Rethinking perceived service quality: An alternative to hierarchical and multidimensional models

Article *in* Total Quality Management & Business Excellence · January 2010

DOI: 10.1080/14783360903492694

CITATIONS
51

READS
211

2 authors:

Jose A. Martinez
Universidad Politécnica de Cartagena
139 PUBLICATIONS **2,009** CITATIONS

SEE PROFILE

Laura Martínez Caro
Universidad Politécnica de Cartagena
66 PUBLICATIONS **1,309** CITATIONS

SEE PROFILE

Total Quality Management
Vol. 21, No. 1, January 2010, 93–118

Routledge
Taylor & Francis Group

Rethinking perceived service quality: An alternative to hierarchical and multidimensional models

Jose Antonio Martínez García* and Laura Martínez Caro

Department of Management and Marketing, Technical University of Cartagena, Paseo Alfonso XIII, 50.30203, Cartagena, Spain

Measuring perceived service quality continues to be a controversial topic in management literature. Brady and Cronin's (2001) hierarchical and multidimensional model overcame several limitations of previously proposed models. Nevertheless, Brady and Cronin's conceptualisation has been the object of substantial criticism. This research describes the most important limitations of the Brady and Cronin's (2001) model, pointing out some newly identified drawbacks to this kind of conceptualisation and to using the methodology for analysing service quality models. To overcome these shortcomings, a new procedure is proposed, based on the proper identification of service quality attributes and on the study of unobserved heterogeneity in customer perceptions. An application of this procedure to the insurance industry shows the possible advantages of using this method for studying perceived quality in services. The results of the empirical study confirm the presence of several latent classes formed by customers with disparate service quality perceptions.

Keywords: service quality; unobserved heterogeneity; multidimensional and hierarchical model

Introduction

Measuring service quality has been one of the most recurrent topics in management literature during the last three decades, for two main reasons: (1) the need to develop valid instruments for the systematic evaluation of firms' performance from the customer point of view; (2) the association between perceived service quality and other key organisational outcomes. The former has been enhanced by the proliferation of quality management systems such as ISO certification, self-assessment models (e.g. EFQM), and other quality management practices, as well as the increased importance of customer relationship marketing; the latter reason has been supported by relevant research that relates service quality with market share, profitability, customer loyalty, competitive price premium and probability of purchase (Brady & Cronin, 2001).

Several service quality models have been proposed in recent years and have been broadly considered in applied research, although their validity has also been widely debated. Grönroos' (1984) service quality model was the first attempt, and later other important researchers proposed their conceptualisation: Parasuraman et al. (1985), Cronin and Taylor (1992), Teas (1993), Rust and Oliver (1994), Dabholkar et al. (1996), Brady and Cronin (2001). Brady and Cronin's (2001) model is the most important recent contribution; their hierarchical and multidimensional model of service quality overcomes several of the criticisms made of earlier models. Brady and Cronin's (2001)

*Corresponding author. Email: josean.martinez@upct.es

ISSN 1478-3363 print/ISSN 1478-3371 online

DOI: 10.1080/14783360903492694
http://www.informaworld.com

model has been adapted and modified by other researchers who have contributed to the dissemination of the new conceptualisation. Nevertheless, several shortcomings of this conceptualisation have been identified, calling into question the appropriateness of the implementation of this methodology.

This article describes the most important criticism of the Brady and Cronin (2001) model and points out newly identified drawbacks to this kind of conceptualisation and to using the methodology for analysing service quality models. Moreover, to overcome these shortcomings a new procedure is proposed which is based on the proper identification of service quality attributes using a portion of Brady and Cronin's (2001) framework and on the study of unobserved heterogeneity in customer perceptions. An application of this procedure to the insurance industry shows the possible advantages of using this method for studying perceived quality in services. Therefore, the aim of this research is to advance the study of perceived service quality, proposing a method that minimises its limitations.

Literature review

Service quality models

The perception of service quality has been extensively studied during the past three decades. Owing to the intangible, heterogeneous and inseparable nature of services, service quality can be defined as 'the consumer's judgment about a product's overall excellence or superiority' (Zeithaml, 1988, p. 3) or 'the consumer's overall impression of the relative inferiority/superiority of the organisation and its services' (Bitner & Hubbert, 1994, p. 77). Many models have been developed to measure customer perceptions of service quality. The first attempt to measure service quality was based on Gronroos' (1984) service quality paradigm. He distinguished between technical quality, which refers to the outcome of the service performance, and functional quality, which relates to the subjective perception of how the service is delivered. Rust and Oliver (1994) expanded Grönroos' model (1984) by adding a service environment dimension.

Later, Parasuraman et al. (1985) developed the SERVQUAL model which breaks down the notion of service quality into five constructs as follows: tangibles, reliability, responsiveness, assurance and empathy (Parasuraman et al., 1988). SERVQUAL represents service quality as the discrepancy between a customer's expectations for a service offering and the customer's perceptions of the service received.

Despite the fact that SERVQUAL has been the model most widely used and disseminated by academics, it has also been extensively criticised (Buttle, 1996; Carman, 1990; Karatepe et al., 2005). SERVQUAL's weaknesses led to the development of alternative models to measure customer perceptions of service quality. Thus, Cronin and Taylor (1992) developed the SERVPERF model, which is a performance-only measure of service quality. This measure obtained a psychometrically superior assessment of service quality in terms of construct validity and operational efficacy through its performance items and explained more of the variance in an overall measure of service quality than SERVQUAL. Equally, Teas (1993) developed the evaluated performance model (EP) in order to overcome some of the problems associated with the gap in conceptualisation of service quality (Grönroos, 1984; Parasuraman et al., 1985, 1988). This model measures the gap between perceived performance and the ideal amount of a feature, rather than the customers' expectations.

Later, Dabholkar et al. (1996) proposed a multilevel model labelled RSQS (retail service quality scale). In this model, service quality is viewed as a higher-order factor that is defined

by two additional levels of attributes. The first level is made up of five primary dimensions such as physical characteristics, reliability, personal interaction, problem solving and policy; and the second is represented by six sub-dimensions, i.e. appearance, convenience, promises, doing it right, inspiring confidence and courteous/helpful.

Most recently, Brady and Cronin (2001) suggested a hierarchical and multidimensional model. These authors combined the traditional approach of service quality – i.e., the tri-component model of service quality by Rust and Oliver (1994) with the multi-level conceptualisation of service quality (Dabholkar et al., 1996). They described a third-order factor model, in which service quality is explained by three primary dimensions, i.e. interaction quality, physical environment quality and outcome quality. Each of these dimensions consists of three corresponding sub-dimensions such as attitude, behaviour and experience (interaction quality); ambient conditions, design and social factors (physical environment quality); waiting time, tangibles and valence (outcome quality).

The first dimension is 'interaction quality'. The interpersonal interactions have been identified as the employee–customer interface (Hartline & Ferrell, 1996) and the key element in a service exchange (Czepiel et al., 1985). Three distinct factors constitute customer perceptions of interaction quality: employee *attitudes*, *behaviours* and *expertise. Attitudes* allude to willingness to help customers; *behaviours* refer to employees taking actions to address customers' needs; and *expertise* relates to the knowledge of employees about the service.

The second dimension is 'environment quality'. As service is intangible and often requires the customer to be present during the process, the surrounding environment can have a significant influence on perceptions of the overall quality of the service encounter. Their qualitative study revealed that three factors influence the perceived quality of the physical environment: *ambient conditions* pertain to non-visual aspects, such as temperature, scent and music (Bitner, 1992). *Facility design* refers to the layout or architecture of the environment and can be either functional or aesthetic. *Social conditions* refer to the number and type of people evident in the service setting as well as their behaviours (Aubert-Gamet & Cova, 1999; Grove & Fisk, 1997).

The last dimension of service quality is 'outcome quality'. This factor is labelled 'technical quality' by Grönroos (1984, p. 37), who defines it as 'what the consumer is left with when the production process is finished'. Based on their qualitative research, Brady and Cronin (2001) identified three sub-dimensions: *waiting time, tangible* and *valence. Waiting time* refers to the amount of time it takes to provide a service. *Tangible* relates to physical evidence, which is used by customers as a proxy for judging performance (Booms & Bitner, 1981). And finally, *valence* captures attributes that show whether customers believe that service outcome is good or bad, regardless of their evaluation of any other aspect of the experience.

According to these authors, their hierarchical and multidimensional model improves the understanding of three basic issues about service quality: (1) what defines service quality perceptions; (2) how service quality perceptions are formed; and (3) how important it is where the service experience takes place.

Brady and Cronin's (2001) model is superior to previous models at measuring perceived service quality, with the following points standing out:

- It is necessary to develop industry-specific service quality models. There are no universal dimensions/factors/attributes of service quality, although 'personal interaction quality', 'environmental quality' and 'outcome quality' are the key factors of quality that are most likely to be considered by the majority of research contexts.

- It is preferable for the model to be performance based, measured with respect to expectations minus perceptions gap.
- There is different level of abstraction in customer perceptions, from the most aggregate level (overall service quality) to the most disaggregate level (sub-dimensions of quality). This model permits measurement of the flexibility of implementing measures as a function of the level of aggregation required.
- The importance of qualitative research as the first stage in the scale development.

Replications and modifications of Brady and Cronin's (2001) model

Despite the relative novelty of Brady and Cronin's (2001) study, several authors have taken it into account in their investigations, emphasising the hierarchical and multidimensional nature of perceived service quality (e.g. Kim & Jin, 2002; Ko & Pastore, 2004). The impact of Brady and Cronin's conceptualisation has been reflected in some replications/adaptations or modifications of their hierarchical model.

Authors like Lui (2005), Jones (2005) or Fullerton (2005) have adopted the hierarchical conceptualisation for their focused research. Lui (2005) claimed that the concept of perceived service quality is multidimensional and multi-level. This author proposed the same dimensions and sub-dimensions of Brady and Cronin's (2001) quality model in several areas: fast food restaurants, gasoline stations, medical clinic, photo shops, mobile phone repair shops and 24-hour grocery stores. However, he did not test the conceptualisation with a confirmatory method.

