

10 YEARS OF SERVICE QUALITY MEASUREMENT: REVIEWING THE USE OF THE SERVQUAL INSTRUMENT*

Simon Nyeck

ESSEC BUSINESS SCHOOL, PARÍS

**Miguel Morales
Riadh Ladhari**

UNIVERSITÉ LAVAL, QUEBEC, CANADÁ

Frank Pons

UNIVERSITY OF SAN DIEGO, ESTADOS UNIDOS

Resumen

En 1988, Parasuraman, Zeithaml y Berry elaboraron un instrumento para medir la calidad del servicio. Desde esa fecha, este instrumento ha sido utilizado en numerosos estudios sobre distintas industrias y en diferentes países, tanto por académicos como por profesionales. Sin embargo, a pesar de su amplia difusión, pocos estudios tratan los aspectos de dimensionalidad y validez de esta escala de medición. El presente artículo describe las prácticas observadas con relación a estos aspectos a través del análisis de los estudios que han usado SERVQUAL durante los últimos diez años. A partir de una muestra de 60 trabajos empíricos que usan la escala SERVQUAL, se analiza los principales aspectos de validez tratados por cada autor, empleando una plantilla de análisis adaptada del estudio de Stokes y Miller (1975). Con base en los datos disponibles, el estudio sugiere que la escala desarrollada por Parasuraman, Zeithaml y Berry (1988) no presenta una estructura dimensional estable de cinco factores. Finalmente, el artículo evalúa la influencia de las características del diseño de la investigación sobre la confiabilidad de SERVQUAL.

Introduction

During the last decade, research on service marketing focused mainly on the analysis of service quality. Consequently, the studies conducted by Parasuraman, Zeith-

aml and Berry (1985, 1988, 1991, 1994), Grönroos (1983, 1984, 1993) and Eiglier and Langeard (1987) emphasize the importance of conceptualization and measurement of the service quality construct. Several researchers in this discipline emphasize the explanation of the perceived quality by using the SERVQUAL dimensions, reproducing, in general, the process followed by Parasuraman *et al.* (1988).

* Trabajo presentado a la 31ª conferencia de la European Marketing Association (EMAC), llevada a cabo en Braga, Portugal, del 28 al 31 de mayo de 2002.

The popularity of SERVQUAL with researchers can be explained mainly by its ease of use and by its adaptability to diverse service sectors. Even if certain researchers have only retained the concept of gap analysis as operationalization of perceived service quality, it appears that the SERVQUAL model remains the most complete attempt to conceptualize and measure service quality. Nevertheless, over the years, the acceptance of the model proposed by PZB as a «standard» instrument was called into question. Authors thus proposed other conceptualizations (Grönroos 1993; Haywood-Farmer 1988; Iacobucci et al. 1994; Johnston 1988) as well as other measurement instruments (Cronin and Taylor 1992; Brown et al. 1993).

This study has a twofold objective: first, the research describes a state of practices regarding validity in research that has used SERVQUAL; second, the research evaluates the influence of research design characteristics on SERVQUAL reliability.

1. Conceptual Background: The SERVQUAL Instrument

For PZB (1988), perceived quality is the result of the comparison between what consumers consider the service offered by the company (i.e., their expectations) and their perceptions of the performance of the service provided. Service quality is thus considered to be the difference between the perceptions and the expectations of consumers. In 1988, Parasuraman, Zeithaml and Berry broke down ten dimensions into 97 items (approximately 10 items per dimension). Five dimensions were finally retained: reliability, presence of tangible elements, confidence, help-

fulness, and empathy. These five dimensions are broken into 22 items. Each item is further split into two more items, one measuring the expectations having to do with those companies belonging to the service sector in question, the other serving to measure the perception of the service offered by a particular company xyz.

Since the scale was developed using customers of five service sectors (repair and maintenance of small electrical appliances, banking, long distance telephone, title brokerage, and credit cards), Parasuraman et al. (1988) concluded that «SERVQUAL offers a variety of potential applications. It can be used to evaluate the expectations of customers... as well as their perceptions... for a wide range of services and distribution organizations». Several researchers have therefore used SERVQUAL to measure service quality in various sectors such as health (Babakus and Mangold 1992; Headley et al. 1993), banking (Brown et al. 1993; Pitt et al. 1995), fast food (Lee and Ulgado 1997), professional services (Freeman and Dart 1993), retail trade (Gagliano and Hathcote 1994), and advertising (Quester et al. 1995).

Nevertheless, the dimensionality and the psychometric properties of SERVQUAL have caused a lively controversy, since the studies which have used SERVQUAL do not always mention a standard dimensional structure. In effect, the stability of the factorial structure is not demonstrated, nor is its invariance across various sectors proven. This led Babakus and Boller (1992) to conclude that the measure of service quality offers a challenge. Furthermore, the results regarding validity of the instrument are mitigated. This conclusion concerning the dimen-

sionality and psychometric properties of the scale also appears in several studies (Asubonteg et al 1996; Kettinger and Lee 1994; Csipak et al. 1994) which provide a comparative evaluation of works having used SERVQUAL.

