0.12	0.25
Responsiveness	0.08	0.22
Reliability	0.02	0.12

Table 2.
Basic details of the variables in the model



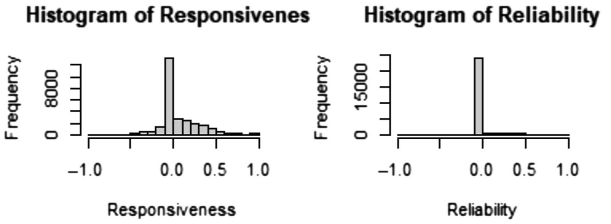
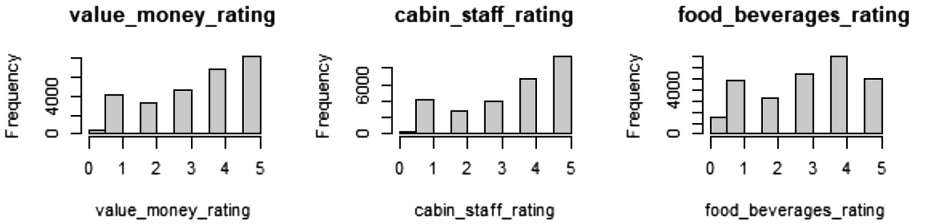


Figure 4.
Histogram of SERVQUAL dimensions



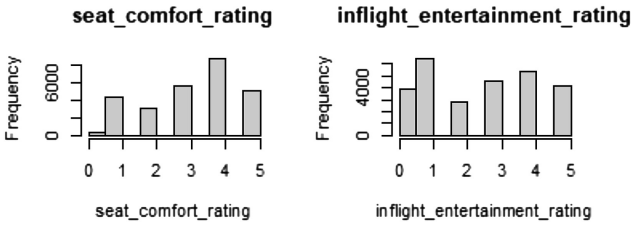


Figure 5.
Histogram of quantitative ratings

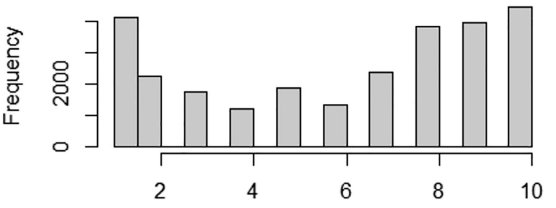


Figure 6.
Histogram of overall rating (dependent variable)

Table 3. Regression model results: Overall satisfaction and recommendation

Models	Model 1	Model 1.1	Model 2	Model 2.1	Model 3	Model 3.1
Dependent variable	Overall satisfaction	Recommendation	Overall satisfaction	Recommendation	Overall satisfaction	Recommendation
Model method	Ordered logistic regression	Binary logistic regression	Ordered logistic regression	Binary logistic regression	Ordered logistic regression	Binary logistic regression
AIC	79,489	11,967	79,379	11,929	52,171	7,061
Sample size	27,052	27,052	27,052	27,052	18,600	18,600
Intercept		−8.57***		−9.7***		−10.36***
Rating for seat comfort	0.57***	0.44***	0.56***	0.42***	0.65***	0.6***
Rating for cabin staff	0.84***	0.78***	0.8***	0.78***	0.92***	0.95***
Rating for food and beverages	0.2***	0.13***	0.38***	0.26***	0.34***	0.26***
Rating for inflight entertainment	0.05***	0.04*	0.11***	0.16***	0.1***	0.12***
Rating for value for money	1.14***	1.24***	1.03***	1.34***	1.07***	1.3***
Tangibles	0.44***	1.21***	0.54***	1.53***	0.35***	0.88***
Responsiveness	0.51***	0.96***	0.37***	1.18***	0.4***	0.89***
Reliability	0.5***	1.02***	0.34	0.67	0.5***	0.87**
Empathy	0.48***	1.11***	0.34***	0.67***	0.49***	0.95***
Assurance	0.26***	1.04***	0.18**	1***	0.16**	0.67***
Econ			0.17	1.4***		
Rating for seat comfort: econ			0.02	0.03		
Rating for cabin staff: econ			0.06**	0		
Rating for food and beverages: econ			−0.22***	−0.16***		
Rating for inflight entertainment: econ			−0.07***	−0.14***		
Rating for value for money: econ			0.14***	0.14***		
Tangibles: econ			−0.14*	−0.42***		
Responsiveness: econ			0.2	−0.26		
Reliability: econ			0.22	0.44		
Empathy: econ			0.12**	0.62***		
Assurance: econ			0.09	0.07		

(*continued*)

Table 3.

Models	Model 1	Model 1.1	Model 2	Model 2.1	Model 3	Model 3.1
Airline low cost					−1.14***	0.83
Rating for seat comfort: airline low cost					−0.12*	−0.27*
Rating for cabin staff: airline low cost					0.2**	0.05
Rating for food and beverages: airline low cost					−0.31***	−0.22*
Rating for inflight entertainment: airline low cost					−0.03	−0.31***
Rating for value for money: airline low cost					0.53***	0.26*
Tangibles: airline low cost					0.08	0.46
Responsiveness: airline low cost					−0.3*	−0.79***
Reliability: airline low cost					0.47	−0.02
Empathy: airline low cost					0.68**	0.79*
Assurance: airline low cost					0.08	1.02
1|2	4.82***		4.98***		5.31***	
2|3	6.16***		6.32***		6.66***	
3|4	7.16***		7.33***		7.76***	
4|5	7.87***		8.05***		8.58***	
5|6	8.95***		9.13***		9.77***	
6|7	9.68***		9.87***		10.58***	
7|8	10.83***		11.01***		11.85***	
8|9	12.32***		12.51***		13.43***	
9|10	13.8***		13.99***		15.02***	

Note(s): *** means $p < 0.001$, ** means $p < 0.01$ and * means $p < 0.05$ econ means "Economy class travelers"

RBVβ = 0.44, $p < 0.001$) and behavior of cabin-staff (CSATβ = 0.84, $p < 0.001$; RBVβ = 0.78, $p < 0.001$) are more strongly related with customer satisfaction as well as recommendation. Ancillary aspects such as food and beverages (CSATβ = 0.2, $p < 0.001$; RBVβ = 0.13, $p < 0.001$) and entertainment inside the flight (CSATβ = 0.05, $p < 0.001$; RBVβ = 0.04, $p < 0.001$) have relatively weaker relationships.

