TOTAL QUALITY MANAGEMENT, VOL. 13, NO. 2, 2002, 241-254

Routledge
Taylor & Francis Group

An empirical assessment and application of SERVQUAL in a Mainland Chinese department store

XIANDE ZHAO,[1] CHANGHONG BAI[2] & Y. V. HUI[3]
[1]*Department of Decision Sciences and Managerial Economics, Chinese University of Hong Kong, Shatin, NT Hong Kong,* [2]*School of International Business, Nankai University, Tianjin, The People's Republic of China &* [3]*Department of Management Sciences, City University of Hong Kong, Kowloon, Hong Kong*

ABSTRACT *More and more retailers in Mainland China are trying to di erentiate their services through quality improvement. In this paper, we report upon the application of SERVQUAL in service quality improvement at a department store in Tianjin, a city in Northern China. We shall also assess the reliability and validity of the SERVQUAL instrument using data collected from this department store. The improvement e ort included comparisons of service quality as perceived by both external customers and internal employees, identification of areas for improvements through focus group discussion of the survey results, and development of a plan for improvements. Statistical analyses of the survey response from this company were also performed to test the validity of the SERVQUAL instrument and the applicability of the five dimensions of service quality in the Chinese retail industry. The results indicate that the gap scores did not merge into five dimensions of service quality; rather, the perception scores roughly merged into six dimensions. The findings from this study indicate that the SERVQUAL instrument and the five dimensions of service quality may not be applicable to the retail sector of Mainland China and further research is necessary to understand service quality in Mainland China.*

Introduction

The retail sector in Mainland China has experienced a period of rapid growth. The relaxation in government policy has led to the establishment of a large number of retail outlets. For example, there were roughly a total of 100 department stores with a business area of more than 10 000 m^2 throughout the entire country in 1994. By the end of 1997, however, there were over 70 such department stores in Beijing alone (Li, 1998).

The rapid increase in the number of retail outlets has inevitably enhanced competition among retailers. Price wars among large department stores have been reported frequently (Li, 1998) and have seriously reduced retailer profitability (Jing, 1999). Adding to the fire was the government's policy to allow foreign investors to enter the retail industry. This

Correspondence: Xiande Zhao, Department of Decision Sciences and Managerial Economics, Chinese University of Hong Kong, Shatin, NT Hong Kong. Tel: 852 2609 7640; Fax: 852 2603 6840; E-mail: xiande@baf.msmail.cuhk.edu.hk

ISSN 0954-4127 print/ISSN 1360-0613 online/02/020241-14
DOI: 10.1080/09544120120102478

brought an influx of foreign retailers, such as Carrefour, Parkson, Price Smart and Istain, causing competition to turn white-hot. These retailers generally have much more experience of the retailing business and place a heavier emphasis on quality and productivity improvement than local retailers. The entry of foreign retailers to the China market has helped raise quality expectations among Chinese customers.

In addition, productivity improvements and an increase in production capacity in recent years have led to a dramatic increase in the supply of different types of products. The previous situation of demand exceeding supply in the centrally planned economy has been completely turned around. The favorable position that China's retailers enjoyed in the centrally planned economy no longer exists. On the contrary, the retail markets are increasingly favorable to customers because of the growth in competition among retailers. Consumers are now enjoying the benefits of having a variety of good quality products available to them at a much lower price.

The pressure of competition in the local retail industry has forced retailers to look for ways to enhance their competitive position. Many have decided to improve service quality in order to differentiate their services from those of the their competitors, rather than initiate a price war. However, the majority of retailers in China are not that familiar with concepts and tools related to service quality improvement and have little experience in making systematic improvements in service quality. Tianjin Leadar department store, the subject of this study, was established in 1986. It was the first and only high-end department store in Tianjin, one of the four cities under the direct administrative guidance of the Chinese Central Government. It was quite profitable during its first few years of operation. In recent years, however, the department store has lost its leadership position in top-end retailing and started losing money due to the increase in competition.

In this paper, we report upon the service quality improvement experience of the Leadar department store in Tianjin. With the help of consultants from the Institute of Modern Management of Nankai University, Leadar embarked on a systematic process of service quality improvement. This process involved the evaluation of service quality as perceived by its customers using SERVQUAL, a comparison of the opinions of customers and staff, the identification of areas for improvement through focus group discussion of the survey results and the development of a plan for improvements. The experience of this company is used to demonstrate how SERVQUAL can be used as a tool for a service company to implement systematic improvements in service quality. Furthermore, statistical analyses were performed to test the dimensionality of service quality and to examine the reliability of SERVQUAL in China. In the following sections, we first summarize related studies, then describe the data collection and analysis procedures that we used. Subsequently, we present the results of our statistical analyses and discuss the managerial implications of the results. Finally, we conclude the paper by summarizing the major contributions of the paper and suggest directions for future research.