2. Research Methodology and Data

This research sets out with a dual purpose: First, we describe research practices regarding the utilization of SERVQUAL. We pay particular attention to the evaluation of reliability, convergent, discriminant and predictive validities, as well as to the dimensionality of the SERVQUAL instrument; second, this research evaluates the existing relationship between the research design criteria and SERVQUAL's reliability. We must mention, however, that our research will be limited to the study of the effect of design on the reliability coefficients because the indicators concerning convergent, discriminant and predictive validities are sometimes nonexistent in various studies.

The studies carried out by Churchill and Peter (1984) and Peterson (1994) have allowed us to select those design criteria that affect reliability. Thus, (1) the sample size, (2) the number of points on the scale, (3) the number of items. The following research propositions are mentioned:

P1. Sample size: Churchill and Peter (1984) found a negative relationship between sample size and the alpha coefficient. However, Peter (1994) observed the absence of such a relationship.

P2. The number of points on the scale used: The available literature does not seem to show any agreement regarding

the effect of the number of points on the scale on reliability. Bendig (1953, 1954) observed that reliability is independent of the number of points on the scale. Further, Churchill and Peter (1984) and Peterson (1994) found significant relationships between reliability and the number of points on the scale. We expect a positive relationship between these two variables.

P3. The number of items: The results obtained by Churchill and Peter (1984) and by Peterson (1994) show the presence of this positive relationship between the two variables. We expect the alpha coefficient will increase as a function of the number of items.

The Sample

The sample is made up of forty (40) articles published since the appearance of SERVQUAL (1988) that we have collected from 18 periodicals. However, an article could contain the analysis of service quality in more than one sector; in such a case we considered the sector examined to be a sample unit, which gives a total of 60 observations. For an article to be retained, it had to fit the following three criteria: use of SERVQUAL or of a modified SERVQUAL scale; study of service quality in a given sector following an empirical method; and supplying, in results, indicators concerning the reliability, validity or dimensionality of the scale (Table 1).

Data Analysis Grid

The research used Stokes and Miller's grid (1975) to evaluate the use of SERVQUAL scale. The evaluation grid was divided into four headings: General char-

Table 1: Description of the sample of the evaluated studies

Publication	Number of articles inventoried	Number of sectors inventoried (n)
<i>Asac proceedings</i>	3	3
<i>Advances in Consumer Research</i>	1	1
<i>Decision Sciences</i>	1	1
<i>International Journal of Service Industry Management</i>	4	5
<i>Journal of Business Industrial Marketing</i>	1	1
<i>Journal of Business Research</i>	1	1
<i>Journal of Health Care Marketing</i>	4	5
<i>Journal of Marketing</i>	1	2
<i>Journal of Professional Services</i>	2	3
<i>Journal of Retailing</i>	7	17
<i>Journal of Services Marketing</i>	4	4
<i>Managing Service Quality</i>	1	2
<i>MIS Quarterly</i>	1	3
<i>MRN</i>	1	1
<i>Public Administration Quarterly</i>	1	1
<i>Service Industries Journal</i>	6	8
<i>Total Quality Management</i>	1	2
Total	40	60

acteristics, formulation of the problem, data collection, and analysis of data. The final part of the evaluation grid contains six criteria. We determined the number of final dimensions presented in each sector inventoried. Finally, we categorized the convergent, discriminant and predictive validities in three ways: the article presented statistical validation results; the article discussed only the validity without supporting the discussion with «statistical» measures; the article made no reference to the validity of the scale. In

all cases, we classified the studies by taking into account only the information provided within each article. As for reliability, we took Cronbach alpha mean, when the alphas were given by dimension.

Finally, to complete the explanatory part of our research, we recoded the initial data in order to regroup it into categories, as was done by Peterson (1994). The research propositions are tested with non-parametric tests due to the small size our sample and sub-groups.

3. Findings and Discussion

The use of SERVQUAL in several sectors raises questions on the number of dimensions and their stability from one context to another. In most of the cases (79%), the number of dimensions varies between one (McAlexander et al. 1994; Simon 1997; Brown et al. 1993) and nine (Carman 1990). This result invalidates the invariance of the scale's structure. Figure 1 illustrates the instability associated with the number of dimensions.

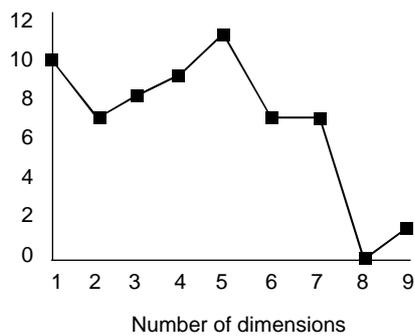


Figure 1: Dimensionality of SERVQUAL

Regarding the dimensionality of SERVQUAL, our study has brought to light the results reached by several authors who duplicated the SERVQUAL scale (Carman 1990; Babakus and Boller 1992). The dimensional structure is very unstable, even within a given sector. While the original study by Parasuraman et al. (1988) proposed five «universal» dimensions which were supposed to measure service quality in any sector, the vast majority of studies report a number of dimensions other than five. This result supports the work of Eiglier et al. (1989), who found that quality is a relative notion with respect to a given client segment.