Further, we also found, that while all the dimensions of SERVQUAL/SERVPERF have positive and significant relationship with customer satisfaction and recommendation, their relative importance varies based on the outcome variable. For customer satisfaction, the most important variable is responsiveness ($\beta = 0.51$, $p < 0.001$), followed by reliability ($\beta = 0.5$, $p < 0.001$), empathy ($\beta = 0.48$, $p < 0.001$), tangibles ($\beta = 0.44$, $p < 0.001$) and assurance ($\beta = 0.26$, $p < 0.001$). On the other hand, for recommendation behavior tangibles ($\beta = 1.21$, $p < 0.001$) have been found to be most important, followed by empathy ($\beta = 1.11$, $p < 0.001$), assurance ($\beta = 1.04$, $p < 0.001$), reliability ($\beta = 1.02$, $p < 0.001$) and responsiveness ($\beta = 0.96$, $p < 0.001$). The result supports H1 to H5.

4.3 Effect of travel class

Next, we have included the effect if the travel class in the model. In our dataset, 70.06% reviewers have reviewed their experience in economy class, and the rest were non-economy (premium, business or first-class). We created a dummy variable that took a value of 1 when the travel-class was economy and used that variable for further modeling. We have included the direct effect of the travel class and the interaction effects of travel class and individual service attributes and SERVQUAL/SERVPERF dimensions.

We found travel class has no significant impact on customer satisfaction, but economy class people have higher recommendation behavior ($\beta = 1.4$, $p < 0.001$). Moreover, for economy class travelers, value for money (CSATβ = 0.14, $p < 0.05$; RBVβ = 0.14, $p < 0.001$) and behavior of cabin-staff (CSATβ = 0.06, $p < 0.01$) are more important and food/beverages (CSATβ = −0.22, $p < 0.001$; RBVβ = −0.16, $p < 0.001$) and entertainment inside the flight (CSATβ = −0.07, $p < 0.001$; RBVβ = −0.14, $p < 0.001$) are less important than those traveling in non-economy class.

In terms of the SERVQUAL/SERVPERF dimensions, economy class people give lesser importance to tangibles (CSATβ = −0.14, $p < 0.05$; RBVβ = −0.42, $p < 0.001$) and more importance to empathy (CSATβ = 0.12, $p < 0.01$; RBVβ = 0.62, $p < 0.001$).

4.4 Effect of airline type

We have also included the effect of the airline type (low-cost vs. full-service) in the model. In our dataset, 8,452 had no airline type mentioned, and we did not use these data. In the rest of the data, 1805 reviews were on low-cost airlines and the rest of full-service airlines. We created a dummy variable that took a value of 1 when the airline was low-cost and used that variable for further modeling. We have included the direct effect of airline type and the interaction effects of airline-type with individual service attributes and SERVQUAL/SERVPERF dimensions.

We found airline type has no significant impact on recommendation behavior, but low-cost airline travelers are generally less satisfied ($\beta = -1.14$, $p < 0.001$). Moreover, for travelers in low-cost airlines, value for money (CSATβ = 0.53, $p < 0.001$; RBVβ = 0.26, $p < 0.05$) and behavior of the staff in cabin (CSATβ = 0.2, $p < 0.01$) are more important and comfort in seating arrangement (CSATβ = −0.12, $p < 0.05$; RBVβ = −0.27, $p < 0.05$), food and beverages (CSATβ = −0.31, $p < 0.001$; RBVβ = −0.22, $p < 0.05$) and entertainment inside the flight (RBVβ = −0.31, $p < 0.001$) are less important than those traveling in full-service airlines.

In terms of the SERVQUAL/SERVPERF dimensions, travelers in low-cost airlines gives lesser importance to responsiveness (CSATβ = −0.3, $p < 0.05$; RBVβ = −0.79, $p < 0.001$) and more importance to empathy (CSATβ = 0.68, $p < 0.01$; RBVβ = 0.79, $p < 0.001$).

4.5 Robustness check: structural model

To study the robustness of the theoretical model proposed, we have conducted PLSSEM with the data. In model 4 (Table 4), we did not consider the mediating effect of customer satisfaction between the SERVQUAL/SERVPERF dimensions and customer loyalty (recommendation behavior) and in model 5 (Table 4) we considered the same. We found that the goodness-of-fit of the model improves with the mediating effect being considered. We found seat comfort rating ($\beta = 0.16, p < 0.001$), cabin staff rating ($\beta = 0.29, p < 0.001$), food and beverages rating ($\beta = 0.06, p < 0.001$), inflight entertainment rating ($\beta = 0.01, p < 0.01$), value for money rating ($\beta = 0.42$, $p < 0.001$), tangibles ($\beta = 0.04$, $p < 0.001$), responsiveness ($\beta = 0.03$, $p < 0.001$), reliability ($\beta = 0.02$, $p < 0.001$), empathy ($\beta = 0.05$, $p < 0.001$) and assurance ($\beta = 0.04$, $p < 0.001$) are positively related with customer satisfaction. We also found that tangibles ($\beta = 0.04$, $p < 0.001$), responsiveness ($\beta = 0.03$, $p < 0.001$), reliability ($\beta = 0.01$, $p < 0.001$), empathy ($\beta = 0.05$, $p < 0.001$) and assurance ($\beta = 0.05$, $p < 0.001$) are positively related with customer loyalty. This supports H1 to H5. Customer satisfaction also has a positive relationship with customer loyalty ($\beta = 0.71$, $p < 0.001$). Moreover, in the presence of customer satisfaction, the relationship of SERVQUAL/SERVPERF dimensions with customer loyalty (recommendation behavior) becomes either weaker or insignificant, thus establishing the mediating effect of customer satisfaction in the relationship of SERVQUAL/SERVPERF dimensions and customer loyalty. Therefore, the results support H6(a) to H6(d) partially and H6(e) fully. The PLSSEM results suggest that the theoretical model is robust and can be validated even with the SERVQUAL/SERVPERF dimensions developed from textual data.

4.6 Robustness check: predictive models

Following our theoretical model, we used both machine learning and econometric algorithms for predicting customer satisfaction and recommendation. We used standard evaluation techniques for analyzing the model performance. We used two sets of models both for customer satisfaction and recommendation: the first one only included the quantitative rating and the second one included the SERVQUAL dimensions.

Models	Model 4		Model 5	
Dependent variable	Overall satisfaction	Recommendation	Overall satisfaction	Recommendation
R^2	0.794	0.670	0.794	0.775
Intercept	0.00	0.00	0.00	0.00
Rating for seat comfort	0.16***	0.11***	0.16***	0.00
Rating for cabin staff	0.29***	0.26***	0.29***	0.05***
Rating for food and beverages	0.06***	0.04***	0.06***	0.00
Rating for inflight entertainment	0.01**	0.01***	0.01**	0.01***
Rating for value for money	0.42***	0.42***	0.42***	0.12***
Tangibles	0.04***	0.04***	0.04***	0.02***
Responsiveness	0.03***	0.03***	0.03***	0.01***
Reliability	0.02***	0.01***	0.02***	0.00
Empathy	0.05***	0.05***	0.05***	0.02***
Assurance	0.04***	0.05***	0.04***	0.02***
Overall satisfaction				0.71***

Note(s): *** means $p < 0.001$, ** means $p < 0.01$ and * means $p < 0.05$

Table 4.
Path model results

Along with linear regression for customer satisfaction and binary logistic regression for recommendation behavior, we also used random forest (RF), artificial neural network (ANN) and support vector machine (SVM), as has been done in the extant literature (Chatterjee, 2019; Kalinić *et al.*, 2020; Simsek *et al.*, 2021; Siering *et al.*, 2018; Yeh and Chen, 2020). We used linear regression instead of ordered logistic regression for customer satisfaction following extant literature which suggests linear regression provides better predictive power than non-linear regression in the context of customer reviews (Chatterjee, 2019). ANN mimics the human neural network. ANN has multiple interlinked layers of nodes (input, hidden and output) and neurons connect such nodes with each other. Each node has an activation function, which changes the input to an output. Each connection is also weighted. The weighted sum of the outputs from a neuron of an earlier layer becomes the input for a node in following layers, which then goes through the activation function and passes to the next layer (Han *et al.*, 2011). This is commonly called the feed-forward mechanism. The randomly assigned weights are optimized in an iterative training process to minimize predictive inaccuracy (Nisbet *et al.*, 2009). We used the sigmoid function as the activation function as it is most frequently used (Fuller *et al.*, 2009). We have used a R package called "nnet" (Venables and Ripley, 2002). While we had 24 neurons in a single hidden layer for customer satisfaction, the same for recommendation was 22. We have also chosen decay value = 0.0005 and training stop = 0.0001 for both the models.

The second machine learning algorithm used by us was RF. It is known for large data handling capabilities and overcoming overfitting (Breiman, 2001). This method creates multiple samples from the training data by choosing with repetition a part of the overall training data (Satapathy *et al.*, 2016). Multiple decision trees are developed on such multiple samples. Ensembling of the results are done to get the final result (Satapathy *et al.*, 2016). For the classification problem, the voting mechanism is used from each tree (Sun and Schulz, 2015), while for regression problem, the weighted average is considered. We employed "randomForest" package in R where the number of predictors taken into consideration at each fork of the tree was equal to 8 and 500 random trees was used during modeling.

SVM has often been considered to be one of the best classifiers available (Kotsiantis *et al.*, 2007; Ghosh *et al.*, 2019). SVM searches the highest margin hyperplanes that maximize the distance between different groups (Huang *et al.*, 2002). SVM can also use non-linear Kernels as often linear hyperplane will not be able to differentiate between groups. An R package called "e1071" was used (Meyer and Wien, 2015). Following extant literature, we selected parameters using the grid-search heuristic and have chosen Radial Basis Function kernel (Siering *et al.*, 2018).

We have used 70% of the data as training data and the rest as testing data and 100-fold repetition. We used root mean square error (RMSE) and accuracy score respectively as the measure of model fit for customer satisfaction and recommendation. As the recommendation behavior was enough balanced (60% recommended and 40% did not), accuracy is a suitable measure of model fit. The mean and standard deviation of the model fit scores of the 100 models are listed in Table 5. In both of the cases, SVM performed the best followed by the regression techniques and then ANN and RF. As we have included the SERVQUAL measures, the model fit improved significantly for most of the models. This suggests that the SERVQUAL dimensions, as measured using the text mining method suggested by us, have good predictive power for customer satisfaction and recommendation behavior.

5. Discussions

The current study focuses on the application of text mining and sentiment mining technique to retrieve insights about the SERVQUAL/SERVPERF dimensions from the textual part of

Model fit scores	Before SERVQUAL dimensions				After SERVQUAL dimensions			
	CusSAT		RecoB		CusSAT		RecoB	
	Mean RMSE	Sd RMSE	Mean accuracy	Sd accuracy	Mean RMSE	Sd RMSE	Mean accuracy	Sd accuracy
Linear regression	1.11	0.29			1.02	0.19		
Logistic regression			0.81	0.09			0.83	0.07
ANN	1.35	0.25	0.79	0.12	1.18	0.17	0.82	0.11
RF	1.27	0.31	0.79	0.15	1.12	0.26	0.81	0.13
SVM	1.03	0.35	0.86	0.11	0.86	0.22	0.91	0.11

Note(s): CusSAT = customer satisfaction and RecoB = recommendation behavior

Table 5. Model fit scores from predictive models

the consumer reviews. It also uses such measures of the SERVQUAL/SERVPERF dimensions to find their relative importance for customer satisfaction and recommendation behavior. Along with information from the textual review, quantitative ratings were also included in the models as covariates.

Among the quantitative attribute-wise-rating core attributes such value for money, comfort in seating arrangement and behavior of cabin-staff is found to be more strongly related to customer satisfaction and recommendation than augmented aspects like food/beverages and entertainment inside the flight. These results are in line with extant literature, which suggests that core aspects are the primary factors that lead to service evaluations as they are more accessible and diagnostic (Chatterjee, 2019; Lynch, 2006; Seiring *et al.*, 2018). Moreover, core aspects become furthermore important for low-cost airline travelers and economy class travelers in both (low-cost and full-service) types of airlines, as the accessibility and diagnosticity increases in the service context of economy-class or low-cost-airline (Lynch, 2006).

Among the SERVQUAL/SERVPERF dimensions, we found differential importance for customer satisfaction and recommendation behavior. This can be because satisfaction is an attitudinal construct, while the recommendation is part of loyalty behavior (Kumar *et al.*, 2013). The attitude is formed by the expectancy-value process, where individuals compare the accessible evaluative information with their expectations (Ajzen and Fishbein, 2000). However, the behavior is affected by attitude, norms and control perceptions. Therefore, the way information (SERVQUAL/SERVPERF dimensions) may influence attitude (customer satisfaction) may be different from the way it influences behavior (recommendation). Therefore, the result is in line with the extant literature of attitude-behavior relationships and differences (Ajzen and Fishbein, 2000).

For customer satisfaction, the most important variable is responsiveness, followed by reliability, empathy, tangibles and assurance. The core service airline provides fast travel, and timeliness is the most important factor of such a service, followed by safety. For attitude formation, information about these aspects are most accessible (Ajzen and Fishbein, 2000; Lynch, 2006). Therefore, a higher effect of responsiveness and reliability on customer satisfaction is expected.

On the other hand, for recommendation behavior, tangibles have been found to be most important, followed by empathy, assurance, reliability and responsiveness. Moreover, economy class travelers are less affected by tangibles, and low-cost airline customers are less concerned about responsiveness. Moreover, both of these customers give huge importance to empathy. Along with attitude, the behavior is affected by the norm and control perceptions (Ajzen and Fishbein, 2000). In case of airline reviews, norms may play

a crucial role. While the responsiveness and reliability perception about an airline is a part of personal experience, tangibles are more visible and diagnosable by others. Moreover, empathy is also related to the norm of reciprocation. We reciprocate empathy to someone who shows us empathy. Therefore, if the airline shows empathy to a traveler, the traveler will show empathy towards the airline by recommending it, irrespective of the satisfaction level. Thus, our finding is in relation to the theory of planned behavior (Ajzen and Fishbein, 2000).

We also observed improved predictive outcomes of the econometric and machine learning models if we include SERVQUAL dimensions while predicting customer satisfaction and recommendation behavior. This suggests that the proposed method is an essential addition to the customer satisfaction and recommendation prediction problem.

5.1 Theoretical contributions

Our study has multiple of theoretical and methodological contributions. First, the paper contributes towards the existing literature on service quality, SERVQUAL, customer satisfaction and recommendation (Cenfetelli *et al.*, 2008; Chesbrough and Spohrer, 2006; Delone and Mclean, 2004; Rust and Miu, 2006; Xu *et al.*, 2013; Wang, 2008). Most of the existing studies on SERVQUAL/SERVPERF and customer satisfaction have used quantitative measures or survey-based measures of service quality and have ignored qualitative and textual data (Luo *et al.*, 2012; Palese and Usai, 2018). However, textual data have a considerable contribution to the literature of marketing and information systems (Chatterjee, 2019). This creates a gap in the SERVQUAL/SERVPERF and customer satisfaction literature, which this study intends to fill, thus contributing to the extant literature.

Second, the findings of our study also provide new insights to the stream of literature that focused only on the aggregate measure of service quality (Luo *et al.*, 2012) in the context of online reviews. Our study provides insights about individual SERVQUAL/SERVPERF dimensions and their influence on customer satisfaction and recommendation behavior, which is relatively new in online reviews literature (Xu *et al.*, 2013; Petter *et al.*, 2012). Little study has been done on the inter-connections of service quality dimensions, customer satisfaction and loyalty using textual data from UGC on the Internet (Palese and Usai, 2018); therefore, this study is an important addition to this literature.

Third, our findings also strengthen the attitude-behavior paradigm of satisfaction and loyalty (Kumar *et al.*, 2013). It also strengthens the expectancy-value framework of attitude formation and the theory of planned behavior for behavior-formation (Ajzen and Fishbein, 2000). Thus, the paper contributes to the literature on the theory of planned behavior and expectancy-value framework (Ajzen and Fishbein, 2000).

Fourth, our study also contributes to the literature of service quality, which majorly focuses on quantitative measures (Anderson and Jacobsen, 2000; Calisir *et al.*, 2016; Namukasa, 2013). Quantitative measures are more common in service-quality research because such measures are accessible and diagnosable. We found, among the quantitative attribute wise rating core attributes are found to be more strongly related to customer satisfaction and recommendation than augmented aspects. Moreover, core aspects become furthermore important for low-cost airline travelers and economy class travelers in both (low-cost and full-service) types of airlines. This finding strengthens extant literature of application of AD models, which suggests such information is more persuasive that are accessible and diagnostic (Chatterjee, 2019; Lynch, 2006; Seiring *et al.*, 2018).

Finally, the study provides a better methodology to predict customer satisfaction and recommendation by combining quantitative ratings and qualitative reviews and at the same time exploring the theoretical explanations behind the same. Explainable machine learning

and artificial intelligence have been among the chief topics of analytics research within the last decade (Rai, 2020). The current study intends to contribute to the same.

5.2 Managerial implications

The current work has multiple managerial implications. First, as of now, the service managers did not have any mechanism or methodology to measure SERVQUAL/SERVPERF dimensions using the textual review data. The only way the service managers could measure such dimensions was via survey, which is costly and have many limitations and biases. Although, online reviews can also be fake and biased, true picture of consumer responses can be found once those are eradicated. This paper is the first paper which provides a mechanism to measure SERVQUAL/SERVPERF dimensions from textual reviews.

Second, the study helps the service managers to focus on the qualitative information of the online reviews to generate insights about their service quality and customer satisfaction. This will give them an idea of which SERVQUAL/SERVPERF dimensions they should focus on or which dimension is more important for their customers in terms of driving customer satisfaction. Variations of the relative importance of the SERVQUAL/SERVPERF dimensions depending on the airline type or the traveler class contributes towards the above objective of a better understanding of service quality management and customer satisfaction. Therefore, the methodology suggested herein is essential for service managers.

Third, not only the study tries to help managers understand the relative importance of SERVQUAL/SERVPERF dimensions in different business contexts, as mentioned in point 2, the study also suggests that the relative importance can vary based on business goals too. It helps customer relationship managers to understand that the drivers of the consumer attitude (customer satisfaction) and consumer loyalty behavior (recommendation) are different, and they may have to focus on different sets of SERVQUAL/SERVPERF dimensions to ensure satisfaction and loyalty. Moreover, the quantitative evaluations, which are better accessible and more diagnostic, have been used classically to do the above. Textual reviews have not contributed much to the decision-making process of the customer relationship managers. This method will contribute to the same.

Finally, the study proposes a methodology which can be used as an addition on to any service quality survey. Until now, most service quality surveys done by business research managers are done using close-ended qualitative questions, irrespective of the well-known biases of the close-ended questions. Our methodology can be used in situations beyond online reviews too, where the close-ended service quality surveys are appended with open-ended questions. Text mining and econometric techniques can then be used for insight generation from both open and close-ended survey responses in the context of service quality measurement.

5.3 Limitations and future scope of research

This research article has a few limitations too. First, we could choose only one industry for our analysis-airline industry. Although the above choice does not impact the mechanism of service quality measurement which is the main contribution of the study, the relative importance of various aspects of SERVQUAL/SERVPERF may not be generalizable from the results of this study. Future researchers should replicate the method in other industries to get an idea about the applicability and generalizability of the methods in various industries. Second, online textual reviews also have different formats. Some collect overall reviews; some collect pros and cons separately; some also collected tips given in the customers. However, in this methodology, we do not separate out textual reviews as a response to different types of open-ended questions. Future studies can focus on how to handle these different texts differently. Third, the paper uses the bag-of-words method for feature creation in SERVQUAL/SERVPERF. This makes the process manual and not automated, at least till

the library of the words expressing a SERVQUAL/SERVPERF dimension is created. Future researchers can try to make the process fully automated by adopting advanced topic modeling techniques such as latent Dirichlet allocation, latent semantic analysis etc. Topic modeling can provide us the latent topics in the text which can be further used to find the predictive models to measure the importance of such topics. We have also not considered any filter for bad quality English or grammatically erroneous text. We have also used any filter for fake reviews. However, the above will not impact the key findings of the paper.

References

Ajzen, I. and Fishbein, M. (2000), "Attitudes and the attitude-behavior relation: reasoned and automatic processes", *European Review of Social Psychology*, Vol. 11 No. 1, pp. 1-33.

Anderson, H. and Jacobsen, P.N. (2000), "Creating loyalty: its strategic importance in your customer strategy", *Customer Relationship Management*, Vol. 55, p. 67.

Anderson, C.R. and Zeithaml, C.P. (1984), "Stage of the product life cycle, business strategy, and business performance", *Academy of Management Journal*, Vol. 27 No. 1, pp. 5-24.

Asubonteng, P., McCleary, K.J. and Swan, J.E. (1996), "SERVQUAL revisited: a critical review of service quality", *Journal of Services Marketing*, Vol. 10 No. 6, pp. 62-81.

Ba, S. and Pavlou, P.A. (2002), "Evidence of the effect of trust building technology in electronic markets: price premiums and buyer behavior", *MIS Quarterly*, Vol. 26 No. 3, pp. 243-268.

Bairi, J. and Murali Manohar, B. (2011), "Critical success factors in gaining user customer satisfaction in outsourced IT services", *Journal of Enterprise Information Management*, Vol. 24 No. 6, pp. 475-493, doi: 10.1108/17410391111166530.

Breiman, L. (2001), "Random forests", *Machine Learning*, Vol. 45 No. 1, pp. 5-32.

Calisir, N., Basak, E. and Calisir, F. (2016), "Key drivers of passenger loyalty: a case of Frankfurt–Istanbul flights", *Journal of Air Transport Management*, Vol. 53, pp. 211-a217.

Cenfetelli, R.T., Benbasat, I. and Al-Natour, S. (2008), "Addressing the what and how of online services: positioning supporting-services functionality and service quality for business-to-consumer success", *Information Systems Research*, Vol. 19 No. 2, pp. 161-181.

Chatterjee, S. (2019), "Explaining customer ratings and recommendations by combining qualitative and quantitative user generated contents", *Decision Support Systems*, Vol. 119, pp. 14-22.

Chen, F.Y. and Chang, Y.H. (2005), "Examining airline service quality from a process perspective", *Journal of Air Transport Management*, Vol. 11 No. 2, pp. 79-87.

Chen, P.T. and Hu, H.H.S. (2013), "The mediating role of relational benefit between service quality and customer loyalty in airline industry", *Total Quality Management and Business Excellence*, Vol. 24 Nos 9-10, pp. 1084-1095.

Chesbrough, H. and Spohrer, J. (2006), "A research manifesto for services science", *Communications of the ACM*, Vol. 49 No. 7, pp. 35-40.

Cronin, J.J. Jr and Taylor, S.A. (1992), "Measuring service quality: a reexamination and extension", *Journal of Marketing*, Vol. 56 No. 3, pp. 55-68.

Curry, N. and Gao, Y. (2012), "Low-cost airlines—a new customer relationship? An analysis of service quality, service satisfaction, and customer loyalty in a low-cost setting", *Services Marketing Quarterly*, Vol. 33 No. 2, pp. 104-118.

Dabholkar, P.A. (2006), "Factors influencing consumer choice of a 'rating Web site': an experimental investigation of an online interactive decision aid", *Journal of Marketing Theory and Practice*, Vol. 14 No. 4, pp. 259-273.

Dang, Y., Zhang, Y. and Chen, H. (2010), "A lexicon-enhanced method for sentiment classification: an experiment on online product reviews", *IEEE Intelligent Systems*, Vol. 25 No. 4, pp. 46-53.

Davies, B., Baron, S., Gear, T. and Read, M. (1999), "Measuring and managing service quality", *Marketing Intelligence and Planning*, Vol. 17 No. 1, pp. 33-40.

Delone, W.H. and Mclean, E.R. (2004), "Measuring e-commerce success: applying the DeLone & McLean information systems success model", *International Journal of Electronic Commerce*, Vol. 9 No. 1, pp. 31-47.

Duan, W., Cao, Q., Yu, Y. and Levy, S. (2013), "Mining online user-generated content: using sentiment analysis technique to study hotel service quality", *2013 46th Hawaii International Conference on System Sciences*, IEEE, pp. 3119-3128.

Fuller, D.O., Troyo, A. and Beier, J.C. (2009), "El Nino Southern Oscillation and vegetation dynamics as predictors of dengue fever cases in Costa Rica", *Environmental Research Letters*, Vol. 4 No. 1, 014011.

Gharib, R.K., Garcia-Perez, A., Dibb, S. and Iskoujina, Z. (2020), "Trust and reciprocity effect on electronic word-of-mouth in online review communities", *Journal of Enterprise Information Management*, Vol. 33 No. 1, pp. 120-138, doi: 10.1108/JEIM-03-2019-0079.

Ghose, A. and Ipeirotis, P.G. (2006), "Designing ranking systems for consumer reviews: the impact of review subjectivity on product sales and review quality", *Proceedings of the 16th annual workshop on information technology and systems*, pp. 303-310.

Ghose, A. and Ipeirotis, P.G. (2010), "Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23 No. 10, pp. 1498-1512.

Ghosh, R., Sinha, N. and Biswas, S.K. (2019), "Automated eye blink artefact removal from EEG using support vector machine and autoencoder", *IET Signal Processing*, Vol. 13 No. 2, pp. 141-148.

Gilbert, D. and Wong, R.K. (2003), "Passenger expectations and airline services: a Hong Kong based study", *Tourism Management*, Vol. 24 No. 5, pp. 519-532.

Gourdin, K.N. (1988), "Bringing quality back to commercial air travel", *Transportation Journal*, Vol. 1, pp. 208823-208829.

Grönroos, C. (1982), *Strategic Management and Marketing in the Service Sector*, Swedish School of Economics and Business Administration, Helsinki (accessed 20 May 2013).

Grönroos, C. (1984), "A service quality model and its implications in marketing", *European Journal of Marketing*, Vol. 18 No. 4, pp. 36-44.

Hamouda, M. (2019), "Omni-channel banking integration quality and perceived value as drivers of consumers' satisfaction and loyalty", *Journal of Enterprise Information Management*, Vol. 32 No. 4, pp. 608-625, doi: 10.1108/JEIM-12-2018-0279.

Han, J., Pei, J. and Kamber, M. (2011), *Data Mining: Concepts and Techniques*, Elsevier.

Han, H. and Ryu, K. (2009), "The roles of the physical environment, price perception, and customer satisfaction in determining customer loyalty in the restaurant industry", *Journal of Hospitality and Tourism Research*, Vol. 33 No. 4, pp. 487-510.

He, W., Zhang, W., Tian, X., Tao, R. and Akula, V. (2019), "Identifying customer knowledge on social media through data analytics", *Journal of Enterprise Information Management*, Vol. 32 No. 1, pp. 152-169, doi: 10.1108/JEIM-02-2018-0031.

Hsu, L.C. (2021), "Effect of eWOM review on beauty enterprise: a new interpretation of the attitude contagion theory and information adoption model", *Journal of Enterprise Information Management*, Vol. ahead-of-print No. ahead-of-print.

Hu, N., Pavlou, P.A. and Zhang, J. (2006), "Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication", *Proceedings of the 7th ACM conference on Electronic commerce*, pp. 324-330.

Hu, N., Liu, L. and Zhang, J.J. (2008), "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects", *Information Technology and Management*, Vol. 9 No. 3, pp. 201-214.

Huang, C., Davis, L.S. and Townshend, J.R.G. (2002), "An assessment of support vector machines for land cover classification", *International Journal of Remote Sensing*, Vol. 23 No. 4, pp. 725-749.

Jabr, W. and Zheng, Z. (2014), "Know yourself and know your enemy", *Mis Quarterly*, Vol. 38 No. 3, pp. 635-A10.

Jiang, Z. and Benbasat, I. (2004), "Virtual product experience: effects of visual and functional control of products on perceived diagnosticity and flow in electronic shopping", *Journal of Management Information Systems*, Vol. 21 No. 3, pp. 111-147.

Kalinić, Z., Marinković, V., Djordjevic, A. and Liebana-Cabanillas, F. (2020), "What drives customer satisfaction and word of mouth in mobile commerce services? A UTAUT2-based analytical approach", *Journal of Enterprise Information Management*, Vol. 33 No. 1, pp. 71-94, doi: 10.1108/JEIM-05-2019-0136.

Kohli, R., Devaraj, S. and Mahmood, M.A. (2004), "Understanding determinants of online consumer satisfaction: a decision process perspective", *Journal of Management Information Systems*, Vol. 21 No. 1, pp. 115-136.

Kotsiantis, S.B., Zaharakis, I. and Pintelas, P. (2007), "Supervised machine learning: a review of classification techniques", *Emerging Artificial Intelligence Applications in Computer Engineering*, Vol. 160 No. 1, pp. 3-24.

Kumar, N. and Benbasat, I. (2006), "Research note: the influence of recommendations and consumer reviews on evaluations of websites", *Information Systems Research*, Vol. 17 No. 4, pp. 425-439.

Kumar, V., Dalla Pozza, I. and Ganesh, J. (2013), "Revisiting the satisfaction–loyalty relationship: empirical generalizations and directions for future research", *Journal of Retailing*, Vol. 89 No. 3, pp. 246-262.

Ladhari, R. (2009), "A review of twenty years of SERVQUAL research", *International Journal of Quality and Service Sciences*, Vol. 1 No. 2, pp. 172-198.

Lee, H., Lee, Y. and Yoo, D. (2000), "The determinants of perceived service quality and its relationship with satisfaction", *Journal of Services Marketing*.

Lehtinen, U. and Lehtinen, J.R. (1991), "Two approaches to service quality dimensions", *Service Industries Journal*, Vol. 11 No. 3, pp. 287-303.

Liou, J.J. and Tzeng, G.H. (2007), "A non-additive model for evaluating airline service quality", *Journal of Air Transport Management*, Vol. 13 No. 3, pp. 131-138.

Liou, J.J., Hsu, C.C., Yeh, W.C. and Lin, R.H. (2011), "Using a modified grey relation method for improving airline service quality", *Tourism Management*, Vol. 32 No. 6, pp. 1381-1388.

Luo, J., Ba, S. and Zhang, H. (2012), "The effectiveness of online shopping characteristics and well-designed websites on satisfaction", *MIS Quarterly*, Vol. 36 No. 4, pp. 1131-1144.

Lynch, J.G. Jr (2006), "Accessibility-diagnosticity and the multiple pathway anchoring and adjustment model", *Journal of Consumer Research*, Vol. 33 No. 1, pp. 25-27.

Martín, J.C., Román, C. and Espino, R. (2011), "Evaluating frequent flyer programs from the air passengers' perspective", *Journal of Air Transport Management*, Vol. 17 No. 6, pp. 364-368.

Mason, K.J. (2001), "Marketing low-cost airline services to business travellers", *Journal of Air Transport Management*, Vol. 7 No. 2, pp. 103-109.

McAfee, A., Brynjolfsson, E., Davenport, T.H., Patil, D.J. and Barton, D. (2012), "Big data: the management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 60-68.

Meyer, D. and Wien, F.T. (2015), "Support vector machines", *The Interface to Libsvm in Package e*1071, Vol. 28.

Mohammad, S.M. and Turney, P.D. (2013), "Crowdsourcing a word–emotion association lexicon", *Computational Intelligence*, Vol. 29 No. 3, pp. 436-465.

Mostafa, M.M. (2013), "More than words: social networks' text mining for consumer brand sentiments", *Expert Systems with Applications*, Vol. 40 No. 10, pp. 4241-4251.

Mudambi, S.M. and Schuff, D. (2010), "Research note: what makes a helpful online review? A study of customer reviews on Amazon.com", *MIS Quarterly*, Vol. 34 No. 1, pp. 185-200.

Namukasa, J. (2013), "The influence of airline service quality on passenger satisfaction and loyalty", *The TQM Journal*, Vol. 25 No. 5, pp. 520-532.

Ndubisi, N.O., Wah, C.K. and Ndubisi, G.C. (2007), "Supplier-customer relationship management and customer loyalty: the banking industry perspective", *Journal of Enterprise Information Management*, Vol. 20 No. 2, pp. 222-236, doi: 10.1108/17410390710725797.

Nguyen, D., Rieu, I., Mariani, C. and van Dam, N.M. (2016), "How plants handle multiple stresses: hormonal interactions underlying responses to abiotic stress and insect herbivory", *Plant Molecular Biology*, Vol. 91 No. 6, pp. 727-740.

Nisbet, E.K., Zelenski, J.M. and Murphy, S.A. (2009), "The nature relatedness scale: linking individuals' connection with nature to environmental concern and behavior", *Environment and Behavior*, Vol. 41 No. 5, pp. 715-740.

Oliver, R.L. (1981), "Measurement and evaluation of satisfaction processes in retail settings", *Journal of Retailing*, Vol. 57 No. 3, pp. 25-48.

Ostrowski, P.L., O'Brien, T.V. and Gordon, G.L. (1993), "Service quality and customer loyalty in the commercial airline industry", *Journal of Travel Research*, Vol. 32 No. 2, pp. 16-24.

Palese, B. and Usai, A. (2018), "The relative importance of service quality dimensions in E-commerce experiences", *International Journal of Information Management*, Vol. 40, pp. 132-140.

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. 4, pp. 41-50.

Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: a multiple—item scale for measuring customer perceptions of quality of service", *Journal Retailing*, Vol. 64, pp. 12-40.

Pavlou, P.A. and Dimoka, A. (2006), "The nature and role of feedback text comments in online marketplaces: implications for trust building, price premiums, and seller differentiation", *Information Systems Research*, Vol. 17 No. 4, pp. 392-414.

Pavlou, P.A. and Gefen, D. (2004), "Building effective online marketplaces with institution-based trust", *Information Systems Research*, Vol. 15 No. 1, pp. 37-59.

Petter, S., DeLone, W. and McLean, E.R. (2012), "The past, present, and future of 'IS success'", *Journal of the Association for Information Systems*, Vol. 13 No. 5, p. 2.

Rai, A. (2020), "Explainable AI: from black box to glass box", *Journal of the Academy of Marketing Science*, Vol. 48 No. 1, pp. 137-141.

Rust, R.T. and Miu, C. (2006), "What academic research tells us about service", *Communications of the ACM*, Vol. 49 No. 7, pp. 49-54.

Satapathy, S.M., Acharya, B.P. and Rath, S.K. (2016), "Early stage software effort estimation using random forest technique based on use case points", *IET Software*, Vol. 10 No. 1, pp. 10-17.

Seelhorst, M. and Liu, Y. (2015), "Latent air travel preferences: understanding the role of frequent flyer programs on itinerary choice", *Transportation Research Part A: Policy and Practice*, Vol. 80, pp. 49-61.

Senecal, S. and Nantel, J. (2004), "The influence of online product recommendations on consumers' online choices", *Journal of Retailing*, Vol. 80 No. 2, pp. 159-169.

Siering, M., Deokar, A.V. and Janze, C. (2018), "Disentangling consumer recommendations: explaining and predicting airline recommendations based on online reviews", *Decision Support Systems*, Vol. 107, pp. 52-63.

Simsek, S., Albizri, A., Johnson, M., Custis, T. and Weikert, S. (2021), "Predictive data analytics for contract renewals: a decision support tool for managerial decision-making", *Journal of Enterprise Information Management*, Vol. 34 No. 2, pp. 718-732, doi: 10.1108/JEIM-12-2019-0375.

Sun, L. and Schulz, K. (2015), "The improvement of land cover classification by thermal remote sensing", *Remote Sensing*, Vol. 7 No. 7, pp. 8368-8390.

Szwarc, P. (2005), *Researching Customer Satisfaction and Loyalty: How to Find Out what People Really Think*, Kogan Page Publishers, London.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. (2011), "Lexicon-based methods for sentiment analysis", *Computational Linguistics*, Vol. 37 No. 2, pp. 267-307.

Tian, X., He, W., Tang, C., Li, L., Xu, H. and Selover, D. (2020), "A new approach of social media analytics to predict service quality: evidence from the airline industry", *Journal of Enterprise Information Management*, Vol. 33 No. 1, pp. 51-70, doi: 10.1108/JEIM-03-2019-0086.

Tse, D.K. and Wilton, P.C. (1988), "Models of consumer satisfaction formation: an extension", *Journal of Marketing Research*, Vol. 25 No. 2, pp. 204-212.

Venables, W.N. and Ripley, B.D. (2002), "Exploratory multivariate analysis", in *Modern Applied Statistics with S*, Springer, New York, NY, pp. 301-330.

Wang, Y.S. (2008), "Assessing e-commerce systems success: a respecification and validation of the DeLone and McLean model of IS success", *Information Systems Journal*, Vol. 18 No. 5, pp. 529-557.

Weigend, A. (2009), "The social data revolution (s)", *Harvard Business Review*, available at: https://hbr.org/2009/05/the-social-data-revolution (accessed 28 June 2015).

Xu, J., Benbasat, I. and Cenfetelli, R.T. (2013), "Integrating service quality with system and information quality: an empirical test in the e-service context", *MIS Quarterly*, Vol. 37 No. 3, pp. 777-794.

Yeh, J.-Y. and Chen, C.-H. (2020), "A machine learning approach to predict the success of crowdfunding fintech project", *Journal of Enterprise Information Management*, Vol. ahead-of-print No. ahead-of-print.

Zeithaml, V.A., Parasuraman, A., Berry, L.L. and Berry, L.L. (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*, Simon & Schuster, Free Press.

Zhu, F. and Zhang, X. (2010), "Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics", *Journal of Marketing*, Vol. 74 No. 2, pp. 133-148.

Further reading

Berry, L.L. (1988), "SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.

Jiang, H. and Zhang, Y. (2016), "An investigation of service quality, customer satisfaction and loyalty in China's airline market", *Journal of Air Transport Management*, Vol. 57, pp. 80-88.

Piccoli, G. (2016), "Triggered essential reviewing: the effect of technology affordances on service experience evaluations", *European Journal of Information Systems*, Vol. 25 No. 6, pp. 477-492.

Corresponding author

Shivam Gupta can be contacted at: shivam.gupta@neoma-bs.fr