Related studies

Since services are intangible, heterogeneous and inseparable, it is difficult to measure service quality objectively. Over the years, many researchers have proposed and evaluated alternative service quality models and instruments for measuring service quality. Among these models, SERVQUAL (Parasuraman *et al.*, 1985) is the most prominent and the most widely used. The authors of this model proposed that the consumer's opinion of quality be formed by an internal comparison of performance with expectations. Good service quality means that the customers' perceptions of service performance meet or exceed their expectations of what the service firm should provide. Through in-depth interviews and focus group discussions with

firms in four different service industries, Parasuraman *et al.* (1985) identified 10 determinants of service quality, which included access, communication, competence, courtesy, credibility, reliability, responsiveness, security, tangibles and customer knowledge. Subsequently, Parasuraman *et al.* (1988) reduced these 10 determinants to five using factor analyses. This led to the development of a 22-item SERVQUAL scale for measuring service quality. The five dimensions include tangibles, reliability, responsiveness, assurance and empathy.

The SERVQUAL scale has been widely used to measure service quality in different service contexts, such as professional services (Freeman & Dart, 1993), health care (Lam, 1997), tourism (Tribe & Snaith, 1998), business school (Pariseau & McDaniel, 1997) and information systems (Kettinger & Lee, 1994). It has also been widely tested for its validity and reliability (Babakus & Boller, 1992; Bolton & Drew, 1991; Cronin & Taylor, 1992, 1994). In spite of the fact that some of these studies failed to support the five dimensional factor structures, Parasuraman *et al.* (1993) defended the five-factor structure of service quality on conceptual and practical grounds.

Although the SERVQUAL instrument has been applied in the study of service quality for many different types of service, it has been the subject of a number of criticisms. For example, Reeves and Bednar (1994) considered the strengths and weaknesses of SERVQUAL and related instruments. The issue of how best to conceptualize and operationalize service quality is still a subject of heated debate (Cronin & Taylor, 1994; Parasuraman *et al.*, 1994; Teas, 1994). However, it is generally agreed that service quality is a multi-dimensional or multi-attribute construct (Cronin & Taylor, 1992; Parasuraman *et al.*, 1988).

The generalizability of SERVQUAL in different service industries has also been questioned. Babakus and Boller (1992) used the SERVQUAL scale to measure service quality in an electric and gas utility company. They found that the proposed five-factor structure of SERVQUAL is problematic and doubted the suitability of the SERVQUAL scale for measuring quality across a wide range of services. The applicability of SERVQUAL across different cultures is also an issue. Because SERVQUAL was developed in a Western environment and due to cultural differences, it is likely that cultural factors will influence its applicability. Donthu and Yoo (1998) studied the effect of the cultural orientation of consumers on their service quality expectations. Based on Hofstede's dimensions of culture, they hypothesized and tested the influence of culture on consumer service quality expectations and found that, as a result of cultural orientation, consumers varied in both their overall expectations with regard to service quality and their expectations of each of the service quality dimensions.

Mattila (1999) also examined the impact of culture, but on customer evaluations of complex services. She evaluated the trade-offs that Western and Asian customers were willing to make between personalized service and pleasant physical environment in the context of luxury hotels. She found that customers with a Western cultural background might be more likely to rely on tangible cues than their Asian counterparts, and that the hedonic dimension of the consumption experience might be more important for Western consumers than for Asians.

Based on the above review of the literature, we found that although the SERVQUAL scale is the most popular instrument for measuring service quality and has been used widely to measure service quality, its reliability and applicability in different cultural and industrial environments still require investigation. In this paper, we investigate the reliability of the instrument and the dimensions of service quality in a Chinese retailing service setting.

