Journal of Retailing and Consumer Services 17 (2010) 29–42

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

journal homepage: www.elsevier.com/locate/jretconser

Some insights on conceptualizing and measuring service quality

Jose A. Martínez *, Laura Martínez [1]

Department of Management and Marketing, Technical University of Cartagena, Spain

ARTICLE INFO

Keywords:
Service quality
Multidimensional models
Reflective and formative models
Realist and constructivist perspectives

ABSTRACT

We discuss past work in the conceptualization and measurement of perceived service quality and describe the most important models proposed in the last 25 years. We infer the general equations that can be derived from each conceptualization. Finally, we summarize the shortcomings and contradictions of each model as well as conclusions reached so far by a certain consensus of researchers using different models. In order to provide a framework for understanding service quality models, we discuss the service quality paradigm from the realist and constructivist perspective along with the multidimensional nature of service quality implicit in the reflective versus formative debate. We conclude by recommending the development of more creative models of service quality, proposing three different options for quantitative analysis that minimize the various limitations that characterize the most widely used models.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

What is perceived service quality? How must service quality be measured? These two questions have been intensely debated by academics over the last three decades and are among the most recurrent topics in management and marketing literature (e.g. Brady and Cronin, 2001; Ekinci, 2001; Seth et al., 2005; Parasuraman et al., 1985, 1988, 1994). Perceived quality is also an important topic for practitioners. Note that, for example, according to the International Organization for Standardization (ISO), at the end of 2006 there were 897,866 certified companies in the world. With a growing annual rate of 20% for the period, 1995–2006, the number of ISO certified companies has been growing at a rate far higher than the economic growth (Martínez-Costa et al., in press).

In order to shed some light on these questions, several service quality models have been proposed and widely tested in applied research (see Seth et al., 2005, for a review). Grönroos' (1982, 1984) service quality model was the first attempt, and later other important researchers proposed their own conceptualizations (e.g. Parasuraman et al., 1985, 1988, 1994; Cronin and Taylor, 1992; Rust and Oliver, 1994; Dabholkar et al., 1996; Brady and Cronin, 2001). All these models share a common feature: They propose a multidimensional service quality conceptualization that it is inherently linked to the measurement of consumer quality perceptions. Therefore, service quality models offer a framework for understanding what service quality is, as well as how to measure service quality in each proposed conceptualization.

It is generally accepted that all these models have contributed to better understanding the service quality concept. However, there are still several issues that are matters of debate. For example: What is the relationship between service quality and its dimensions? Dabholkar et al. (1996) suggest a reflective model. Parasuraman et al. (1985) propose a formative one. Brady and Cronin (2001) propose both reflective and formative perspectives. We argue that this lack of consensus is partly the result of the fact that the philosophical framework adopted for developing the service quality models is not specified for the aforementioned researchers. Thus, adopting a realist or constructivist position (Hunt, 1991), is of key concern in evaluating and comparing service quality models. We call this issue the "multidimensional problem".

In this paper, we explain the most important service quality models, summarizing the conclusions that have been reached by consensus, and we infer the general equations derived from each conceptualization. This was required to lay a foundation for the two main contributions of our article to the service quality debate: (1) we discuss the service quality paradigm from a positivist and constructivist perspective; (2); we discuss the multidimensional nature of service quality pointing out the advantages and disadvantages of using a reflective or a formative perspective. These two points have not been addressed in previous service quality reviews, such as Ekinci (2002), Finn and Kayande (1998), or Seth et al. (2005). Only the work of Schembri and Sandberg (2002) made a critical review of the most dominant service quality models from an epistemological viewpoint. We expand the criticism made by

* Corresponding author at: Facultad de Ciencias de la Empresa, Universidad Politécnica de Cartagena, Paseo Alfonso XIII, 50, 30203 Cartagena, España. Tel.: +34 968 32 57 76; fax: +34 968 32 70 81.
E-mail addresses: josean.martinez@upct.es (J.A. Martínez), laura.martinez@upct.es (L. Martínez).
[1] Tel.: +34 968 32 59 39; fax: +34 968 32 70 81.

0969-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jretconser.2009.09.002

Schembri and Sandberg (2002) including a discussion about the formative–reflective debate. The inferred equations serve to support our criticism. In addition, we offer guidelines for further research regarding service quality measurement, in order to avoid the methodological problems derived from the multidimensional conceptualization of service quality.

2. 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 has been 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 organization and its services" (Bitner and Hubbert, 1994, p.77). Many models have been developed to measure customer perceptions of service quality. We will briefly describe the most important models in this section:

2.1. Grönroos' Model (GM) and rust and oliver's model (rom)

The first service quality model was proposed by Grönroos (1982, 1984). This author adapts the disconfirmation paradigm from customer satisfaction literature in order to propose that the quality of the service is dependent on expected service and perceived service. Expected service quality is formed by variables such as word of mouth, corporate image, advertising, pricing or personal factors, whereas perceived quality is the result of consumer's view of a bundle of service dimensions, some of which are technical and some of which are functional in nature. Technical quality refers to the outcome of the service performance or what the customer receives in the service encounter. Functional quality relates to the subjective perception of how the service is delivered and defines customers' perceptions of the interactions that take place during service delivery. Grönroos also claims that under certain conditions corporate image can act as another service quality dimension, although, in reality, it is a variable that moderates the relationships between quality dimensions (technical and functional) and perceived quality.

GM proposes that technical solutions or technical abilities of the employees are examples of variables that influence the technical quality dimension, whereas customer-oriented physical and technical resources, accessibility of the firm's services, the consumer orientation of self-service systems, and the firm's ability to maintain continuous contact with its customers are examples of ways of influencing the functional quality dimensions. Fig. 1(a) shows the graphic schema.

Grönroos found empirical support for his model in the responses from a sample of experts and practitioners. