

---

# Comparing tools for service quality evaluation

Fiorenzo Franceschini and Marco Cignetti

*Politecnico di Torino, Torino, Italy and*

Mara Caldara

*MTS Systems S.R.L., Torino, Italy*

## 1. Introduction

In recent times we have observed a growing importance of service, offered as support for products delivered, or as an entity in itself given by tertiary industries to firms and societies (Franceschini and Rossetto, 1995a; Hauser and Clausing, 1988; Parasuraman *et al.*, 1996).

In spite of this, the problem of service quality evaluation still exists.

The matter is particularly delicate for two reasons: first, human presence and service intangibility and, second, the dependence on the delivering process (Grönroos, 1982; Parasuraman *et al.*, 1996).

The attempt to define an evaluation standard independent of any particular service context has stimulated the setting up of several methodologies (Carman, 1990; Cronin and Taylor, 1994; Hayes, 1992; Parasuraman *et al.*, 1991; 1994; 1996; Teas, 1994).

On one hand, all these papers declare the great interest in this problem, and on the other hand, they cause potential users a serious difficulty in choosing a proper tool for their particular needs.

This paper is aimed at providing an orientation map for anyone faced with the problem of service quality evaluation. The map points out the outstanding features and weaknesses for each evaluation tool.

Two of the methods presented are then used to evaluate service quality of a customer post-sales assistance and servicing.

Finally we present some results and considerations of our experimentation.

## 2. Quality evaluation models

Available literature provides plenty of service quality evaluation methodologies. Some come as a result of the realization of conceptual models produced to understand the evaluation process (Parasuraman *et al.*, 1985), and others come from empirical analysis and experimentation on different retailing sectors (Cronin and Taylor, 1992; Franceschini and Rossetto, 1997b; Parasuraman *et al.*, 1988).

Parasuraman, Berry and Zeithaml first attempted to compare and classify different methods in 1991. In this paper, a set of different versions of one of the most famous tools (SERVQUAL) was evaluated according to some analysis criteria such as data collection, sample size, questionnaire format used to collect

---

data, items number, questionnaire dispensing, data analysis, dimensions number considered for service evaluation and questionnaires' reliability.

Starting from this first comparison, we have tried to extend the analysis to other methods. A further set of analysis criteria such as theoretical method baseground, customer-tool interference degree, idiosyncratic effect on interviewed customers, and data pre-elaboration, had also been considered in the analysis.

Except for QUALITOMETRO which is still under an advanced experimentation, the other methods are typically used in practice. A detailed description of each method can be found in Cronin and Taylor (1992); Franceschini and Rossetto (1997a); Parasuraman *et al.* (1991); Schvaneveldt *et al.* (1991); and Teas (1994). Here we summarize only some important features.

SERVQUAL was developed by Parasuraman, Zeithaml and Berry (PZB-model) (Parasuraman *et al.*, 1991; 1993; 1994). SERVQUAL was inspired by a conceptual model offered in 1985 by the same authors. Service quality is evaluated by calculating the difference (gap) between what the customer expects and what he/she really perceives.

SERVQUAL (1991 version) is structured into three sections. The first and third sections propose 22 questions for the evaluation of expectations and perceptions respectively. The second section asks the customer for the importance of each service quality dimension. Service quality evaluation is obtained by comparing expectations and perceptions values.

Schvaneveldt *et al.* (1991) evaluated service quality from two perspectives. The first "objective" involved the presence or absence of a particular quality dimension, and the second "subjective" involved the users' resulting sense of satisfaction or dissatisfaction. A questionnaire was administered to customers to evaluate service quality.

Cronin and Taylor (1992) proposed a method called SERVPERF (Cronin and Taylor, 1994). The main feature of SERVPERF is its focus on customers' perceptions. According to Cronin and Taylor, this procedure gives better results than SERVQUAL and reduces the number presented to service users.

To better define the meaning of expectations, Teas (1993) proposed the NQ model (Normed Quality). Expectations may be interpreted by customers in two different ways: at the ideal level, by giving each attribute the highest score, or at the feasible level when considered under the actual conditions in which service may be delivered. The NQ method focuses interviewees' attention towards two kinds of expectations, but asks the customer for another set of questions, stimulating potential idiosyncratic effects.

The last observed tool is QUALITOMETRO, conceived for evaluation and "on-line" service quality control. The tool was developed and proposed by Franceschini and Rossetto (1997b).

An interesting feature of this method is the possibility of a separate "measurement" of expected ( $Q_e$ ) and perceived ( $Q_p$ ) quality without the potential for cross-influence.  $Q_e$  is observed as *ex-ante* service use, and the second as *ex-post* on the same questionnaire. It is important to remember that all other tools ask for a contemporary *ex-post* evaluation.

---

The QUALITOMETRO method is based on service quality dimensions (determinants) proposed in the PZB-model (Parasuraman *et al.*, 1985; 1988). It allows an online quality monitoring of the differential  $\partial Q$  between expected and perceived quality, and it may also be used in situations where there are periodical service users (Franceschini and Rossetto, 1997a; Oliver, 1981).

Online monitoring is developed by means of a “*p*” control chart. Table I shows similarities and differences among analyzed methods. For completion, it is important to remember that every tool gives a short introduction about reasons and ways for questionnaire dispensing (Cignetti, 1996).

Table I illustrates the variety of service sectors considered in the analysis, from telephone companies to supermarkets, to libraries, in both public and private enterprises.

The methods compared present a great difference in the number of questions delivered to customers. One passes from a minimum of 8 + 8 questions (8 for expectations and 8 for perceptions) for QUALITOMETRO to a maximum of 10 + 10 + 10 + 10 + 10 questions for NQ method.

The number of items proposed is an extremely delicate factor for a questionnaire. If it is true that the more the items dispensed the higher is the “information” available, it is also true that items in themselves may stimulate a clear idiosyncrasy and tiredness during administration (Drew and Castrogiovanni, 1995). This fact indicates a lowering of interviewee involvement and a loss of information trustworthiness.

A typical problem for the measurement of a physical magnitude is that the effectiveness of a questionnaire is as high or low as the interactions between measure and instrument; in this sense Table I gives a qualitative index for the degree of customer-tool interference.

An important issue emerging from Table I regards data pre-elaboration and subsequent aggregation. It is opportune to remember here that in subjective evaluations a metrological reference chain does not exist as in the case of physical quantities (temperature, length, etc. (Franceschini and Rossetto 1997a)). Each individual gives indications according to his or her own reference system that is usually unique to that person. The homogeneity hypothesis adopted for individuals’ reference systems is then critical for the aggregation and interpretation of data collected from different individuals.

A second delicate problem is the numerical coding of judgements given by evaluators.

As Table I shows, every tool uses semantic evaluation point scales (i.e. 1-7 or 1-5) to qualify the particular scale level. During data pre-elaboration, qualitative scales are converted into numerical interval scales (a linear interval scale allows object setting so that the differences between side elements are the same. Interval scales, without any scale origin, allow equality/inequality, ordering and subtraction operations) and any symbol is interpreted as a number. On these numbers statistical elaboration is then carried out.

The scalarization of collected data presents two main problems. The first concerns the introduction of an arbitrary metrics system (Franceschini, 1996;

	Revised SERVQUAL Parasuraman <i>et al.</i> (1991)	Two-Way Schvaneveldt <i>et al.</i> (1991)	SERVPERF Cronin and Taylor (1992)	Normed Quality Teas (1994)	QUALITOMETRO Franceschini and Rossetto (1997b)
Theoretical background	The determinants method of service quality and gap theory. Service quality is calculated as difference between perceptions and expectations with importance weights given to each dimension according to the formula $QS = \sum_i I_i (P_i - E_i)$	Latent evaluation factors: service quality is evaluated by answers given by customers to questions about "objective" (quality attributes) and "subjective" (satisfaction levels)	Service quality is evaluated by perceptions only without expectations and without importance weights according to the formula $QS = \sum_i P_i$	The problem for expectation run to a redefinition of this component and discriminate between ideal expectation $I$ and feasible expectation $A_e$ to calculate service quality according to the formula $QS = - \sum_i I_i ((P_i - I_i) - (A_{ei} - I_i))$	The determinants of service quality. Customer expectations and perceptions are evaluated in two distinct moments. Quality evaluation is carried out by means of a comparison between quality expectations and perceptions profiles using MCDA
Data collection sample features	two telephone companies, two insurance companies, two banks	banks, restaurants, laundries, supermarkets	two banks, two pest control companies, two laundries, two fast food	three big department stores	Library facility at DISPEA Department
Sample size	290 to 487 according to companies	330	660	120	100
Items number (expectations plus perceptions)	22 + 22	not declared	22	10 + 10 + 10 + 10 + 10	8 + 8
					(Continued)

**Table I.**  
Comparison of some  
methods for service  
quality evaluation  
(Cronin and Taylor,  
1992; Franceschini  
and Rossetto 1997a;  
Parasuraman *et al.*,  
1991; Schvaneveldt  
*et al.*, 1991; Teas, 1994)

Table I.

	SERVQUAL Parasuraman et al. (1991)	Two-Way Schvaneveldt et al. (1991)	SERVPERF Cronin and Taylor (1992)	Normed Quality Teas (1994)	QUALITOMETRO Franceschini and Rossetto (1997b)
Response scale	7-point semantic differential	5-point semantic	7-point semantic differential	7-point semantic differential	7-point semantic comparative
Dimensions importance	Weights evaluation with constant sum	not needed	Weights evaluation with constant sum	Weights evaluation with constant sum	7-point semantic comparative
Questionnaire dispensing	Mail	not declared	Mail	interview	Expectations before service use and perceptions after delivering
Customer-tool interference degree	High	Medium	High	High	Low
Idiosyncratic effect	High	Medium	Medium	High	Low
Data pre-elaboration	Scalarization	Scalarization	Scalarization	Scalarization	Without scalarization
Data analysis	Factorial analysis followed by oblique rotation	Factorial analysis	Factorial analysis followed by oblique rotation	Factorial analysis followed by oblique rotation	MCD methods and "p" control chart
Reliability: (Cronbach's alpha coefficient)	0.8 to 0.93	not declared	0.63 to 0.98	Calculated other validity and reliability coefficients	Global quality indicators as reliability factor
Dimensions number	five: Tangibles Reliability Assurance Responsiveness Empathy	five: Performance Security Completeness Ease of use Emotivity/environment	five: Tangibles Reliability Assurance Responsiveness Empathy	five: Tangibles Reliability Assurance Responsiveness Empathy	five: Tangibles Reliability Assurance Responsiveness Empathy

---

Franceschini and Rossetto 1995a); the second concerns the assumption for an identical scale “interpretation” by any interviewee.

The scalarization procedure may generate a “distortion” effect, which can lead to a partly or completely bad interpretation of collected data. Critical matter of the question is that, usually, the extent of distortions that have been introduced is not clear. In other words, original information that is “arbitrarily” enriched in order to simplify its aggregation and elaboration might be highly modified if compared with the one really expressed by customers, with imaginable consequences.

QUALITOMETRO seeks to solve some of these expressed difficulties. For instance, numerical scalarization of collected data is avoided due to the use of multiple criteria analysis techniques (Ostanello, 1985). Data analysis and questionnaire reliability are then carried out.

Table I shows the list of dimensions considered by each method. As we can see, there are great similarities and some differences too, mainly about the meaning assigned to each dimension.

Finally, Table I shows the number and kind of considered dimensions for each method. There are great similarities and some differences too, mainly about the meaning assigned to the considered dimension.

### **3. Comparative experimentation**

The aim of this paper is to evaluate the features of these methods by carrying out a parallel experimentation of SERVQUAL and QUALITOMETRO.

The study is based on a sample of customers of an international enterprise that deals with technical assistance on material testing facilities and laboratory simulations.

Data for analysis were gathered from a sample of 15 customers. SERVQUAL and QUALITOMETRO questionnaires were administered to this group. Ten customers returned completed questionnaires.

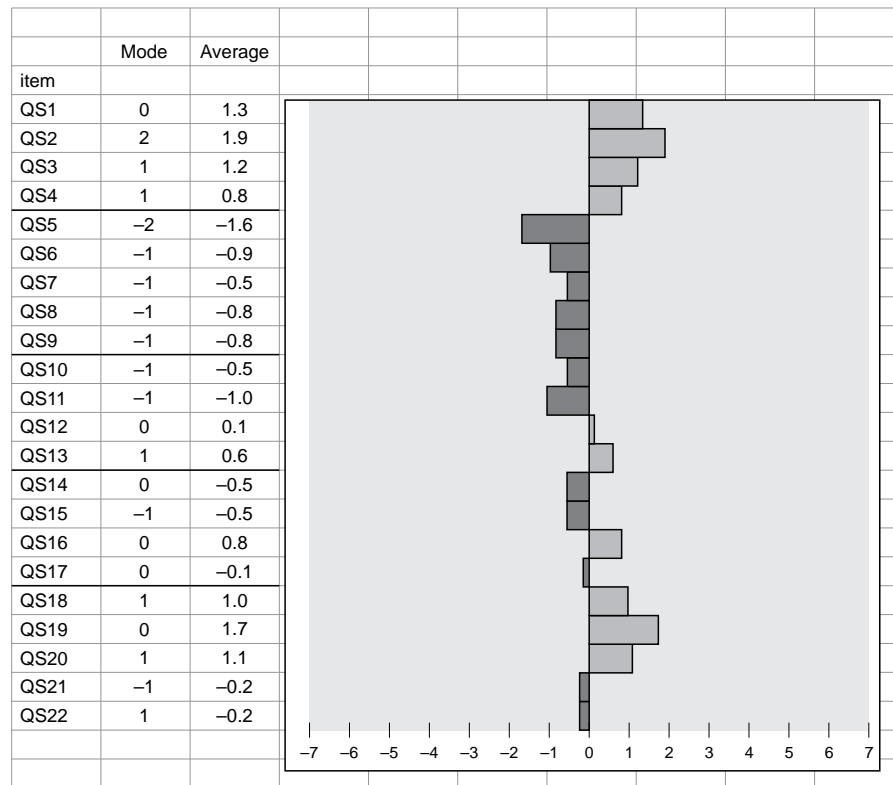
Notwithstanding the limited sample size, the statistical analysis revealed some interesting results. Figures 1 and 2 present the results obtained with SERVQUAL for each item (*QS1*,...*QS22*) and for each dimension (tangibles, reliability, responsiveness, assurance, empathy). Inside each dimension, items with discordant sign are highlighted.

Figures 3 and 4 give the average weights assigned by customers to the importance of each dimension. It is observed that the shapes of the histograms for the two questionnaires are similar. The result also indicates a greater importance for reliability and responsiveness compared to the other three dimensions.

Further investigation shows some differences between QUALITOMETRO and SERVQUAL values. The range of weights is about 12 percentage points for QUALITOMETRO and about 24.5 percentage points for SERVQUAL. This last method discriminates more efficiently the weight of dimensions.

These differences are probably due to the different ways of dispensing questionnaires. SERVQUAL asks customers to share 100 points among the five

**Figure 1.**  
SERVQUAL items score  
(QS1,... QS22) obtained  
for a sample of ten  
questionnaires for a  
post-sales assistance  
service



dimensions, forcing them to give a clear relevance to the most important dimension. QUALITOMETRO, on the contrary, asks for an independent 1-7 score for each dimension.

#### 4. Customers' quality profiles

The purpose of quality profiles is to show at first sight the relevance of perceptions on expectations or vice versa for the five dimensions of service quality.

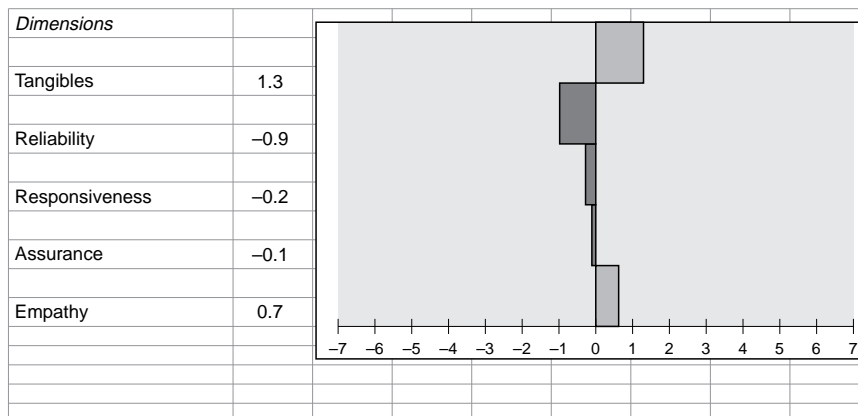
Figure 5 illustrates data obtained with SERVQUAL for each customer and for each dimension. Excepting customers #3 and #6, we can observe some conflict situations where there is not a global relevance between expectations and perceptions.

Figure 6 shows weighted profiles obtained using QUALITOMETRO under the assumption of collected data scalarization.

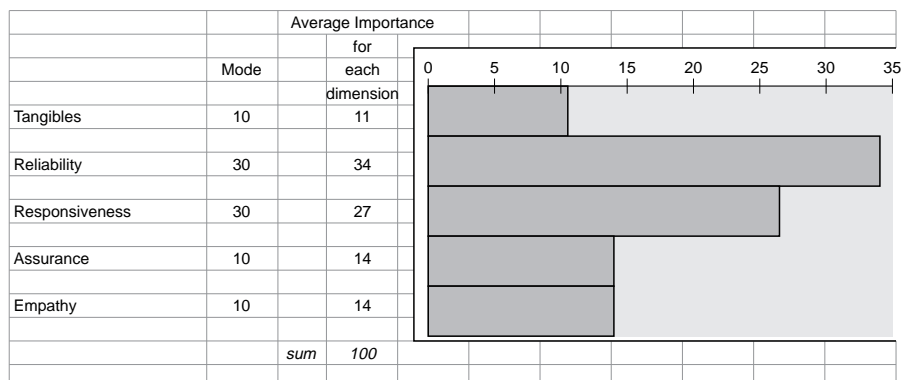
Excepting customers #6 and #10 whose interpretation is not immediate, there is a clear separation between expectations and perceptions profiles.

This result is not surprising because of the propitious features of QUALITOMETRO, consisting of a separate measurement of expectations and perceptions on the same scale before and after service delivery.

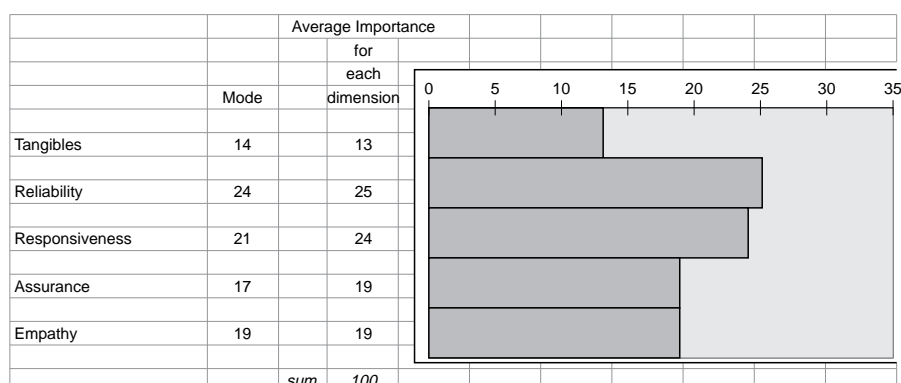
This procedure “guides” customer evaluation when really comparing perceptions with expectations, as highlighted by customer profiles #1, #2, #3, #4, #5, #7, #8, #9 (see Figure 6).



**Figure 2.**  
SERVQUAL dimensions score obtained for a sample of ten questionnaires for a post-sales assistance service

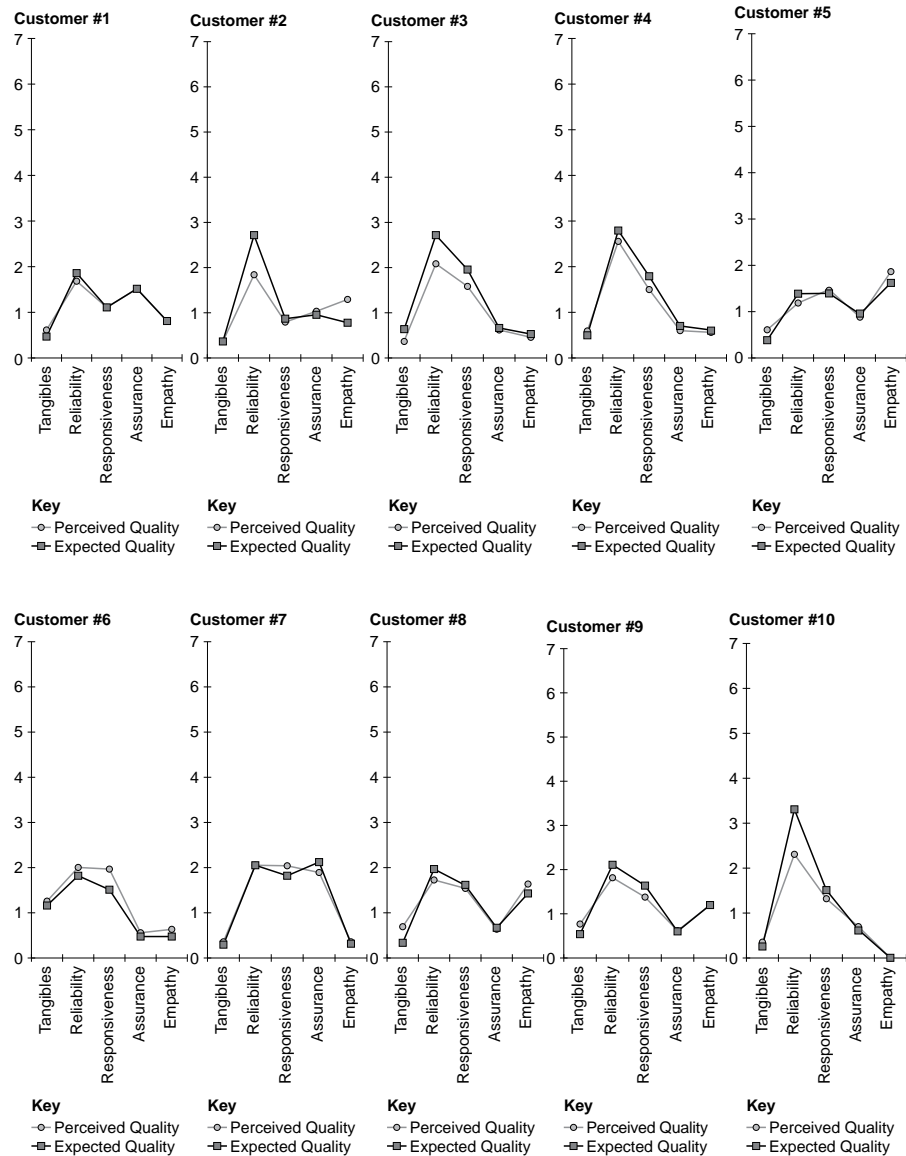


**Figure 3.**  
Importance weights obtained by means of SERVQUAL



**Figure 4.**  
Importance weights obtained by means of QUALITOMETRO

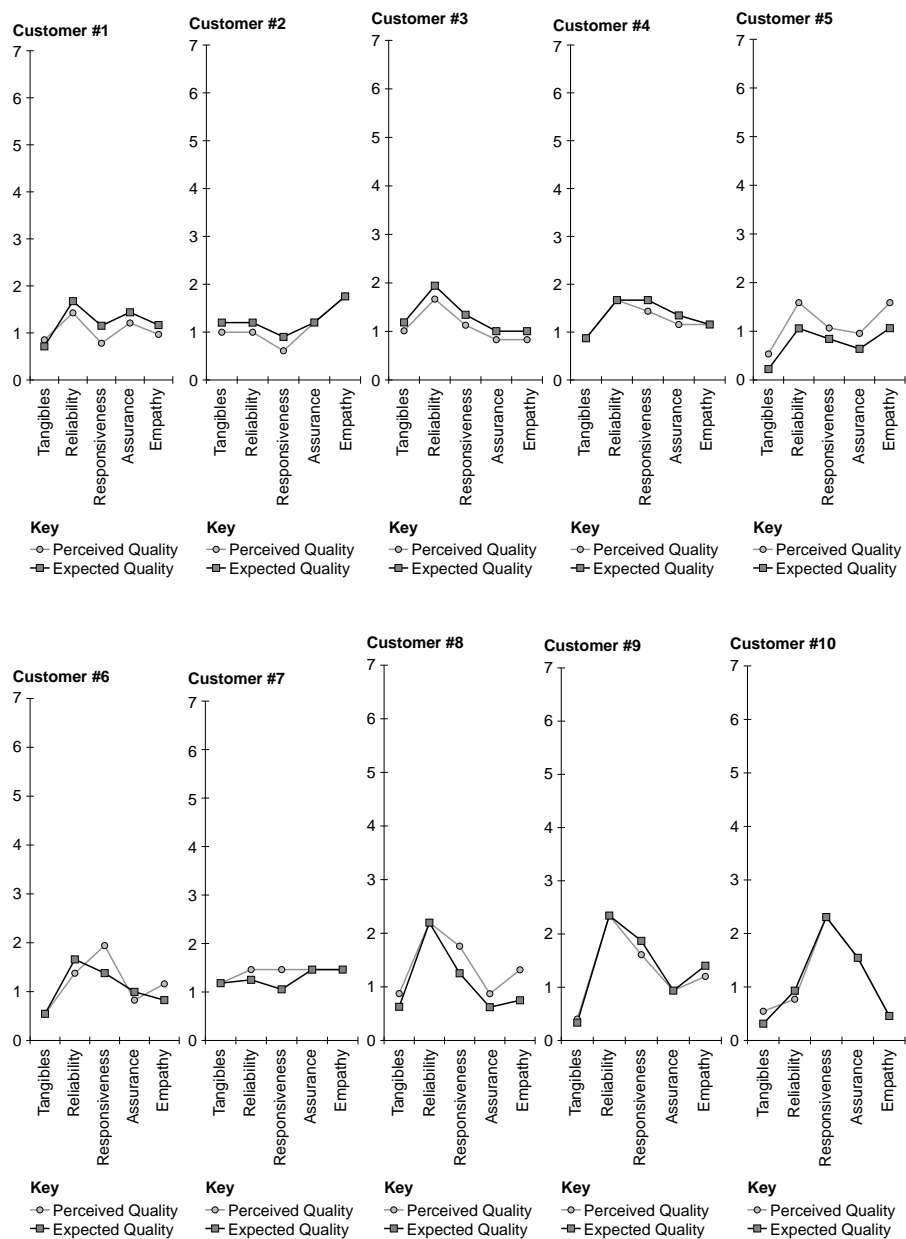




**Figure 5.**  
Weighted profiles of ten  
customers obtained by  
means of SERVQUAL

Other useful considerations may be noted by examining the shape of profiles. Each shape reveals how a certain dimension is considered important compared to the others.

A flat shape shows, for example, a similar importance for each dimension (see customer #7 in Figure 6), while a strong maximum means a clear relevance of one of them. So, for example, customer #10 in Figure 6 gives a greater



**Figure 6.**  
Weighted profiles of ten  
customers obtained by  
means of  
QUALITOMETRO

attention to responsiveness, neglecting other dimensions such as tangibles, thus showing a positive value between perceptions and expectations.

The shape of the profile shows a prompt understanding of customers' needs, allowing the company to pursue a customized and tailored service.

---

From Figure 6, for instance, we can classify customers #1, #3 and #4 under the same group revealing a clear relevance for reliability, then customers #2 and #6 in another group considering empathy as a determining factor for service quality.

Customers #8, #9, and #10 give a greater attention to either reliability or responsiveness.

The analysis of profiles is capable of supporting customers' portfolio segmentation.

## 5. Conclusions

Parallel experimentation carried out by examining QUALITOMETRO and SERVQUAL allowed a confirmation of qualities of both questionnaires, but showed some problems too.

The impressions obtained from customers indicate the usefulness of both tools, but they "complained" that SERVQUAL requires an excessive length of time to answer. On the other hand, QUALITOMETRO appears to be easy to use.

A clear advantage of SERVQUAL is its ability to obtain the importance of the weights for the dimensions in a better way than QUALITOMETRO. The experimentation conducted here confirmed this advantage.

Our results show the possibility of using quality profiles to cluster groups of customers with similar needs, thus enabling the company to customize its service delivery.

The orientation map described here may be a useful tool in helping and guiding users in the selection of service quality evaluation methods. The selection of the most appropriate tool depends on the particular context where service quality is to be evaluated.

Further development of the work will be on the direction of increasing the sample size and improving the statistical analysis of quality profiles.

## References

- Carman, J.M. (1990), "Consumer perceptions of service quality: an assessment of the SERVQUAL scale", *Journal of Retailing*, Vol. 66 No. 1, Spring, pp. 33-55.
- Cignetti, M. (1996), "La valutazione della Qualità nei servizi: una applicazione nel settore assistenza clienti", Thesis degree, Politecnico di Torino.
- Cronin, J.J. and Taylor S.A. (1992), "Measuring service quality: a reexamination and extension", *Journal of Marketing*, Vol. 56, pp. 55-68.
- Cronin, J.J. and Taylor, S.A. (1994), "SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality", *Journal of Marketing*, Vol. 58, January.
- Drew, J.H. and Castrogiovanni, C.A. (1995), "Quality management for services: issues in using customer input", *Quality Engineering*, Vol. 7 No. 3, pp. 551-66.
- Franceschini, F. (1996), *Quality Function Deployment: Qualità e Innovazione*, Ed. Politeko, Torino.
- Franceschini, F. and Rossetto, S. (1995a), "QFD: the problem of comparing technical/engineering design requirements", *Research in Engineering Design*, Vol. 7, pp. 270-8.
- Franceschini, F. and Rossetto, S. (1995b), "Quality and innovation: a conceptual model of their interaction", *Total Quality Management*, Vol. 6 No. 3, pp. 221-9.

- 
- Franceschini, F. and Rossetto, S. (1997a), "Design for quality: selecting products' technical features", *Quality Engineering*, Vol. 9 No. 4, pp. 681-8.
- Franceschini, F. and Rossetto, S. (1997b), "On-line service quality control: the 'Qualitometro' method". *De Qualitate*, Vol. 6 No. 1, pp. 43-57, forthcoming in *Quality Engineering*.
- Grönroos, C. (1982), "Strategic management and marketing in the service sector", Swedish School of Economics and Business Administration.
- Hauser, J. and Clausing, D. (1988), "The house of quality", *Harvard Business Review*, Vol. 66 No. 3, pp. 63-73.
- Hayes, B.E. (1992), *Measuring Customer Satisfaction*, ASQC Quality Press, Milwaukee, WI.
- Oliver, R. (1981), "Measurement and evaluation of satisfaction process in retail settings", *Journal of Retailing*, Vol. 57 No. 1, Fall, pp. 25-48.
- Ostanello, A. (1985), "Multiple criteria decision methods and application", in Fandel, G. and Spronk, J. (Eds), *Outranking Methods*, Springer-Verlag, Berlin, pp. 41-60.
- Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1991), "Refinement and reassessment of the SERVQUAL scale", *Journal of Retailing*, Vol. 67 No. 4, pp. 420-50.
- Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1993), "More on improving service quality measurement", *Journal of Retailing*, Vol. 69 No. 1, pp. 140-7.
- Parasuraman, A., Berry, L.L. and Zeithaml, V.A. (1996), "The behavioral consequences of service quality", *Journal of Marketing*, Vol. 60, April.
- 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, Fall, pp. 41-50.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: a multiple-item scale for measuring consumer perception of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1994), "Reassessment of expectations as a comparison standard in measuring service quality: implications for future research", *Journal of Marketing*, Vol. 58 No. 1, pp. 11-124.
- Schvaneveldt, S.J., Enkawa, T. and Miyakawa, M. (1991), "Consumer evaluation perspectives of service quality: evaluation factors and two-way model of quality", *Total Quality Management*, Vol. 2 No. 2.
- Teas, R.K. (1993), "Expectations, performance, evaluation, and consumers' perceptions of quality", *Journal of Marketing*, Vol. 57, July, pp. 132-9.
- Teas, R.K. (1994), "Expectations as a comparison standard in measurement of service quality: an assessment of a reassessment", *Journal of Marketing*, Vol. 58, January.

(Fiorenzo Franceschini is an associate professor in the Department of Manufacturing Systems and Economics at the Polytechnic School of Turin. His current research interests are in the area of service quality, quality engineering and control. He previously held positions at Telettra Telecomunicazioni S.P.A. and other private companies. He may be contacted at [fiorenzo@athena.polito.it](mailto:fiorenzo@athena.polito.it).

Marco Cignetti graduated in mechanical engineering at the Polytechnic School of Turin. His main scientific interest currently concerns total quality management and service quality.

Mara Caldara graduated in mechanical engineering at the Polytechnic School of Turin. She holds positions at MTS Inc. Her main scientific interests are in the area of total quality management and service quality.)