Data collection

To help improve the service quality at Leadar department store, we conducted a service quality survey among customers and employees to find the differences in opinions between these two groups. To do this we used the 22-item survey questionnaire as used by Parasuraman

Table 1. *Demographic characteristics of customer respondents (total number = 273, in percentages)*

Sex		Age (years)		Family income per month (RMB)		Occupation	
Male	28.6	Below 20	18.3	Below 1000	23.8	Workers	15.8
Female	65.6	21-35	47.6	1000-1500	31.5	Teachers	7.3
Missing	5.9	36-45	21.2	1500-2000	16.8	Officials	11.6
		46 and above	8.4	2000-3000	8.4	Businessmen	2.9
		Missing	4.4	3000 and above	6.6	Engineers	10.0
				Missing	12.8	Managers	13.9
						Housewives	1.5
						Retirees	1.5
						Students	20.9
						Others	10.6
						Missing	5.1

Table 2. *Characteristics of employee respondents (total number = 201, in percentages)*

Length of service (years)		Position	
Below 1	5.5	Managers	21.4
1-3	6.0	Salespeople	62.2
4-7	11.9	Back-office staff	9.5
8 and above	71.1	Missing	7.0
Missing	5.5		

et al. (1988). Respondents were asked to indicate their expectations and perceptions for each of the 22 items in the questionnaire using a seven-point Likert scale, with '1' indicating 'strongly disagree' and '7' indicating 'strongly agree' for each of the 22 statements. The questionnaire was first translated into Chinese by the author from Nankai University and the Chinese version of the questionnaire was then translated back into English by the authors at the Chinese University of Hong Kong to ensure the accuracy of the translation. The questionnaire was first pilot-tested among 25 respondents. Based on the results of the pilot test, some questions were reworded to avoid confusion.

The questionnaires were distributed to customers in the stores by trained student helpers. The respondents were asked to fill out the questionnaire in the store. Student helpers were there to clarify the questions and collect the completed questionnaires. For employees, questionnaires were distributed through their managers. The author from Nankai University answered any questions about the questionnaire. Out of a total of 300 questionnaires distributed, 273 were completed and collected from customers who had experienced the store's services. A total of 201 questionnaires were distributed among and collected from the store's employees. Tables 1 and 2 show some of the key demographic characteristics of the store's customers and employees, respectively.

Data analysis and results

Reliability assessment

To test the reliability of the SERVQUAL scale and the internal consistency of the five factors as suggested by Parasuraman *et al.* (1988), we computed Cronbach's coefficient α for each

Table 3. *Reliability coefficients (alpha) of SERVQUAL scale for customers (n = 235)*

A *priori* dimension	Perception (*P*)	Expectation (*E*)	Gap (*P*-*E*)
Tangible (T1, T2, T3, T4)[a]	0.7723	0.7280	0.6879
Reliability (RL1, RL2, RL3, RL4, RL5)[a]	0.7785	0.8478	0.6491
Responsiveness (RS1, RS2, RS3, RS4)[a]	0.7656	0.7026	0.5090
Assurance (A1, A2, A3, A4)[a]	0.5941	0.6619	0.5050
Empathy (E1, E2, E3, E4, E4)[a]	0.8272	0.7927	0.2644
Overall (22 items)	0.8370	0.8416	0.6934

[a]The symbols in parentheses indicate the questions included in the dimension. These are the questions contained in the original SERVQUAL instrument as used by Parasuraman *et al.* (1988). For example, T1, T2, T3 and T4 represent the first, second, third and fourth questions related to the tangibles dimension in the SERVQUAL instrument. The simplified questions are shown in Table 18. Please note that the questions on empathy and responsiveness are negatively worded. The questionnaire used is available from the authors upon request.

Table 4. *Reliability coefficients (alpha) of SERVQUAL scale for employees (n = 201)*

A *priori* dimension	Perception (*P*)	Expectation (*E*)	Gap (*P*-*E*)
Tangible (T1, T2, T3, T4)[a]	0.8363	0.6746	0.7391
Reliability (RL1, RL2, RL3, RL4, RL5)[a]	0.8253	0.7861	0.8001
Responsiveness (RS1, RS2, RS3, RS4)[a]	0.7678	0.7388	0.5620
Assurance (A1, A2, A3, A4)[a]	0.6855	0.8529	0.6903
Empathy (E1, E2, E3, E4, E4)[a]	0.8500	0.8529	0.6903
Overall (22 items)	0.8560	0.8655	0.8551

[a]As Table 3.

of the five *a priori* dimensions using data on perceptions, expectations and the differences between the perceptions and expectations. The reliability coefficients for the store's customers and its employees are shown in Tables 3 and 4, respectively. The results in Table 3 show that the internal consistencies of the perception and expectation scales are all quite high. Four out of the five dimensions have internal consistency measures higher than 0.70. The only dimension that has an α below 0.70 was assurance. Misunderstanding of the questions related to this dimension might have caused this. The overall reliability of the perception (0.837) and expectations (0.8416) scales were also quite high. The reliability coefficients for the gap scores (*P*-*E*) were much lower. All dimensions have an α coefficient of below 0.70. Three out of the five dimensions (responsiveness, assurance, empathy) measured reliability coefficients of below 0.60, which is the minimum acceptable value, even for exploratory research. Therefore, the internal consistency of the gap scores is quite low.

When the reliability coefficients of the data from the employees were examined, we saw that the reliability of the perceptions and the gap scores were consistently higher than among the customers for all five dimensions. This may indicate that the employees had a better understanding of the questions, which resulted in the higher consistency of the answers. However, the reliabilities of the expectation scores for tangibles, reliability and assurance factors among the employees were no better than the corresponding reliability of those for the customers.

Overall, the results in Tables 3 and 4 do indicate that there are significant reliability problems in using the gap model to measure service quality. The problems were greater with regard to the reliability of the results for the dimensions that were measured using negative wording (empathy and responsiveness). The reliability of the assurance dimension was also low.

Factor analyses

To test the validity of the five-factor structure in service quality within the Chinese retail industry, we performed both exploratory and confirmatory factor analyses on the perception, expectation and the gap scores for both the customers and the employees. All exploratory factor analyses were initially performed using the principal axis factoring method and Oblimin rotation with the Kaiser Normalization. We used the same methods as in Parasuraman *et al.* (1988) with the hope of obtaining the same factor structure. Other rotational methods were also used to improve the factor loading. The results of the factor analyses for the gap scores, the perception scores and the expectation scores are shown in Tables 5, 6 and 7, respectively.

The exploratory factor analysis gap score results in Table 5 indicate that the 22 items do not match the five-factor structure as described by Parasuraman *et al.* (1998). For example, the third question on assurance (A3) does not load on the same factor as the other three questions on assurance. Some of the items on empathy match the items on responsiveness. The negative wording in the items on empathy and responsiveness may have caused them to merge as one factor. Babakus and Boller (1992) also found that the items worded with negatives tended to merge to the extent that they formed a single factor. The factor loading for the employees shows that the five dimensions are even more problematic. Items in different dimensions have become mixed and many items have a high loading for a number of factors. Table 5 clearly indicates that the gaps between perceptions and expectations do not support the five dimensions of service quality as suggested by Parasuraman *et al.* (1998). Other rotation methods also failed to improve the factors' loading and factor structure.

We also performed confirmatory factor analysis on the gap score of the store's customers to see whether these scores conformed to the five-factor structure. Bentler's comparative fit index is 0.8441 and Bentler and Bonnett's non-normed index measured 0.819. These figures show that the gap scores do not conform to the five-factor structure well.

Table 6 shows the factor analyses results of customer perceptions. The factor analyses results in Table 6 show the customer perception scores basically to support the five dimensions of service quality as proposed by Parasuraman *et al.* (1998). The only item not loaded properly was the fourth item under 'assurance' (A4). This item asked about the degree of agreement with the statement that Leadar department store provides sufficient support for employees to make improvements. This item has a very low loading (< 0.30) for all five factors. Another item that showed a problem with loading on one factor was the fourth item under 'responsiveness' (RS4), which had an almost equal loading for factors 1 and 4.

To improve the factor loading, we tried different rotation methods. Using Equamax rotation with Kaiser Normalization and an eigenvalue cut-off point of 1.0, we obtained the exploratory factor analysis results as shown in the right-hand side column of Table 6. From this table, we can see that the perception scores are distributed among six factors. The assurance factor is divided into two factors (A1 and A2). A1 represents the assurance dimension related to customer perception as to whether the firm is trustworthy and whether customers feel secure doing business with the store. A2 is concerned with whether the employees are polite and whether they get adequate support from the store to do their jobs.

Confirmatory factor analyses of the perception scores for both the original five-factor structure, as well as the six-factor structure discovered through this study, were also

Table 5. *Results of factor analysis of GAP scores (extraction method: principal axis factoring; rotation method: Oblimin with Kaiser normalization)*

Groups	Customers						Employees				
Items/factors	1	2	3	4	5	Items/factors	1	2	3	4	5
T3	0.671			−0.302	−0.329	Rl1	0.867				0.457
T1	0.537			−0.414		Rl3	0.760				0.400
T4	0.520			−0.436		T4	0.427				0.447
T2	0.511	0.322				Rl2	0.719				0.303
A3	0.485	0.426	0.362	−0.397	−0.338	A2	0.696				
Rs2		0.507				A1	0.664				
E1		0.455				A4	0.616				
E5	0.320	0.436			−0.409	A3	0.603				
Rs2		0.435				Rl4	0.578				
E2		0.385				Rl5	0.462				
Rs3		0.352	0.344			E4		0.711		−0.313	
Rs4		0.319				E2		0.707			
A1	0.359		0.626	−0.353		E3		0.663		−0.383	
A2			0.590			E1		0.442	0.356		
A4			0.331			E5		0.378			
Rl2				−0.579		Rs4			0.766		
Rl4	0.421		0.409	−0.568	−0.321	Rs1		0.309		−0.879	
Rl3				−0.521		Rs2		0.568		−0.662	
Rl1			0.309	−0.483		Rs3		0.414		−0.471	
Rl5				−0.393		T2	0.535				0.760
E3		0.460		−0.333	−0.728	T3	0.610				0.681
E4					0.548	T1	0.311				0.663
Eigenvalues	3.74	2.07	1.58	1.46	1.29	Eigenvalues	6.58	3.02	1.44	1.18	1.07
Variance extracted %	17.01	9.41	7.17	6.65	5.87	Variance extracted %	29.90	13.71	6.55	5.37	4.86

Note: Factor loadings below 0.30 are not shown in the table.

Table 6. *Results of factor analysis of perception scores by customer (extraction method: principal axis factoring)*

Rotation	Oblimin with Kaiser normalization					Equamax with Kaiser normalization						
Items/factors	1	2	3	4	5	Items/factors	1	2	3	4	5	6
E3	0.744					E3	0.732					
E2	0.724					E2	0.696					
E4	0.705			0.327		E4	0.690	0.346				
E1	0.589			0.363		E1	0.591	0.442				
E5	0.478					E5	0.515					
T3		0.651				RS2		0.810				
T1		0.631				RS3	0.391	0.681				
T4		0.580	0.333			RS1		0.534				
T2		0.565				RS4	0.354	0.420				
A4						T1			0.716			
Rl4		0.365	0.683			T3			0.577			
Rl3			0.625			T2			0.568			
Rl2			0.508			T4			0.565			
Rl1			0.504			RL4			0.353	0.610		
Rl5			0.471			RL1				0.598	0.324	
Rs2				0.825		RL3			0.400	0.479		
Rs3	0.457			0.599		RL5				0.461		
Rs1				0.557		RL2				0.447		
Rs4	0.372			0.382		A1					0.829	
A1					0.731	A2					0.496	
A2					0.682	A3						0.712
A3		0.345			0.428	A4						0.353
Eigenvalues	5.28	3.97	1.45	1.08	1.02	Eigenvalues	5.42	3.85	1.44	1.15	1.06	1.03
Variance extracted %	23.99	18.06	6.61	4.93	4.62	Variance extracted %	24.64	17.49	6.56	5.21	4.81	4.68

Note: Factor loadings below 0.30 are not shown in the table.

Table 7. *Results of factor analysis of expectations (extraction method: principal axis factoring; rotation method: Oblimin with Kaiser normalization)*

Groups	Customers					Employees				
Items/factors	1	2	3	4	5	Items/factors	1	2	3	4	5
RL2	0.792					E3	0.824				
RL3	0.697		0.324			E2	0.799			0.304	
RL1	0.627		0.318			E4	0.772			0.411	
RL5	0.584		0.337			E1	0.734			0.345	
E4		0.792				RS3	0.657			0.533	
E3		0.721				RS4	0.565				
E2		0.574		0.308		E5	0.537				
E1		0.519		0.391		T2		0.781			
E5		0.440				T1		0.688			0.329
T3			0.721			T3		0.629			0.510
T2			0.635			T4		0.450			0.372
RL4	0.402		0.558			A2			0.824		0.324
T4	0.309		0.485			A1			0.665		
T1			0.462			RS1	0.447			0.809	
RS2				0.784		RS2	0.594			0.720	
RS1				0.650		RL4	0.316	0.422			0.749
RS3		0.411		0.472		RL3		0.370			0.721
RS4				0.310		RL1		0.576			0.681
A2					0.631	RL2	0.372	0.345			0.544
A4					0.612	RL5					0.534
A1					0.513	A3			0.353		0.444
A3					0.475	A4					0.388
Eigenvalues	5.75	3.27	1.68	1.13	1.05	Eigenvalues	6.45	2.92	1.53	1.28	1.14
Variance extracted %	26.15	14.84	7.64	5.12	4.77	Variance extracted %	29.31	13.25	6.93	5.82	5.72

Note: Factor loadings below 0.30 are not shown in the table.

performed. For the five-factor structure, Bentler's comparative fit index and Bentler and Bonnet's non-normed index measured 0.9185 and 0.9054, respectively. For the six-factor structure, Bentler's comparative fit index and Bentler and Bonnet's non-normed index measured 0.9296 and 0.9162, respectively. These results indicate that the perception scores fit the six-factor structure better than they fit the five-factor structure. These results also indicate that perception scores conform to the five factor-structure much better than do the gap scores.

We also conducted factor analyses of the perception scores by the store's employees. Many items have a mixed loading on multiple factors. Alternative rotation methods failed to improve the loading. We can therefore conclude that the employees' perception scores failed to match the five established factors of service quality.

Table 7 shows the results of exploratory factor analysis of expectation scores for the store's customers. The expectation scores conform to the five-factor structure better than gap scores do. The only item that did not match was RL4, for which the loading was highest on 'tangibles'. However, its reliability factor loading was also fairly high. The factor loading of expectation scores for the store's employees is also shown in Table 7. Results in Table 7 indicate that many items were not clearly loaded to any factor and all items that were worded with negatives merged into one factor. Apparently, this result does not support the five-dimension structure of service quality.

Alternative rotational methods were tried to improve factor loading for the expectation scores for both the store's customers and for its employees; but this failed to improve the factor loading. Confirmatory factor analysis of the expectation scores for customers shows that Bentler's comparative fit index measured 0.9043 and Bentler and Bonnet's non-normed index measured 0.8889. These indices show that expectation scores conform to the five-factor structure better than the gap scores do, but do not conform as well as the perception scores.

Overall, the results of our factor analysis presented in Tables 5–7 show that while the gaps between perceptions and expectations do not match the five factors of service quality, the perception scores did match six factors. These results indicate potential problems in using the gap model to measure service quality. We also found that employee opinions did not support the structure of the five dimensions of service quality either.

GAP analysis and identification of areas for improvement

Since factor analysis did not support the five-factor structure of service quality, we have analyzed the gap scores for individual items in the questionnaire and for each dimension using the simple averages of the scores for all items that belong to that dimension. The results are shown in Table 8.

From Table 8, we see that the highest average gap between customer perceptions and expectations exists in the reliability dimension. Among the five items in the reliability dimension, customer responses indicated that the greatest gap existed in the area of being sympathetic to customers and showing a sincere interest in helping them to solve problems. The next greatest gap existed in the area of doing something by the time promised. The third largest gap was indicated to be in the area of being dependable. To reduce the gap in reliability, the company therefore needs to make improvements in these areas. However, employees have a slightly different perspective. They indicated that the greatest gap in their view was in the area of 'keeping accurate records', followed by 'dependability' and 'doing things within a certain time as promised'.

These differences in opinion indicate that it is advisable for a company to conduct

Table 8. *Gaps between perceptions and expectations (P-E) for customers (n = 235) and employees (n = 200)*

Dimensions and items	Customers	Employees
Tangibles (average of T1, T2, T3 and T4)	*−0.34*	*−0.75*
T1. Has up-to-date equipment	−0.37	−0.61
T2. Physical facilities are visually appealing	−0.25	−0.59
T3. Employees are well dressed and neat in appearance	−0.42	−0.94
T4. Appearance of physical facilities is in keeping with the type of services provided	−0.18	−0.84
Reliability (average of RL1, RL2, RL3 and RL4)	*−0.43*	*−0.59*
RL1. When employee promises to do something, he/she does so	−0.55	−0.65
RL2. Be sympathetic to customers and show a sincere interest in resolving customers' problems	−0.59	−0.35
RL3. Be dependable	−0.31	−0.57
RL4. Provide services at the time promised	−0.22	−0.49
RL5. Keep records accurately	−0.21	−0.79
Responsiveness (average of RS1, RS2, RS3 and RS4)	*−0.08*	*−0.18*
RS1. Not be expected to tell you when services will be performed	−0.10	−0.36
RS2. Not realistic for you to expect prompt services	−0.23	−0.12
RS3. Not always willing to help you	−0.04	−0.14
RS4. OK if too busy to respond to your requests promptly	0.26	−0.17
Assurance (average of A1, A2, A3 and A4)	*−0.26*	*−0.64*
A1. Employees can be trusted	−0.05	−0.30
A2. Feel safe in your transaction with this store	0.14	−0.31
A3. Employees are polite	−0.49	−0.77
A4. Employees get adequate support from the firm to do their jobs well	−0.52	−1.15
Empathy (average of E1, E2, E3 and E4)	*−0.05*	*−0.12*
E1. Cannot be expected to give individual attention	−0.15	0.16
E2. Employees cannot be expected to give personal attention	0.12	−0.16
E3. Not realistic to expect employees to know your needs	−0.28	−0.19
E4. Not realistic to expect them to have your best interests at heart	0.35	−0.15
E5. Not to be expected to have operating hours convenient to all	−0.47	−0.24

surveys among its customers to identify the most important areas for improvement. If a company is not aware of its customers' requirements, it cannot devote resources effectively to improving those areas that may have the most significant impact on customer satisfaction.