On the other hand, Jones (2005) integrated communications with Brady and Cronin's three dimensions of overall service quality to investigate the critical role of this factor in the services industry. Communications quality was divided into two sub-dimensions consisting of interpersonal and impersonal communications quality. The results showed that communications quality was a significant driver of overall service quality in three of the four industry samples (banking, fine dining, physician and sports).

Finally, Fullerton (2005) considered the dimensions of service quality proposed by Brady and Cronin (2001), i.e. interaction quality, physical environmental quality and outcome quality, to develop a model of service quality–loyalty relationships in retail services.

Nevertheless, Brady and Cronin's (2001) conceptualisation has been modified by other authors who have tried to improve the original proposal. Thus, Martínez and Martínez (2007, 2008) have proposed third-order reflective hierarchical models for measuring service quality in different industries (urgent transport services and travel agencies). There are two methodological differences between the original conceptualisation and the modifications made by Martínez and Martínez (2007, 2008): the philosophy of measurement and the nature of causal relationships.

Regarding measurement, Martínez and Martínez (2007, 2008) assume that the proposed dimensions are different forms manifested by perceived service quality. Likewise, these dimensions are explained by their respective sub-dimensions, and finally the indicators are observable representations of each sub-dimension. This view agrees with the taxonomy of reflective multidimensional constructs by Law et al. (1998) and Edwards (2001), and it reproduces the proposal of Dabholkar et al. (1996). Nevertheless, this does not correspond with the Brady and Cronin (2001) perspective of hierarchical structure, in spite of the fact that they base their structure on Dabholkar et al.'s model. Thus, Brady and Cronin (2001) (and the cited authors that have replicated their model) add items to represent the primary dimensions and the overall service quality as an implicit assumption of the dimensions as antecedents of service quality. Martínez and Martínez

(2007, 2008) claim that dimensions are not antecedents of service quality but rather expressions of the complexity of the construct. Service quality is a higher-order factor underlying the dimensions. This view regarding measurement agrees with the entity realism of latent variable theory and the ontological concept of validity (Borsboom et al., 2003, 2004), as opposed to instrumentalism and the formative models (Figure 1). Taking into account that service quality has been defined in a similar way to attitude (Parasuraman et al., 1988), the implementation of measures about attitudes is more consistent with the reflective view. Therefore, changes in attitude towards service excellence have to be reflected in changes in the dimensions, sub-dimensions and observable indicators, following the hierarchy of perceptions developed in different levels of abstraction.

Regarding the nature of causal relationships, Martínez and Martínez (2008) assert that studies that use their proposed dimensions of service quality for predicting an overall evaluation of quality (e.g. Brady & Cronin, 2001; Cronin & Taylor, 1992), a satisfaction judgement (e.g. Cronin & Taylor, 1992; Lam & Zhang, 1999), or a corporate image evaluation (e.g. Bloemer & de Ruyter, 1998) could be open to severe criticism from researchers who highlight the difficulties of inferring causality in cross-sectional designs (Kaplan et al., 2000; Kline, 2006). This problem is described extensively within the context of service quality literature by Martínez (2006a). For example, several authors propose that service quality perception is a determinant of corporate image evaluation (e.g. Andreassen & Lindestad, 1998; Bhattacharya et al., 1995; Ferreira, 1996; Turner, 1982; Weiner, 1985), and other researchers propose just the opposite (e.g. Bitner, 1992; Nguyen & Leblanc, 1998; Selnes, 1993; Smith & Park, 1992). Martínez (2006a) argues that this linkage between these two concepts is mediated by expectations, so that there is a cyclic process in the consumer mind after the first service encounter between customer and firm (Figure 2). This author proposes an approximation to system dynamics to describe the relationships between different attitudes toward the service provided by companies based on the criticism of cross-sectional designs for making causal inferences (e.g. Kaplan et al., 2000; Netemeyer, 2001). If time is divided in k distinct discrete entities, the researcher does not know the exact k moment of time corresponding to the data

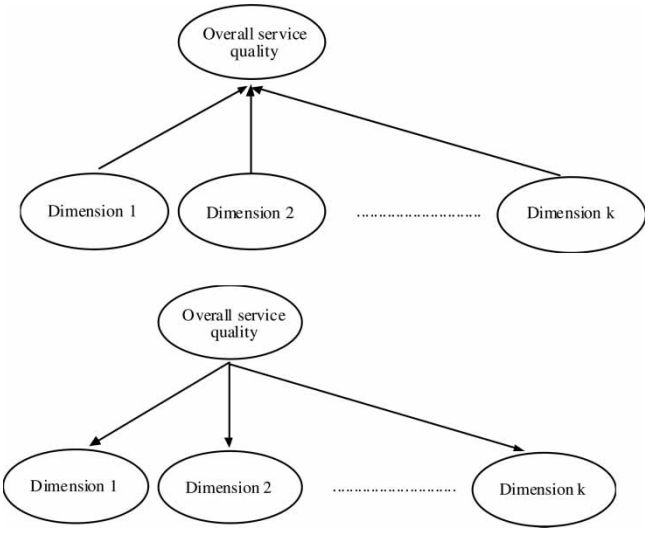

Figure 1. Reflective (a) and formative (b) service quality models.

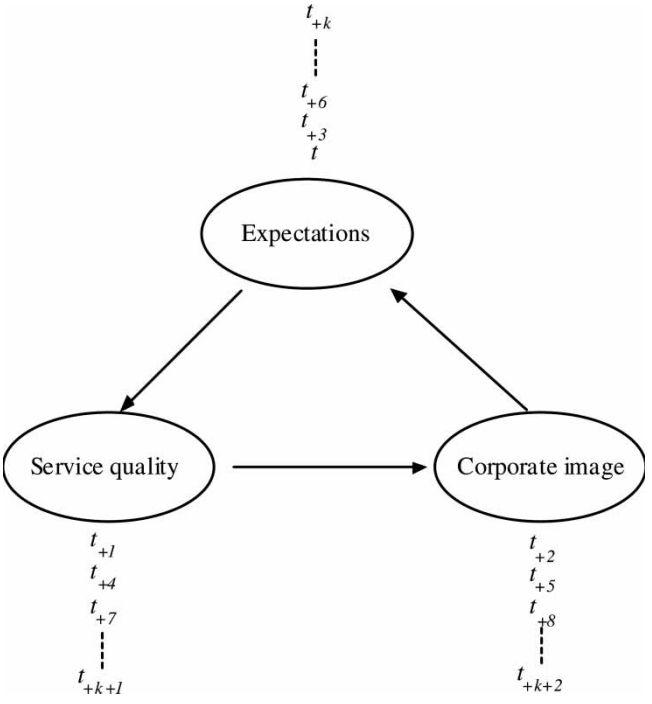

Figure 2. Example of reciprocal relationships between performance measures.
Note: Service quality → Corporate image (Bitner, 1992; Nguyen & Leblanc, 1998; Selnes, 1993; Smith & Park, 1992). Corporate image → Expectations (Bloemer & de Ruyter, 1998; Grönroos, 1984; Oliver, 1997). Expectations → Service quality (Brady & Cronin, 2001; Cronin & Taylor, 1992; Parasuraman et al., 1985, 1988; Zeithaml & Bitner, 1996).

collection. In addition, while some causal processes act almost instantaneously in the consumer mind, other causal mechanisms are subjects to a longer lag period. Therefore, causal relationships between variables that are measured at the same time (cross-sectional designs), representing attitudes or subjective perceptions are not asymmetric but cyclic or non-recursive. This fact invalidates a great part of applied research that proposes causal modelling of relationships between attitude variables and, specifically in the context of service quality models, it severely questions the formative view represented by the majority of service quality models.

Finally, other authors such as Ko and Pastore (2004, 2005) have also modified Brady and Cronin's (2001) conceptualisation using a reflective hierarchical model for measuring service quality in the recreation sport industry. However, these authors do not explain the reason they altered the original proposal nor do they explain the weakness or strength of this modification.

Problems with hierarchical reflective models for measuring service quality

Hierarchical reflective models for measuring service quality are not a panacea, and they are subject to several problems. We are going to describe them in the following paragraphs.

First of all, if we consider service quality a multidimensional construct, it does not exist separately from its dimensions (Edwards, 2001). The challenge is therefore to consider dimensions of quality as separate from overall service quality; we mean to

distinguish the attribute-level performance from global judgements about service as disparate entities, with divergent meaning, antecedents and consequences. Mittal et al. (1998, p. 34) indicate how a customer can evaluate two attributes of a service in a disparate form:

> Mixed feelings toward a product exist because a consumer may be satisfied with one attribute but dissatisfied with another. For example, in a restaurant, a customer may be highly satisfied with the food but highly dissatisfied with the service at the same time.

This reasoning is easily extendable to the service quality concept, because both satisfaction and service quality are attitudes toward the service/company. Thus, if two attributes are very low correlated, the reflective view is open to debate. Could the customer's evaluations of personal interaction be, for example, low correlated with facilities layout or tangible elements in a specific service? In a study about customer evaluation of service performance in the public sport service, Martínez (2006b) found that there was a non-significant correlation among customer perception of personal interaction and two tangible aspects of the service (changing rooms and physical environment hygiene). This evidence reinforces the reasoning of Mittal et al. (1998), and severely questions the hierarchical multidimensional reflective models as a universal representation of the service quality construct.

Second, the utility of multidimensional constructs has been deeply questioned (see Edwards, 2001, or Hayduk et al., 1995), specifically for the operational problems for determining causal relationships between the multidimensional construct and other divergent concepts. Hayduk et al. (1995) made an excellent criticism of the higher-order factor structure in the representation of the ideological basis of attitudes. They indicated how the worth of higher-order factors is purchased at the price of admitting that they make smaller contributions to the behaviour of the observed items. Certainly, this fact is observed in the studies of Martínez and Martínez (2007, 2008) where there are standardised total effects of service quality on some indicators below a value of 0.5. In addition, as Hayduk et al. (1995, p. 482) pointed out: 'what stops the counting of factors?'. Imagine that factors/sub-dimensions/latents in the lower-order level were analysed in a more detailed form. For example, the 'tangibles' sub-dimension in the study of Martínez and Martínez (2008) could be divided into more disaggregated factors like 'ambient conditions' and 'equipment', and these new factors could be considered the lower-order latents (we can label them as 'sub-sub-dimensions'). Consequently, service quality would be a fourth-order concept manifested by dimensions (third-order level), sub-dimensions (second-order level), sub-sub-dimensions (first-order level) and observable indicators. It is quickly noted that this procedure could be repeated and the difficulties of both the conceptualisation and the statistical procedure for empirically testing the model would increase dramatically.

Third, the meaning of the 'valence' sub-dimension questions its inclusion in a multidimensional reflective model. Valence captures attributes that control whether customers believe the service outcome is good or bad, regardless of their evaluation of any other aspect of the experience (Brady & Cronin, 2001). Martínez and Martínez (2008) indicate that the meaning of the valence sub-dimension questions the implementation of service quality studies to fully analyse the company performance. These authors recommend designing questionnaires that also consider those things that are outside companies' control and could have an influence on overall customer evaluation of the service. This draws on the locus-of-causality literature in attribution theory (Oliver, 1993), where the locus dimension is concerned with the source of causality; that is, whether the cause

resides in the customer, in the situation, or in the company. Thus, it is important to implement measures of company performance that inform companies not only about the third source but also about the other two. If we are interested only in the service quality field, it seems clear that 'valence' and other service quality sub-dimensions are distinct entities, because (and in accordance with Brady & Cronin, 2001) the customer may have a positive perception of each service quality dimension, but the negative valence of the outcome can ultimately lead to an unfavourable service experience. Therefore, some factors that shape the valence of the outcome are outside the direct control of service management. This is inconsistent with the features of a multidimensional third-order reflective concept.

Authors such as Patterson and Spreng (1997) or Grace and O'Cass (2005) have defined a similar concept in the context of overall customer satisfaction evaluation. The valence of the outcome is a global evaluation of service performance, similar to a customer satisfaction measure. And perceived quality and satisfaction are not necessarily associated. Customer satisfaction evaluation implies the consideration of aspects such as convenience, price, emotions that do not necessarily have to be considered in the evaluation of service excellence. Some of these aspects may be outside the company's control and cannot be easily handled by management policies. This distinction is easily understood in the context of retail stores: a customer could evaluate the service of a high discount store badly (customer is admitting that there are other types of department stores that provide a very different type of service quality), but at the same time, the customer could be very satisfied, because he/she does not need high levels of quality to fulfil his/her purchasing goals for this specific purchase situation.

And fourth, all service quality models proposed in the literature have been developed not only for academic purposes but for practitioners' use too. Practitioners are mainly interested in using an instrument for making a systematic assessment of service quality, in order to obtain a dynamic view of the evolution of customer service perception. This means that they need valid instruments based on a good theoretical basis, but preferably also with the following features: (1) short questionnaire length; (2) uncomplicated statistical procedure for data analysis. Although the first feature may not be a problem if each distinct sub-dimension is measured by one or two items[1] (Martínez & Martínez, 2008), the second is of concern because each proposed model has to be tested for each particular sample using structural equation modelling. Third-order reflective models do indeed have the particularity that they are very susceptible to non-convergence problems and improper solutions, which makes their application more difficult.

An additional problem: the presence of unobserved heterogeneity

The study of the unobserved heterogeneity has been gaining importance in the social sciences in the past few years (e.g. Allenby et al., 1998; Lubke & Muthén, 2005; Otter et al., 2004). Traditionally, in the context of structural equation modelling (the confirmatory method used to validate the service quality scales), researchers deal with heterogeneity through two main procedures: (1) to establish *a priori* criteria for segmentation (observed heterogeneity) and then to carry out a multi-group analysis; (2) to implement some cluster algorithms that yield different groups of individuals *a posteriori* (unobserved heterogeneity), and then to perform a multi-group analysis.

A priori segmentation has been criticised because it is not capable of correctly explaining the difference in the individual responses (Jedidi et al., 1997). This method is based on non-statistical criteria. In addition, on the majority of occasions, the researcher does not

know the sources of heterogeneity. Furthermore, *a posteriori* segmentation achieved with fuzzy cluster techniques (e.g. k-means cluster) presents several shortcomings (Vermunt & Magidson, 2002): (1) it is not based on a statistical model so that the cluster selection is somewhat arbitrary; (2) it does not permit different scale types for the variables included in the analysis; (3) it does not permit the joint estimation of covariates.

Latent class models and finite mixture models are now used in lieu of traditional cluster algorithms used for *a posteriori* segmentation. These models do not rely on the traditional model assumptions which are often violated in practice (linear relationships, normal distribution, homogeneity) (Vermunt & Magidson, 2002). Particularly interesting is the methodology developed for analysing latent class regression models in the context of structural equation modelling (Arminger & Stein, 1997; Dolan & Van der Maas, 1998; Jedidi et al., 1997). Nevertheless, we do not know of any applications to hierarchical higher-order reflective models; there are excellent works in the literature regarding the applications of this methodology to models with cause–effects relationships (Ansari et al., 2000; Jedidi et al., 1997) but always using first-order latents in the model specification. The studies of Ansari et al. (2000) and Jedidi et al. (1997) show how results can be biased when the assumption of causal homogeneity is violated.

Latent class cluster models are another alternative for dealing with this problem. This type of analysis requires that there are no distinctions between independent and dependent variables (cause–effect relationships). It is based in a model-based clustering approach and it is assumed that the data are generated by a mixture of underlying probability distributions. Therefore, the choice of number of clusters is less arbitrary (Vermunt & Magidson, 2002).

The application of this methodology can be very beneficial in service quality research because it permits the study of the presence of unobserved heterogeneity in non-hierarchy data, and when there is no causal relationship to postulate. Thus, it is applicable to the attributes or factors of service quality that are evaluated by consumers.

The consideration of the mixture of normal distributions when dealing with continuous variables (the most frequent application of Likert-type scales) is a very important strength of this methodology; because when using structural equation modelling, it is commonly supposed that there is a centred normal distribution of underlying data. Taking into account that Likert-type scales are arbitrary scales, and considering that their distribution can be linearly transformed in a normal distribution with zero mean, the general assumption of not modelling the means in the specification of structural equations is not very restrictive under the hypothesis of homogeneity. However assuming this, researchers are saying that, for example, there cannot be two disparate groups of individuals that perceive service quality in a contradictory way, i.e. quality perceptions have to be normally distributed around a unique centred value (high, medium or low). Obviously, this could be highly questionable because of the heterogeneity of customer preferences and attitudes (Allenby et al., 1998). The implementation of advanced structural equation modelling techniques such as multivariate analysis of covariance structure (MACS), multi-level modelling and latent growth modelling take means into account. Nevertheless, in the first and second cases an *a priori* segmentation is needed, and the third case refers to longitudinal designs.

To sum up, latent class cluster analysis performs many useful functions well. It permits the consideration of different groups of individuals with respect to the pattern of responses, is not based on restrictive assumptions regarding the nature of data, is not subject to criticism about the nature of causality and provides practitioners with a valuable picture about the profile of clusters with respect to service quality attributes.

Proposed instrument and methodology

Taking into account: (1) the superiority of Brady and Cronin's (2001) conceptualisation with respect to earlier service quality models (e.g. SERVQUAL, SERVPERF); (2) the modifications made based on multidimensional reflective models that improve the original conceptualisation; (3) the weaknesses of these new proposals; and (4) the role of the unobserved heterogeneity on individuals responses, we propose a simpler form for measuring service quality that overcomes some of the limitations and weaknesses of other conceptualisations.

We propose the following criteria for designing instruments for measuring perceived service quality in each research context and a methodology of analysis:

(1) Building specific context instruments, i.e., a tailored instrument for each specific sector or industry.
(2) Using only performance measures and not a discrepancy function between expectations and performance.
(3) Identifying the attributes or factors of service quality that are susceptible to being evaluated, using qualitative research and literature revision, in the framework of Brady and Cronin's (2001) hierarchical and multidimensional model.
(4) Implementing measures of 'valence' and global service quality evaluation in order to compare these measures with the service factors evaluated.
(5) Achieving a latent class cluster analysis, in order to identify possible groups with a disparate pattern of response with respect to relevant variables, and profiling the clusters with appropriate additional covariates.

We are going to put into practice this sequence of analysis with a real investigation in the insurance industry.

Service quality in the insurance industry

According to Sachdev and Verma (2004) insurance companies are making strides in raising awareness levels, introducing innovative products and increasing the penetration of the market. Some insurers have fulfilled their mission to be a scale player in the mass market by introducing a range of products to meet the needs of each customer in the effort to serve them better. Many others have taken a more focused approach, introducing select products that they believe hold potential and fill market gaps (Sharma, 2002). Also, technology (websites, email, etc.) has enabled the customer to get in touch with insurance companies quickly, easily and directly, but in turn it will enable customers to make comparisons quickly and accurately (Sachdev & Verma, 2004). Because of this situation, the improvement of the quality has become the key differentiating factor in this sector, as much for ethical reasons as for the belief that a good service constitutes the most effective way to get loyal clients.[2]

In insurance companies, improving quality has become an essential aspect in this industry and has a major effect on differentiation. However, there are few studies in the literature focused on the conceptualisation and measurement of service quality perception.

An example is the study carried out by Parasuraman et al. (1991). The purpose of this investigation was to refine and study the reliability and validity of the SERVQUAL model which was developed by these authors in 1988. This service quality model was pre-tested and the results suggested several refinements. After these modifications, the new model, the *revised SERVQUAL* is applied in five different customer samples (one telephone company, two insurance companies and two banks). The revised model was based on

the five dimensions of original SERVQUAL (1988) which are measured by 22 items (these are the original items, modified as above mentioned). The factor analysis showed the five-dimensional structure of the original scale, with one key exception, namely, the dichotomisation of tangibles into two sub-dimensions (one pertaining to physical facilities/equipment and another pertaining to employees/communication materials). The results obtained led the authors to reaffirm the utility of the scale and the method of measurement. However, the authors did not analyse the differences among the different industries studied, although they recognised that minor modifications in the wording of items to adopt them to specific settings were appropriate.

Later, Parasuraman et al. (1994) again used the insurance companies in one of their studies. The primary objective of this study was compare alternative service quality measurement scales on psychometric and diagnostic criteria to address the unresolved methodological issues. These authors established three alternative questionnaire formats, one incorporating the difference-score formulation and the other two incorporating direct measures (one of them generates direct ratings of the service superiority and service adequacy gaps; the other generates direct ratings of the perceptions). All these formats contained the 22 items for measuring service quality over the five dimensions of revised SERVQUAL (Parasuraman et al., 1991). After several minor modifications were made to the scale items, each of the three formats was evaluated by 800 customers in four different sectors: retail chain, computer manufacturer, auto insurer and life insurer. However, the results showed that all three questionnaire formats possess high psychometric properties; the authors consider that the companies should adopt a service quality measurement system that produces separate measures of adequate-service and desired-service expectations and perceptions.

The insurance industry has also been used by Zeithaml et al. (1996). Through an empirical study these authors examined relationships between service quality and customers' behavioural intentions in several industries (retail chain, computer manufacturer, automobile insurer and life insurer). Several measures of service quality were used: (1) an overall single-item; (2) a multiple-item scale of perceived service from the revised SERVQUAL (1991) representing the five SERVQUAL dimensions; (3) two categorical questions to measure whether respondents had experienced a recent service problem with the company and, if so, whether the problem was resolved to their satisfaction. Also, respondents were asked to indicate separate ratings of their adequate, desired and perceived service. However, contrary to previous studies, service quality was operationalised as a weighted-average perceived performance score across the dimensions; the results showed that service quality is associated positively with favourable behavioural intentions and negatively with unfavourable behavioural intentions. With few exceptions, the better a company's service quality scores are, the higher its mean scores are for loyalty and pay, and the lower they are for switching and external response. However, this study does not analyse the intercompany differences in the service quality.

In the same context, Sachdev and Verma (2004) developed an empiric investigation whose objective was to explore relative importance of service quality dimensions across a 'select' service context: (1) intangible-dominant (banking and insurance); (2) tangible-dominant (fast food and beauty salon). Just as occurred in Zeithaml et al. (1996) study, in Sachdev and Verma (2004) service quality was operationalised using perceptions only. Also, from adequate service and desired service a zone of tolerance was built which was used to determine the relative importance of service quality dimensions. Service quality was measured in different ways: (1) the 22-item scale and five-dimension structure proposed by Parasuraman et al. (1988); (2) an overall service quality single-item (to reduce

shared-method variance). Also, the respondents were asked to distribute 100 points over the five dimensions considering their relative importance. The results showed that in the case of insurance the mean points allocated in the order of importance are: reliability, assurance, responsiveness, tangibility and empathy.

Other research concerning insurance companies was elaborated by Lin et al. (2001). This study examines the relationship between the personality of the service providers and perceived service quality. To conceptualise service quality, a version of SERVQUAL (Parasuraman et al., 1994) was used; however, taking into account the objective of this research, a tangible dimension was deleted because of the very slight relationship between this dimension and the employee's personality. Just as in previous studies, based on the recommendations of Parasuraman et al. (1994), Lin et al. (2001) used a perceptions-only score rather than expectations minus perceptions gap to measure service quality. The data were collected from four different service sectors in Taiwan: life insurance, real estate agencies, information services and securities. The overall results showed that openness is a valid predictor of assurance, conscientiousness correlates with reliability, extraversion is positively correlated with responsiveness, and agreeableness is a valid predictor of assurance and empathy. However, this study does not go more deeply into the impact or the significance of service quality dimensions in the different sectors used.

To sum up, in insurance industry literature, service quality has been conceptualised in different ways. A common thread among most studies is that none of them applies a specific model to measure the quality perception in insurance companies. These studies were multi-sectorial and applied the SERVQUAL scale, in its different versions (Parasuraman et al., 1991, 1994); most of them only used the perceptions scale (Lin et al., 2001; Sachdev & Verma, 2004; Zeithaml et al., 1996). This perspective contrasts with the arguments of other authors who consider that service quality dimensions are not universal (Babakus & Mangold, 1992; Buttle, 1996; Carman, 1990). According to Babakus and Boller (1992), there is a need to develop measures of service quality that are industry-specific, since the use of generic models (such as SERVQUAL or SERVPERF) forces researchers to drastically alter the items chosen to apply these models in a specific context (Babakus & Boller, 1992; Babakus & Mangold, 1992; Carman, 1990; McAlexander et al., 1994). Moreover, there is a need to develop service quality measures that are country/culture specific, since the measures that are developed in one culture may not be applicable in a different cultural setting (Karatepe et al., 2005; Mattila, 1999). According to Mattila (1999), cultural factors are said to have greater influence on people's evaluation of services than on their evaluations of physical goods because of the involvement of customer contact and interaction with employees while a service is delivered. Also, Ueltschy and Krampf (2001) considered that differences in culture affect the measure of quality in a service sector.

However, the studies developed in the context of insurance companies have the weakness that none of them used a qualitative research that allowed identification of service quality dimensions. These dimensions could be critical for consumer evaluations, and therefore they could overcome the definition of service quality provided by SERVQUAL or SERVPERF models.

Consequently, a deeper and more rigorous study of this service is necessary, for the purpose of obtaining a specific scale that shows the relevant features of the insurance service. Hence, we have adopted Brady and Cronin's (2001) methodology to identify the key attributes of service quality, and then we have implemented the modifications

proposed in this research: (1) to separate the key attributes from measures of valence and global service quality; and (2) to analyse the data using latent class cluster analysis.

Identification of the key attributes

Qualitative research

The number and the nature of service quality dimensions are directly related to the service under investigation. Qualitative research was carried out to identify the factors that determine the service quality perceptions of insurance customers.

For qualitative research, we used in-depth interviews of insurance companies' employees and customers. In order to obtain information from the insurance companies' point of view, five managers were interviewed. The managers were interviewed about several service quality issues: (1) what they perceived to be service quality from the consumer's perspective; (2) what steps they took to control or improve service quality; (3) what problems they faced in delivering high quality services (Parasuraman et al., 1985). With regard to the customers' side, 10 customers who had used this service one or more times within the previous 12 months were interviewed (Ryan & Cliff, 1997). Following Brady and Cronin (2001), the respondents were encouraged to list all factors that influenced their perception according to their experience.

To code the qualitative data thus obtained, similar to prior studies (Brady & Cronin, 2001; Karatepe et al., 2005; Richins, 1997), a content analytic approach was employed. In order to ensure the validity of this research, a semi-structured questionnaire was used in the interviews. This was formulated from the emerged framework of references (insurance service quality literature and Brady and Cronin's (2001) study) in order to answer to research problems and questions. Furthermore, in order to ensure the reliability of this research, a tape recorder was also used during the interviews. This method provided us with the opportunity to double check the answers and to ensure that we did not miss any important information.

Carman (1990) suggests context customisation may involve adding or dropping relevant dimensions and appropriately adjusting the attributes considered to make those dimensions. Therefore, after listening to the tapes, three independent coders prepared field notes based on Brady and Cronin's (2001) study and on the rest of the literature about insurance service quality. Then, taking all that information into account, we proposed six quality factors for insurance companies.

Proposed attributes

After combining the findings from our qualitative research with the service quality literature revision we propose the following attributes or factors of service quality: conduct, expertise, problem solving, tangible evidence, information and waiting time. In addition, it is necessary to implement measures of valence and global service quality. We are going to explain the selection of each separate attribute considered from the standpoint of the hierarchical conceptualisation of service quality. Therefore, depending on the level of aggregation, we can propose a more or lesser detailed battery of attributes. In this case, we start with the dimensions of service quality (higher-order level of abstraction) by proposing the attributes or factors of service quality that can be measured.

The first dimension is 'personal interaction'. Several researchers have indicated the importance of this factor in the delivery of services and have identified it as having the most significant effect on service quality perceptions (Bigné et al., 1996; Grönroos,

1982; LeBlanc, 1992). The contact of the insurer with the client is fundamental. The insurer is present at all times in the life of the insurance policy. Before the contract is signed, he/she advises the client about the conditions of the policy or of the service; then during the time it is in effect, he/she revises the rates and the documentation, and he/she holds consultations or resolves doubts, and finally when the situation insured for takes place, he/she assists with the claim and ensures prompt resolution. Both the service quality literature and our qualitative research suggest that there are three more detailed facets or attributes that can be considered as a part of 'personal interaction': *conduct*, *expertise* and *problem solving*. The first two factors coincide with Capelleras and Veciana's (2002) proposal. According to Grönroos (1990), attitudes, behaviour and skills of employees are factors that largely influence service quality perceptions. Based on our qualitative study and literature review, we identified a third attribute, problem solving. The results of this qualitative study reveal that both insurers and users pointed out that when an incident occurs it is very frequent for conflicts to exist and their resolution is very important. Including problem solving in 'personal interaction' is endorsed by the research carried out by other investigators, such as Westbrook (1981) and Siu and Chow (2004). According to Chan (1999) complaints reflect service quality to the extent that consumers make complaints when the quality or quantity of service delivered by the insurance falls short of certain levels of their expectations.

The second dimension is 'physical environment'. This factor is considered as one of the most important and influential in the evaluation of service quality (Bitner, 1990, 1992; Brady, 1997; Howat et al., 1996). Taking into account the existing literature, as well as the results of the qualitative study, we adapted Dabholkar et al.'s (1996) proposal to the specifics of the service offered by insurance companies. So *appearance* was labelled *tangible evidence*, and *convenience* was renamed *information*. On the one hand, the qualitative research reveals that having modern equipment or using advanced technology are outstanding attributes for both the companies and the customers. On the other hand, it showed that the consumers demand exact and precise information about all conditions of the policy throughout the whole process in order to avoid misunderstandings in the event that the incident occurs. The *information* factor was used by other authors in the environment of perception of service quality, such as Howat et al. (1996) and Ko and Pastore (2005). According to Ponce et al. (1993), information and communication are key factors of insurance services.

The last dimension proposed is 'outcome'. This dimension is centred on the consequence of the act of the service, and it indicates what the consumer receives when the productive process and the buyer–salesperson interactions have concluded (Grönroos, 1984). *Waiting time* is the factor that defines 'outcome'. Because in this service what is hired is a certain type of coverage, whether for robbery, catastrophe or fire, it may be that it is never used, but if the incident occurs, the time needed to resolve it should be minimal. This factor has been considerably supported in the literature, for example, in Brady and Cronin (2001) and Taylor and Claxton (1994).

Selection of the items and measures

The creation of a list of items was developed adapting the items of existing generic scales (e.g. Brady & Cronin, 2001; Dabholkar et al., 1996; Parasuraman et al., 1988) and specific insurance scales (Bigné et al., 1996; LeBlanc, 1992; Parasuraman et al., 1991, 1996; Ryan & Cliff, 1997). On the basis of our literature review, we generated a pool of 30 items using a performance-only measure, which were assessed by a panel of experts. These items were

indicators of each attribute. A five-point Likert scale was applied to measure the different items on a scale from strongly disagree to strongly agree.

We selected the 'gold standard indicator' for each concept from the set of items (Hayduk, 1996). Our view of each attribute (considering attributes as latent variables) determines the proximity between that abstract concept and the real world manifestation. Thus, the definition of the theoretical attribute determines the choice of the best indicator. Appendix A shows the selected items.

Brady and Cronin (2001) considered *valence* as a facet of 'outcome'. However, we have explained how the meaning of this factor is different from the others service quality attributes. The evaluation of *valence* is very interesting because it can provide a different perspective about service performance. Providing a high level of service quality may not be enough to satisfy customers.

Additionally, we have included a measure of *overall service quality* (Babakus & Boller, 1992; Bolton & Drew, 1991), based on the following set of reasons: (1) to check if global evaluation of service quality can be different from specific evaluations of attributes; (2) to analyse the degree of robustness of overall service quality evaluation if some attributes are poorly rated; (3) to compare it with the customer's judgement of valence.

Moreover, we included covariates for the analysis. These covariates act as exogenous variables in the analysis but their inclusion is less restrictive than in other confirmatory methodologies like structural equation modelling, because no *a priori* causal hypothesis is tested. Therefore, three variables were considered as active covariates: possession of ISO certification by the companies, customer experience with the evaluated service and customer experience with other companies.

ISO certification can be used by companies to improve their quality system management. If this quality system is successfully adopted, customers may evaluate more positively the service provided by certified companies (Martínez, 2006c).

Customer experience with the company is another relevant variable in the context of service evaluation (Rust et al., 1999). These authors show how companies should pay more attention to the quality delivered to customers with a low degree of experience with the service because they are more sensitive to what they perceive as risk associated with the brand. In addition, with repeated interactions, firms and customers develop bonds, and customers with positive experiences over time are more forgiving and less likely to defect (Anderson & Sullivan, 1993).

Finally, experience with other companies is considered an important variable because it plays a fundamental role in the formation of normative expectations regarding service performance (Ekinci, 2003). Expertise provides customers with information of competing alternatives (East, 1992) and experts have a more developed and complex cognitive structure than novices (Alba & Hutchinson, 1987).

In addition, socio-demographic variables were considered as inactive covariates in order to help in the characterisation of the distinct clusters of customers. Sex, age and income can be important variables in the context of customer evaluation (Cooil et al., 2007).

Data collection

A sample was collected from users of insurance companies in Spain. The procedure to obtain the data was a personal interview with the individuals and a self-administered questionnaire. We trained business and marketing students for this purpose. Respondents had

to fulfil the condition of having used the service within the previous 12 months (Ryan & Cliff, 1997). An initial convenience sample of 221 respondents was obtained, although 17 questionnaires related to online insurance were dropped. According to service quality literature, to measure online service quality it is necessary to develop a specific scale for these types of services (Grönroos et al., 2000; Liljander et al., 2002; Zeithaml et al., 2002). The final sample included 204 subjects.

Results

Latent class cluster analysis was carried out using Latent GOLD 4.0 (Vermunt & Magidson, 2005). We followed the model estimation methodology suggested by these authors. We began with the sequential analysis of disparate latent class models. The Bayesian information criterion based on the log-likelihood (*BIC-LL*), the classification error index and the average weight of evidence (*AWE*) were used as the criteria for model selection (Vermunt & Magidson, 2005). The lower *BIC*, classification error and *AWE*, the better the model.

Table 1 shows the statistics of the different estimated models. It is not uncommon that the various statistical criteria do not always agree with each other, thus there is room for a subjective element in determining the number of classes. In this case, M_4 was the final model selected because this model has the second minor *AWE*, the *BIC-LL* is not widely distant from the lesser value, because it presents the lowest index of classification error among models with more than one latent class, and the distinct latent classes do not have relatively small sizes.

Given that M_5 yielded a better fit, a conditional bootstrap analysis was performed between M_4 and M_5; their significant results would support the choice of M_5. Nevertheless, the interpretation of M_5 was not as clear as M_4, being two latent classes practically equivalent in their profile. Taking into account that indicators showed variances could be a bit inflated because of measurement error, we thought that a clear distinction would be necessary in the profile of latent classes in order to ensure that confidence intervals around mean were as separate as possible. In addition, the study of bivariate residuals indicated that only seven of the 28 residuals for M_4 were larger than the usual threshold of 3.84 (Vermunt & Magidson, 2005). Although the estimated model cannot adequately explain these seven associations among variables at a significance level of 0.05, the deviation from the threshold value was not severe.

For M_4, all log-linear parameters are significant for all indicators. Thus, knowledge for the response in each indicator contributes in a significant way to the ability to discriminate between clusters. R^2 indicates how much the variance of each indicator is explained by this

Table 1. Estimated models and statistics.

		BIC-LL	Number of parameters	Classification error	*AWE*
M_1	1 cluster	4392.673	16	0.000	4525.526
M_2	2 clusters	3814.875	36	0.019	4133.165
M_3	3 clusters	3455.368	56	0.001	3929.168
M_4	4 clusters	3138.584	76	0.002	3772.981
M_5	5 clusters	2915.297	96	0.003	3714.329
M_6	6 clusters	2969.556	116	0.002	3936.818
M_7	7 clusters	2973.969	136	0.021	4125.911

Table 2. Parameters of four-cluster model.

Indicators	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Wald	*p*-value	R^2
Conduct	0.332	−0.295	−1.118	1.081	156.03	<0.001	0.457
Expertise	0.392	−0.462	−0.939	1.009	175.16	<0.001	0.464
Problem solving	0.386	−0.28	−1.280	1.172	200.78	<0.001	0.534
Tangibles	0.233	−0.370	−0.560	0.704	59.78	<0.001	0.238
Information	0.218	−0.243	−0.829	0.855	90.62	<0.001	0.297
Punctuality	0.407	−0.418	−0.781	0.793	114.22	<0.001	0.382
Valence	0.478	−0.428	−1.350	1.301	454.53	<0.001	0.714
Overall quality	0.542	−0.457	−1.628	1.542	15466.52	<0.001	0.955
Covariates							
ISO certification	−0.458	0.499	0.563	−0.605	12.088	0.007	
Experience with the company	0.007	0.095	−0.097	−0.005	2.382	0.500	
Experience with other companies	0.138	0.000	−0.160	0.023	0.513	0.920	
Cluster size	12%	41%	31%	16%			

four-cluster model, being an index of effect size, i.e. the larger R^2, the better the discriminating power (Table 2).

As can be seen in Table 2, the evaluation of 'valence' and 'global service quality' are the variables that most contribute to the cluster discrimination, as the most important indicators of individual differences among groups. Moreover, only the 'possession of ISO certification' has a significant associated Wald statistic, as it is the unique covariate that yields differences among latent classes.

Once the significance of parameters estimates is analysed, the next step is cluster characterisation. Figure 3 details the profile of the relevant variables of our investigation

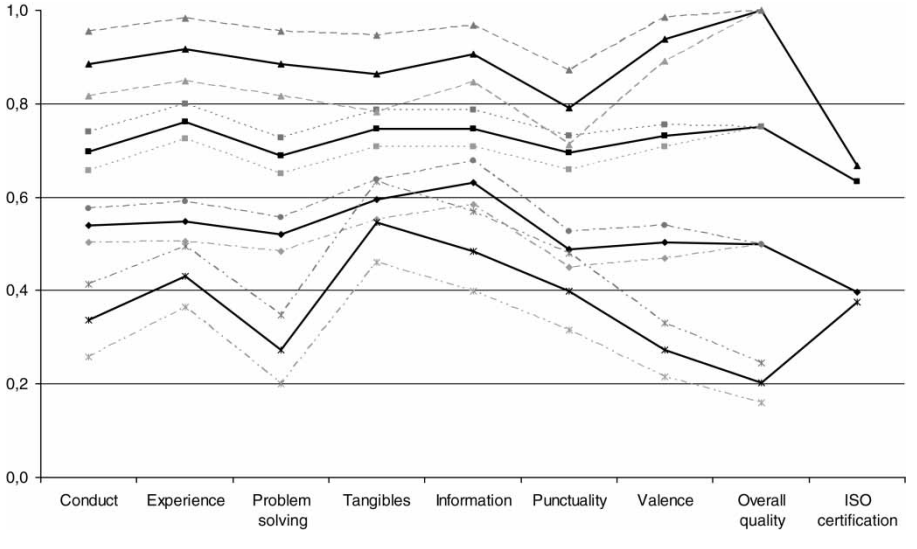

Figure 3. Profile of the latent classes.
Note: Dashed arrows represent 95% confidence interval around the mean. Clusters are ordered by levels of perceived service quality.

in the four latent classes. Following Cohen et al. (1999), we have converted all measures within the closed range (0, 1).

The first cluster is composed of 12% of the cases, and individuals within this group perceived very high levels of service quality for all attributes. Cluster 2 is the most numerous (41% of the cases) group and is also distinguished by high levels of perceived service quality but with significantly fewer values than Cluster 1. Cluster 3 (31%) presents medium values for all indicators. Finally, 16% of individuals (Cluster 4) evaluate service performance significantly below the medium level in several indicators.

The interpretation of the role of the unique relevant covariate is straightforward; around two of each three individuals that have evaluated certified companies are in the two first clusters, i.e., they perceive high levels of service quality; therefore, ISO certification could be a determinant of differences in service quality evaluations.

On the other hand, there are no outstanding differences in the profile of the clusters when considering the rest of the inactive covariates. Latent classes are not differentiated by sex, age or income.

Inter-class differentiation is clear because confidence intervals do not overlap for the majority of the indicators. However, the interpretation of intra-class differences cannot provide very much added value owing to the low sample sizes of clusters. Considering that valence and overall service quality are the more discriminating indicators, we have computed the Cohen's *d* effect size (Cohen, 1988) to compare the differences with respect to the proposed attributes of service quality. We have followed the methodological procedure proposed by Dunlap et al. (1996) for this purpose, adding 95% confidence intervals for *d* (Smithson, 2003). In addition, we have created a new variable: 'attributes mean', that it is an arithmetic mean of attributes. This new variable permits a comparison with respect to valence and overall service quality. Table 3 shows these results.

Although most of the effect size confidence intervals overlap, we can highlight that (1) overall service quality and valence are distinctly perceived (non-zero effect sizes) in all clusters; (2) the mean of the evaluated attributes does not coincide with the judgements of valence and overall service quality in spite of the broad width of the confidence intervals around the effect size index.

Discussion and implications

We have proposed a new form for measuring quality in services based on a criticism of the most important service quality models. Specifically, we have centred the main points on the most important contribution in recent years, Brady and Cronin's (2001) hierarchical and multidimensional model, and we have applied our proposal to an empirical study in the insurance industry. Several issues deserve special attention.

First, customer perceptions, preferences and attitudes are heterogeneous. The assumption of homogeneity can draw a distorted picture of the customer evaluations of the service. The application of a methodology like latent class cluster analysis can help to better understand this heterogeneity, as our empirical study has showed.

Second, attributes/factors/dimensions of service quality are not universal but industry specific. However, the efforts made by Brady and Cronin (2001) to develop a comprehensive framework for analysis have to be taken into account. Model development designed by these authors can be an excellent basis for proposing the attributes of service quality that can be measured. Qualitative research and focused literature revision are necessary for this purpose.

Table 3. Intra-class effect sizes.

Indicators	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Valence	Overall quality	Valence	Overall quality	Valence	Overall quality	Valence	Overall quality
Conduct	−0.368 (−0.88; 0.15)	−0.985 (−1.60; −0.34)	−0.216 (−0.07; 0.51)	−0.391 (0.03; 0.65)	0.253 (−0.56; 0.07)	0.391 (−0.80; −0.08)	0.327 (−0.14; 0.79)	0.767 (0.21; 1.30)
Expertise	−0.152 (−0.71; 0.41)	−0.740 (−1.33; −0.12)	0.210 (−0.05; 0.55)	0.101 (0.04; 0.66)	0.284 (−0.07; 0.55)	0.403 (−0.23; 0.46)	0.911 (0.36; 1.44)	1475 (0.86; 2.07)
Problem solving	−0.365 (−0.82; −0.12)	−0.985 (−1.60; −0.34)	−0.288 (−0.16; 0.36)	−0.496 (−0.13; 0.47)	0.112 (−0.62; −0.03)	0.195 (−0.92; −0.20)	0.000 (−0.37; 0.37)	0.399 (0.02; 0.76)
Tangibles	−0.457 (−1.00; 0.10)	−0.983 (−1.60; −0.34)	0.102 (0.20; 0.81)	−0.024 (0.36; 1.01)	0.587 (−0.19; 0.42)	0.790 (−0.37; 0.32)	1.334 (0.74; 1.91)	1.822 (1.13; 2.49)
Information	−0.243 (−0.77; 0.29)	−0.921 (−1.53; −0.28)	0.099 (0.37; 0.98)	−0.024 (0.54; 1.21)	0.777 (−0.17; 0.40)	1.005 (−0.37; 0.32)	1.035 (0.51; 1.52)	1.468 (0.88; 2.30)
Punctuality	−0.900 (−1.40; −0.37)	−1548 (−2.24; −0.80)	−0.263 (−0.36; 0.17)	−0.479 (−0.40; 0.20)	−0.109 (−0.62; 0.02)	−0.111 (−0.90; −0.18)	0.628 (0.17; 1.06)	1013 (0.55; 1.45)
Valence		− 0.799 (−1.39; −0.17)		−0.236 (−0.27; 0.34)		0.041 (−0.62; 0.08)		0.475 (0.17; 0.76)
Overall quality	− 0.799 (−1.39; −0.17)		−0.236 (−0.27; 0.34)		0.041 (−0.62; 0.08)		0.475 (0.17; 0.76)	
Attributes mean	−0.545 (−1.03; −0.02)	−1.492 (−2.19; −0.77)	−0.433 (−0.15; 0.32)	−0.837 (0.05; 0.67)	0.092 (−0.84; −0.13)	0.419 (−1.33; −0.56)	−0.975 (−1.44; −0.49)	−1.780 (−2.37; −1.16)

Third, valence cannot be considered in the same way as other service quality attributes, because its definition is not concordant with the other attributes. Attributes of service quality can be handled by the quality management system. However, the definition of valence is close to the concept of satisfaction and, as we have commented, quality and satisfaction are not necessarily associated. The empirical study shows non-zero effect sizes between valence and the quality attributes, which underlines the conceptual differences between them.

Fourth, it is advisable to include a measure of overall service quality in order to compare this evaluation with the rest of the indicators. This gives somewhat surprising empirical results because it hints at important divergences in the most separate clusters. For customers that perceive a very high level of quality attributes (Cluster 1), overall service quality is evaluated even higher. Just the opposite occurs with customers that perceive the lowest level of service quality (Cluster 4). This is important because customers are making a clear distinction between attributes and an overall evaluation of quality, pointing to some interesting issues that need examination: (1) a halo effect may exist: the evaluation of attributes can be affected by an extreme (high or low) overall evaluation. For example, a very good or very bad service experience could lead to an extreme assessment of overall quality, and this can affect (deservedly or not) the evaluation of attributes at that point of data collection; (2) quality attributes and overall quality are disparate concepts for these two clusters. An explanation is derived if it is assumed that customers think differently when evaluating specific attributes rather than making an overall judgement, i.e. overall judgements are more susceptible to be 'inflated' or 'deflated' by general feelings toward the service (e.g. corporate image, emotions, satisfaction, trust) beyond the mere evaluation of specific factors. The nearness of evaluations of valence and overall service quality could reinforce this last idea.

And fifth, only one covariate, the possession of ISO certification, presents a significant different profile across clusters. These latent classes do not discriminate between the other considered covariates. This complicates the characterisation of the different groups of customers. However, knowing that the majority of the customers that perceive high levels of quality (Clusters 1 and 2) have evaluated certified companies, it is a very good starting point for further research on the effects of management policies like the ISO certification on some intangible firms results (customer perceptions).

On the other hand, there are some implications for management.

First, we have presented a detailed methodology for assessing perceived service quality in a specific service. Practitioners can follow the sequence of analysis detailed in this article for their focused applications.

Second, the identification of disparate clusters of customers is a valuable strategic tool for management and marketing policies. Although our empirical study has been carried out with multiple firms, disparate segments can appear within the customers of a single company. In this case, the characterisation of these clusters plays a more important role, and probably differences exist in some covariates that could have been hidden by the multi-firm data collection.

Third, the feasibility of use of this methodology by non-academic people overcomes the problems associated with a more complicated program syntax tool such as structural equation modelling. Obviously, however, some minimum grounding in statistics and the use of the appropriate software for data analysis is necessary.

And fourth, the questionnaires should not be excessively lengthy, as we have showed in our empirical study. Nevertheless, this requires a careful selection of attributes and

items in order to mitigate possible problems associated with low reliability. Short questionnaires are highly desirable for reducing costs.

Other alternatives

Although this research has focused on the most important proposals for measuring perceived service quality, there are other interesting alternatives, including ones that have not been traditionally considered 'service quality models': importance–performance analysis (IPA) and fuzzy logic.

Importance–performance analysis (IPA) is a simple and useful technique for identifying those attributes of a product or service that are most in need of improvement or that are candidates for possible cost-saving conditions without significant detriment to overall quality. A two-dimensional IPA grid displays the results of the evaluation about the importance and performance of each relevant attribute (Abalo et al., 2007). A weighted combination of these attributes forms the service quality index. This method requires direct customer evaluation of the importance of each attribute and has been the subject of some criticism (Abalo et al., 2007). In our view, if the procedure for identifying the key attributes of service quality is based on Brady and Cronin's (2001) framework, IPA would overcome several of the statistical and methodological limitations that we have criticised. Nevertheless, two important topics for further debate arise: (1) an analysis of unobserved heterogeneity would be desirable; (2) the assumption that a service quality index is a fully deterministic weighted combination of attributes could be contrasted with the divergences found in our investigation among attributes and overall service quality.

Fuzzy logic is a technique for reasoning under the constraints of uncertainty but with different principles with respect to the probability theory and represents 'degrees of truth' taking precedence over the traditional 'true–false' logic. Fuzzy logic incorporates qualitative linguistic information to determine the importance of quality attributes. Like IPA, fuzzy logic could profit from the use of Brady and Cronin's (2001) framework for selecting the attributes and could elicit the same possible debate on the two aforementioned topics. Lozano and Fuentes (2003) is an interesting source of further information about the application of this technique.