Validity of the measuring instrument.

Three indicators of validity are generally mentioned by researchers who use SERVQUAL: convergent validity, discriminant validity and predictive validity. This critical evaluation of the use of SERVQUAL reveals that, in spite of «acceptable» reliability indicators, other psychometric properties of the instrument have not been established. Table 2 illustrates the failure to account for discriminant and convergent validity (only 44,3% of the cases).

Table 2: Convergent, discriminant and predictive validity

	Convergent Validity		Discriminant Validity		Predictive Validity	
	Frequency	%	Frequency	%	Frequency	%
Discussion	08	13,1	15	24,6	–	–
Indicator	26	42,6	11	19,7	10	16,4
No information	26	44,3	34	54,1	50	83,6
Total	60	100	60	100	60	100

Finally, the research did not show any relationship between number of items and reliability. This result differs from those obtained by Churchill and Peter (1984) and by Peterson (1994). The differences between these studies and our own can be explained by the size of our sample (40 articles) and by the use of a mean alpha. For example, the reliability coefficient was generally given by dimension and validity indicators presented in the studies were heterogeneous, when they were not totally absent. The effect of the two other design criteria (sample size and method of administration of the questionnaire) is not significant.

The results suggest that few researchers concern themselves with the validation of the measuring tool. This reinforces the comments made by Brown et al. (1993), who point out that discriminant validity of the measuring tool for service quality ought to be improved. Furthermore, Stokes and Miller's analysis grid (1975) was a useful tool in this evaluation, and could be adapted to evaluate practices related to other measurement scales. We also recommend that researchers explore alternatives to the conceptualization of service quality.

Selected references

- ASUBONTEG, Patrik; MCCLEARY, Karl J. and SWAN, John. 1996. Servqual revisited: A critical review of service quality. *Journal of Services Marketing*. Vol. 10, Iss. 6, p. 62-81.
- BABAKUS, Emin and BOLLER, Gregory W. 1992. An empirical assessment of the SERVQUAL scale. *Journal of Business Research*. Vol. 24, Iss. 3, p. 253-268.
- BROWN, Tom J; CHURCHILL, Gilbert A. Jr. and PETER, J. Paul. 1993. Research note: Improving the measurement of service quality. *Journal of Retailing*. Vol. 69, Iss. 1, p. 127-139.
- CARMAN, James H. 1990. Consumer perceptions of service quality: An Assessment of the SERVQUAL dimensions. *Journal of Retailing*. Vol. 66, Iss. 1, p. 33-55.
- CHURCHILL, Gilbert A. Jr. and PETER, J. Paul. 1984. Research design effects on the reliability of rating scales: A Meta-Analysis. *Journal of Marketing Research*. Vol. 21, Iss. 4, p. 360-375.
- CRONIN, J. Joseph Jr. and TAYLOR, Steven A. 1992. Measuring service quality: A re-examination and extension. *Journal of Marketing*. Vol. 56, Iss. 3, p. 55-68.
- CSIPAK, James; CHEBAT, Jean-Charles and VENKATESAN, Ven. 1994. Measurement of perceived service quality in the purchase of airline tickets: An assessment of SERVQUAL's reliability and validity. *ASAC Proceedings*. Halifax, Nova Scotia. Vol. 14, Iss. 3, p. 48-58.
- FINN, David W. and LAMB, Charles W. 1991. An evaluation of the SERVQUAL scales in retail setting. *Advances in Consumer Research*. Vol. 18, p. 483-490.
- GRÖNROOS, Christian. 1984. A service quality model and its marketing implications. *European Journal of Marketing*. Vol. 18, Iss. 4, p. 36-44.
- GUMMESSON, E. 1992. Quality dimensions: What to measure in service organizations. In Swartz, Bowen, and Brown (editors). *Advances in Services Marketing and Management*. JAI Press Inc. Vol. 1, p. 49-64.
- MITTAL, Banwari and LASSAR, Walfried M. 1996. The role of personalization in service encounters. *Journal of Retailing*. Vol. 72, Iss. 1, p. 95-109.
- PARASURAMAN, A.; ZEITHAML, Valarie A. and BERRY, Leonard L. 1985. A conceptual model of service quality and its implications for future research. *Journal of Marketing*. Vol. 49, Iss. 4, p. 41-50.
- . 1988. SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*. Vol. 64, p. 12-40.
- PETER, J. Paul and CHURCHILL, Gilbert A., Jr. 1986. relationships among research design choices and psychometric properties of rating scales: A meta-analysis. *Journal of Marketing Research*. Vol. 23, Iss. 1, p. 1-9.
- PETERSON, Robert A. 1994. A meta-analysis of Cronbach's Coefficient Alpha. *Journal of Consumer Research*. Vol. 21, Iss. 2, p. 381-391.
- STOKES, C. Shannon and MILLER, Michael. 1975. A methodological review of research in rural sociology since 1965. *Rural Sociology*. Vol. 40, n.º 4, p. 411-434.