According to customer opinion, the second largest gap existed in the tangible dimension, whereas for employees this was the dimension that was of the greatest importance. Among the four items in the tangibles dimension, both customers and employees indicated the greatest gap was related to whether the 'employees are well dressed and neat in appearance'. The second largest gap was found to be in the area of 'having up-to-date equipment' for the customers; the third most important gap for employees. The second most important gap perceived by employees was in the area of appearance of the physical facilities in keeping with the type of services provided. However, customers perceived this to be the area of least importance.

The questions in the areas of responsiveness and empathy were negatively worded. As a result, for both the perception and expectation scales, lower scores were more acceptable than higher scores. A negative gap between perception and expectations means that perception is better than expectations. The negative average gaps in responsiveness indicate that the company is doing better than expected. Among the four individual items of responsiveness, there was a positive gap only between customer perception and expectations in the domain

'it is all right if the employees are too busy to respond to your request promptly'. This means that customers think that the company could do more to meet their expectations. In the area of empathy, the only item showing a positive gap was the statement 'it is not realistic to expect the company to have your best interests at heart'. The company thus did not meet the expectations of its customers in this area. The negative wording may have caused the gaps in these two dimensions. This result provides further evidence of problems associated with the using of negative wording in the SERVQUAL instrument.

For assurance, the greatest gaps perceived by both employees and customers were in relation to the statement 'employees get adequate support from the firm to do their jobs satisfactorily'. Employees perceived a gap in this area measured at -1.15, the biggest gap among all items. This clearly shows that this company needs to provide more support to front-line employees in order that they can improve the quality of the services they are required to provide to customers. The second largest gap was related to 'employees being polite'. It was very interesting to note that the employees themselves knew that they were not being polite enough to customers. Therefore this is another important aspect that can be improved upon.

Focus group discussions and managerial implications

To help the company identify areas for improvement and implement changes to enhance service quality, the results of the analyses were presented to and discussed with a group of senior and middle-level managers. We first asked them whether the five dimensions of service quality made sense to them and whether any items were missing in the SERVQUAL instrument. The majority of the participants felt that the five dimensions did make sense. However, they also felt that some of the items in the SERVQUAL instrument were a little confusing, especially the negatively worded items. They also indicated that more items directly related to the retail industry should be added.

After careful discussion and analysis of the gaps identified above, the group identified the following specific areas as areas that would benefit from improvement:

(1) The company will provide more training to their front-line service staff to enhance their customer service skills. The training should focus on staff ability to help customers resolve their queries and problems quickly. In the process of resolving such problems, they should show a caring attitude and a sincere interest in helping customers. Furthermore, employees should improve their knowledge and skills so that they can provide a fast and dependable service to their customers. When they promise to do something for the customer within a certain time, they must fulfil that promise.
(2) To reduce the gap in the dimension labeled 'tangibles', the group proposed designing new uniforms for front-line service employees. They also suggested developing guidelines for front-line employees aimed at encouraging these employees to look their best at all times. The group also discussed specific suggestions as to how to upgrade their equipment and improve the layout of the shop to improve the tangibles dimension of service quality.
(3) Since the negatively worded items presented some problems of interpretation of the results concerning the responsiveness and assurance dimensions, the group did not recommend any specific suggestions in terms of areas of improvement. Instead, they suggested modification of the wording of the questions of the questionnaire and that the survey be performed once again after the changes had been made.

(4) To improve the assurance dimension, the group suggested that the company provide more support to empower the employees so that they would be able to provide a more prompt and higher quality service. Furthermore, that front-line staff should be trained to be more polite to customers.

Conclusions

SERVQUAL has undoubtedly had a major impact on the business and academic communities. Although this study shows that the data collected do not support the five-factor structure as proposed by Parasuraman *et al.* (1988), the five dimensions are still useful as a foundation for discussion of and determination of areas for improvement in an organization's service quality. Our focus group discussions on service quality with business professionals for whom service quality is important in their businesses show that these professionals also basically agree with the prescription of the five dimensions of service quality. Using the SERVQUAL instrument, we were able to help a company identify important areas for improvement in its business. We also found that employees and customers have significant differences in opinion in terms of the gaps between their perceptions and expectations of that business on their own accounts. This study was therefore able to highlight how important it is for a company to conduct a survey of and consider the opinions of its customers, in particular, and its employees, in identifying areas for service quality improvements.