Limitations and further research

The major shortcoming of this study is the low sample size collected, which is insufficient to clearly distinguish some differences among intra-class indicators. Marketing segmentation studies usually deal with higher sample sizes that permit a more representative picture of the population under consideration. Taking into account the decreasing costs associated with short questionnaire length, firms have an incentive to use this methodology with larger samples of customers.

In addition, although we have used a careful method for selecting the indicators (qualitative research, literature revision and experts' evaluation), reliability is not explicitly considered in the statistical analysis. Reliability does not affect means, but does affect variances and correlations. Nevertheless, we stress that this careful method for selecting indicators enhances content validity and reliability, and at the very least diminishes the non-desired effects of measurement error. With higher sizes of clusters, researchers could make a more robust test of their conceptualisation, achieving a multicluster confirmatory factor analysis adding a 'second best indicator' (Hayduk & Glaser,

2000) per attribute. Then, if the multi-cluster analysis fits, measurement error would not be an important issue for the robustness of the obtained results.

We also emphasise the systematic utilisation of service quality questionnaires by firms to obtain a dynamic picture of evaluations over time, with the aim of enriching the dynamic view of customer attitudes towards service (Johnson et al., 2006; Mittal et al., 2001). The evolution of clusters over time is an interesting point for further research in order to assess the effect of quality management and marketing policies on the customer's perceptions.

Notes

1. See Hayduk (1996) and Hayduk and Glaser (2000) for an in-depth explanation of the use of gold standard indicators and the non-necessity of at least four indicators per latent, as usually recommended when authors handle latent variables. In addition, the recent research of Bergkvist and Rossiter (2007) claims that for the many constructs in marketing that consist of a concrete singular object and a concrete attribute, single-item measures should be used. Thus if we considerer service quality factors as concrete singular attributes, the single-item measure would report advantages.
2. These are the managerial principles of MAPFRE insurance company.

References

Abalo, J., Varela, J., & Manzano, V. (2007). Importance values for importance–performance analysis: A formula for spreading out values derived from preference rankings. *Journal of Business Research*, *60,* 115–121.

Alba, J.W., & Hutchinson, J.W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, *13,* 411–454.

Allenby, G.M., Arora, N., & Ginter, J.L. (1998). On the heterogeneity of demand. *Journal of Marketing Research*, *35,* 384–389.

Anderson, E.W., & Sullivan, M.W. (1993). The antecedents and consequences of customer satisfaction. *Marketing Science, 12*, 125–143.

Andreassen, T.W., & Lindestad, B. (1998). Customer loyalty and complex services: The impact of corporative image on quality, customer satisfaction and loyalty for customer with varying degrees of service expertise. *International Journal of Service Industry Management*, *9,* 7–23.

Ansari, A., Jedidi, K., & Jagpal, S. (2000). A hierarchical Bayesian methodology for treating heterogeneity in structural equation models. *Marketing Science*, *19,* 328–347.

Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors: Log-likelihood function, minimum distance estimation, fit indices, and a complex example. *Sociological Methods and Research*, *26,* 148–182.

Aubert-Gamet, V., & Cova, B. (1999). Servicescapes: From modern non-places to postmodern commonplaces. *Journal of Business Research*, *44,* 37–45.

Babakus, E., & Boller, G.W. (1992). An empirical assessment of the SERVQUAL scale. *Journal of Business Research*, *24,* 235–268.

Babakus, E., & Mangold, W.G. (1992). Adapting the SERVQUAL scale to hospital services: An empirical investigation. *Health Services Research*, *26,* 767–786.

Bergkvist, L., & Rossiter, J.R. (2007). The predictive validity of multiple-item vs single item measures of the same construct. *Journal of Marketing Research*, *44,* 175–184.

Bhattacharya, C.B., Rao, H., & Glynn, M.A. (1995). Understanding the bond of identification: An investigation of its correlates among art museum members. *Journal of Marketing*, *59*(4), 46–57.

Bigné, J.E., Martínez, C., M. Miquel, J., & Belloch, A. (1996). La calidad de servicio en las agencias de viaje: Una adaptación de la escala SERVQUAL. *Revista Europea de Dirección y Economía de la empresa*, *5*(2), 7–18.

Bitner, M.J. (1990). Evaluating service encounters: The effects of physical surrounding and employees responses. *Journal of Marketing*, *54*(2), 69–81.

Bitner, M.J. (1992). Servicespaces: The impact of physical surroundings on customers and employees. *Journal of Marketing*, *56*(2), 57–71.

Bitner, M.J., & Hubbert, A.R. (1994). Encounter satisfaction versus overall satisfaction versus quality. In R.T. Rust & R.L. Oliver (Eds.), *Service quality: New directions in theory and practice* (pp. 72–94). London: Sage.

Bloemer, J., & de Ruyter, K. (1998). On the relationship between store image, store satisfaction and store loyalty. *European Journal of Marketing*, *32,* 499–513.

Bolton, R.N., & Drew, J.H. (1991). A longitudinal analysis of the impact of service changes on customer attitudes. *Journal of Marketing*, *55*(1), 1–9.

Booms, B.H., & Bitner, M.J. (1981). Marketing strategies and organization structures for service firms. In J.H. Donnelly & W.R. George (Eds.), *Marketing of services* (pp. 47–52). Chicago: American Marketing Association.

Borsboom, D., Mellenbergh, G.J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110,* 203–219.

Borsboom, D., Mellenbergh, G.J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111,* 1061–1071.

Brady, M.K. (1997). *Reconceptualizing perceived service quality: Hierarchical model* (unpublished dissertation). The Florida State University.

Brady, M.K., & Cronin, J.J. Jr. (2001). Some new thoughts on conceptualizing perceived service quality: A hierarchical approach. *Journal of Marketing*, *65*(3), 34–49.

Buttle, F. (1996). SERVQUAL: Review, critique, research agenda. *European Journal of Marketing*, *30*(1), 8–25.

Capelleras, J.L., & Veciana, J.M. (2002). *Service quality in university education: An empirical assessment.* Paper presented at the 31st European Marketing Academy Conference, Braga, Portugal.

Carman, J.M. (1990). Consumer perceptions of service quality: An assessment of the SERVQUAL dimensions. *Journal of Retailing*, *66,* 33–55.

Chan, T.F. (1999). Service quality and unfair racial discrimination in homeowners insurance. *The Journal of Risk and Insurance, 66* (1), 83–97.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum Associates.

Cohen, P., Cohen, J., Aiken, L., & West, S. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34,* 315–346.

Cooil, B., Keiningham, T.L., Aksoy, L., & Hsu, M. (2007). A longitudinal analysis of customer satisfaction and share of wallet: Investigating the moderating effect of customer characteristics. *Journal of Marketing*, *71*(1), 67–83.

Cronin, J.J. Jr, & Taylor, S.A. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing*, *56*(3), 55–68.

Czepiel, J.A., Solomon, M.R., & Surprenant, C.F. (1985). *The service encounter*. Lexington, MA: Lexington Books.

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

Dolan, C.V., & van der Maas, H.J.L. (1998). Fitting multivariate normal mixtures subject to structural equation modeling. *Psychometrika, 63*, 227–253.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., & Burke, M.J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods*, *1*(2), 170–177.

East, R. (1992). The effect of experience on the decision making of expert and novice buyers. *Journal of Marketing Management*, *8,* 167–170.

Edwards, J.R. (2001). Multidimensional constructs in organizational behavior research: Towards an integrative and analytical framework. *Organizational Research Methods*, *4,* 144–192.

Ekinci, Y. (2003). Which comparison standard should be used for service quality and customer satisfaction? *Current Issues and Development in Hospitality and Tourism Satisfaction*, *4,* 61–75.

Ferreira, R.R. (1996). The effect of private club members' characteristics on the identification level of members. *Journal of Hospitality and Leisure Marketing*, *4*(3), 41–62.

Fullerton, G. (2005). The service quality–loyalty relationship in retail services: Does commitment matter? *Journal of Retailing and Consumer Services*, *12,* 99–111.

Grace, D., & O'Cass, A. (2005). Service branding: Consumer verdicts on service brands. *Journal of Retailing and Consumer Services*, *12,* 125–139.
Grönroos, C. (1982). *Strategic management and marketing in the service sector*. Helsinki: Swedish School of Economics and Business Administration.
Grönroos, C. (1984). A service quality model and its marketing implications. *European Journal of Marketing*, *18*(4), 36–44.
Grönroos, C. (1990). *Service management and marketing: Managing the moment of truth in service competition*. Lexington, MA: Lexington Books.
Grönroos, C., Heinonen, F., Isoniemi, K., & Lindholm, M. (2000). The netoffer model: A case example from the virtual marketspace. *Management Decision*, *38,* 243–252.
Grove, S.J., & Fisk, R.P. (1997). The impact of other customers on service experiences: A critical incident examination of getting along. *Journal of Retailing*, *73,* 63–85.
Hartline, M.D., & Ferrell, O.C. (1996). The management of customer contact service employees: An empirical investigation. *Journal of Marketing*, *69*(4), 52–70.
Hayduk, L.A. (1996). *LISREL issues, debates and strategies*. Baltimore, MD: Johns Hopkins University Press.
Hayduk, L.A., & Glaser, D.N. (2000). Jiving the four-step, waltzing around factor analysis, and other serious fun. *Structural Equation Modeling: A Multidisciplinary Journal*, *7,* 1–35.
Hayduk, L.A., Ratner, P.A., Johnson, J.L., & Bottorff, J.L. (1995). Attitudes, ideology and the factor model. *Political Psychology*, *16,* 479–507.
Howat, G., Absher, J., Crilley, G., & Milne, I. (1996). Measuring customer service quality in sports and leisure centres. *Managing Leisure*, *1,* 77–89.
Jedidi, K., Jagpal, H.S., & DeSarbo, W.S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, *16,* 39–59.
Johnson, M.D., Herrmann, A., & Huber, F. (2006). The evolution of loyalty intentions. *Journal of Marketing*, *70*(2), 122–132.
Jones, E.T. (2005). *The importance of communication quality in services* (unpublished dissertation). The Florida State University.
Kaplan, D., Harik, P., & Hotchkiss, L. (2000). Cross-sectional estimation of dynamic structural equation models in disequilibrium. In R. Cudeck, S.H.C. Du Toit, & D. Sorbom (Eds.), *Structural equation modeling: Present and future, a festschrift in honor of Karl G. Joreskog* (pp. 315–339). Lincolnville: Scientific Software International.
Karatepe, O.M., Yavas, U., & Babakus, E. (2005). Measuring service quality of banks: Scale development and validation. *Journal of Retailing and Consumer Services*, *12,* 373–383.
Kim, S., & Jin, B. (2002). Validating the retail service quality scale for US and Korean customer of discount stores: An exploratory study. *Journal of Services Marketing*, *16,* 223–237.
Kline, R.B. (2006). Reverse arrow dynamics. In G.R. Hancock & R.O. Mueller (Eds.), *A second course in structural equation modeling* (pp. 43–68). Greenwich, CT: Information Age Publishing.
Ko, Y.J., & Pastore, D.L. (2004). Current issues and conceptualizations of service in the recreational sport industry. *Sport Marketing Quarterly*, *13*(3), 159–167.
Ko, Y.J., & Pastore, D.L. (2005). A hierarchical model of service quality for the recreational sport industry. *Sport Marketing Quarterly*, *14*(2), 84–97.
Lam, T., & Zhang, H.Q. (1999). Service quality of travel agents: The case of travel agents in Hong Kong. *Tourism Management*, *20,* 341–349.
Law, K.S., Wong, C., & Mobley, W.H. (1998). Towards a taxonomy of multidimensional constructs. *Academy of Management Review*, *23,* 741–755.
LeBlanc, G. (1992). Factors affecting customer evaluation of service quality in travel agencies: An investigation of customer perception. *Journal of Travel Research*, *30*(4), 10–17.
Lin, N.-P., Chiu, H.-C., & Hsieh, Y.-C. (2001). Investigating the relationship between service providers' personality and customers' perceptions of service quality across gender. *Total Quality Management*, *12,* 57–68.
Liljander, V., van Riel, A.C.R., & Pura, M. (2002). Customer satisfaction with e-services: The case of an on-line recruitment portal. In M. Bruhn & B. Stauss (Eds.), *Jahrbuch Dienstleistungsmanagement* (pp. 407–432). Wiesbaden: Gabler.
Lozano, M.C., & Fuentes, F. (2003). *Tratamiento borroso del intangible en la valoración de empresas de Internet*. Retrieved from http://www.eumed.net/cursecon/libreria/.

Lubke, G.H., & Muthén, B.O. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10,* 21–39.

Lui, C.-M. (2005). The multidimensional and hierarchical structure of perceived quality and customer satisfaction. *International Journal of Management*, *22,* 426–435.

McAlexander, J.H., Kaldenberg, D.O., & Koenig, H.F. (1994). Service quality measurement. *Marketing Health Services*, *14*(3), 34–44.

Martínez, J.A. (2006a). *La Relación entre la Calidad Percibida, Satisfacción e Imagen. Una Propuesta de Optimación de Modelos Causales en Modelos Transversales* (Pre-Dissertation). Universidad Politécnica de Cartagena.

Martínez, J.A. (2006b). *La calidad percibida en las piscinas de verano del Ayuntamiento de Cartagena*. Report for the City Council of Cartagena, Spain.

Martínez, L. (2006c). *El efecto de las normas ISO 9000 en la percepción de la calidad de los servicios. Modelización jerárquica y multidimensional* (unpublished dissertation). Universidad Politécnica de Cartagena.

Martínez, L., & Martínez, J.A. (2007). Measuring perceived service quality in urgent transport service. *Journal of Retailing and Consumer Services*, *14,* 60–72.

Martínez, L., & Martínez, J.A. (2008). Developing a multidimensional and hierarchical service quality model for the travel agencies industry. *Tourism Management*, *29,* 706–720.

Mattila, A.S. (1999). The role of culture in the service evaluation process. *Journal of Service Research*, *1,* 250–261.

Mittal, V., Katrichis, J.M., & Kumar, P. (2001). Attribute performance and customer satisfaction over time: Evidence from two field studies. *Journal of Services Marketing*, *15,* 343–356.

Mittal, V., Ross, W.T., Jr, & Baldasare, P.M. (1998). The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. *Journal of Marketing*, *62*(1), 33–47.

Netemeyer, R.G., Bentler, P., Bagozzi, R., Cudeck, R., Cote, J., Lehmann, D., . . . Ambler, T. (2001). Structural equations modeling and statements regarding causality. *Journal of Consumer Psychology*, *10*(1–2), 83–100.

Nguyen, N., & LeBlanc, G. (1998). The mediating role of corporate image on customers' retention decisions: An investigation in financial services. *International Journal of Bank Marketing*, *16,* 52–65.

Oliver, R.L. (1993). Cognitive, affective and attribute bases of the satisfaction response. *Journal of Consumer Research*, *20,* 418–430.

Oliver, R.L. (1997). *Satisfaction: A behavioral perspective on the customer*. New York: McGraw-Hill.

Otter, T., Tüchler, R., & Frühwirthz-Schnatter, S. (2004). Capturing consumer heterogeneity in metric conjoint analysis using Bayesian mixture models. *International Journal of Research in Marketing*, *21,* 285–297.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, *49*(4), 41–50.

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,* 12–40.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, *67,* 420–450.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1994). Alternative scales for measuring service quality: A comparative assessment based on psychometric and diagnostic criteria. *Journal of Retailing*, *70,* 193–194.

Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1996). The behavioral consequences of service quality. *Journal of Marketing*, *60*(2), 31–46.

Patterson, P.G., & Spreng, R.A. (1997). Modelling the relationship between perceived value, satisfaction and repurchase intentions in a business-to-business services context: An empirical examination. *International Journal of Service Industry Management*, *8,* 414–434.

Ponce, J.M., Díaz, E., & Martínez, J.L. (1993). *El Sector Asegurador y su Gestión en la Calidad Total*. Seville, Spain: V Encuentros de Profesores Universitarios de Marketing.

Richins, M.L. (1997). Measuring emotions in the consumption experience. *Journal of Consumer Research*, *24,* 127–146.

Rust, R.T., Inman, J.J., Jia, J., & Zahorik, A. (1999). What you don't know about customer-perceived quality: The role of customer expectation distributions. *Marketing Science*, *18,* 77–92.

Rust, R.T., & Oliver, R.L. (1994). Service quality: Insights and managerial implications from the frontier. In R.T. Rust & R.L. Oliver (Eds.), *Service quality: New directions in theory and practice* (pp. 1–19). London: Sage.
Ryan, C., & Cliff, A. (1997). Do travel agents measure up to customer expectation? An empirical investigation of travel agencies service quality as measured by SERVQUAL. *Journal of Travel and Tourism Marketing*, *6*(2), 1–31.
Sachdev, S.B., & Verma, H.V. (2004). Relative importance of service quality dimensions: A multisectorial study. *Journal of Service Research*, *4,* 93–116.
Selnes, F. (1993). An examination of the effect of product performance on brand reputation, customer satisfaction and loyalty. *European Journal of Marketing*, *27*(9), 19–35.
Sharma, S. (2002). Changing face of life insurance in India. *Insurance Chronicle*, *December,* 24–27.
Siu, N.Y.M., & Chow, K.-H. (2004). Service quality in grocery retailing: The study of a Japanese supermarket in Hong Kong. *Journal of International Consumer Marketing*, *16*(1), 71–87.
Smith, D.C., & Park, C.W. (1992). The effects of brand extensions on market share and advertising efficiency. *Journal of Marketing Research*, *29,* 296–313.
Smithson, M.J. (2003). *Confidence intervals.* Quantitative Applications in the Social Sciences Series, No. 140. Thousand Oaks, CA: Sage.
Taylor, S., & Claxton, J.D. (1994). Delays and the dynamics of service evaluations. *Journal of the Academy of Marketing Science*, *22,* 254–264.
Teas, R.K. (1993). Expectations, performance evaluation, and consumers' perceptions of quality. *Journal of Marketing*, *57*(4), 18–34.
Turner, J.C. (1982). Toward a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15–40). Cambridge: Cambridge University Press.
Ueltschy, L., & Krampf, R. (2001). Cultural sensitivity to satisfaction and service quality measures. *Journal of Marketing Theory and Practice*, *9*(Summer), 14–31.
Vermunt, J., & Magidson, J. (2002). Latent class cluster analysis. In J.A. Hagenaars & A.L. McCutcheon (Eds.), *Advances in latent class analysis* (pp. 89–106). Cambridge: Cambridge University Press.
Vermunt, J., & Magidson, J. (2005). *Latent GOLD 4.0 user's guide*. Belmont, MA: Statistical Innovations Inc.
Weiner, B. (1985). An attributional theory of achievement, motivation, and emotion. *Psychological Review*, *4,* 548–573.
Westbrook, R.A. (1981). Sources of consumer satisfaction with retail outlets. *Journal of Retailing*, *57,* 68–85.
Zeithaml, V.A. (1988). Consumer perceptions of price, quality, and value: A means–end model and synthesis of evidence. *Journal of Marketing*, *52*(3), 2–22.
Zeithaml, V.A., Berry, L.L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing*, *60,* 31–46.
Zeithaml, V.A., & Bitner, M.J. (1996). *Service marketing*. New York: McGraw-Hill.
Zeithaml, V.A., Parasuraman, A., & Malhotra, A. (2002). Service quality delivery through web sites: A critical review of extant knowledge. *Journal of the Academy of Marketing Science*, *30,* 362–375.

Appendix A. Indicators

Conduct: *Employees seek the best for the customers.*
Expertise: *The employees have knowledge enough about different services to respond to consumers' requests.*
Problem solving: *When a customer has a problem, this company shows a sincere interest in solving it.*
Tangibles: *The equipment of the company is in good condition.*
Information: *The communication with this company is quick and easy.*
Waiting time: *This company provides service to the consumers punctually.*
Valence: *I would evaluate the outcome of this service favourably.*
Overall service quality: *This company provides an excellent service.*