This study also found that the gap model of service quality does not perform as well as the perception-based performance measures of service quality in terms of its factor structure. Furthermore, we also determined that the negative wording in the SERVQUAL instrument caused serious problems in terms of the reliability and practical interpretation of the data collected and therefore also of the results obtained. We thus concluded that negative wording would be best avoided in future use of this instrument in conducting similar types of research.

This is one of first studies to test the validity of the five-dimension structure of service quality and the reliability of the SERVQUAL instrument in the retail industry in China. Our results indicate that the five dimensions of service quality did not emerge from the data collected in this company. These results may be in part due to cultural differences between China and Western countries. However, questions remain: What is the dimensionality of service quality in China? How do Chinese consumers evaluate service quality? How do cultural factors influence customer evaluation of service quality? How can service quality measurement questionnaires be designed and adapted in order that they suit and provide accurate reflections of a local environment? These questions and others still remain as central questions for future research activities. As China's enters the WTO, companies in China will have to improve significantly the quality of their services in order to enable themselves to compete successfully in the global marketplace. It is therefore very important for them to know how customers evaluate service quality and what they can do to measure and improve service quality. Further research in the area of service quality will in these circumstances be soon in great demand.

Acknowledgements

This research was support by a Faculty Direct Allocation Grant from the City University of Hong Kong (Grant No. 710062) and a direct grant from the Chinese University of Hong Kong (Grant No. 2070235)

References

BABAKUS, E. & BOLLER, G.W. (1992) An empirical assessment of the SERVQUAL scale, *Journal of Business Research*, 24, pp. 253-268.

BOLTON, R.N. & DREW, J.H. (1991) A longitudinal analysis of the impact of service changes on customer attitudes, *Journal of Marketing*, 55, pp. 1-9.

CRONIN, J.J., JR & TAYLOR, S.A. (1992) Measuring service quality: a reexamination and extension, *Journal of Marketing*, 56, pp. 55-68.

CRONIN, J.J., JR & TAYLOR, S.A. (1994) SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality, *Journal of Marketing*, 58, pp. 125-131

DONTHU, N. & YOO, B. (1998) Cultural influences on service quality expectations, *Journal Service Research*, 1, pp. 178-186.

FREEMAN, K.D. & DART, J. (1993) Measuring the perceived quality of professional business services, *Journal of Professional Services Marketing*, 9, pp. 27-47.

JING, S. (1999) Where is the way for large-sized department stores?, *Market Daily* (in Chinese), 22 October.

KETTINGER, W.J. & LEE, C.C. (1994) Perceived service quality and user satisfaction with the information services function, *Decision Sciences*, 25, pp. 737-766.

LAM, S.K. (1997) SERVQUAL: a tool for measuring patients' opinions of hospital service quality in Hong Kong, *Total Quality Management*, 8, pp. 145-152.

LI, M. (1998) The large stores should be adjusted, *Market Daily* (in Chinese), 10 February.

MATTILA (1999) The role of culture in the service evaluation process, *Journal of Service Research*, 1, pp. 250-261.

PARASURAMAN, A., BERRY, L.L. & ZEITHAML, V.A. (1993) Research note: more on improving service quality measurement, *Journal of Retailing*, 69, pp. 140-147.

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, pp. 41-50.

PARASURAMAN, A., ZEITHAML, V.A. & BERRY, L.L. (1988) SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality, *Journal of Retailing*, 64, pp. 12-40.

PARASURAMAN, A., ZEITHAML, V. & BERRY, L.L. (1994) Reassessment of expectations as a comparison standard in measuring service quality: implications for further research, *Journal of Marketing*, 58, pp. 111-124.

PARISEAU, S.E. & MCDANIEL, J.R. (1997) Assessing service quality in schools of business, *International Journal of Quality and Reliability Management*, 14, pp. 204-218.

REEVES, C.A. & BEDNAR, D. (1994) Defining quality: alternatives and implications, *Academy of Management Review*, 19, pp. 419-445.

TEAS, R.K. (1994) Expectations as a comparison standard in measuring service quality: an assessment of a reassessment, *Journal of Marketing*, 58, pp. 132-139.

TRIBE, J. & SNAITH, T. (1988) From SERVQUAL to HOLSAT: holiday satisfaction in Varadero, Cuba, *Tourism Management*, 19, pp. 25-34.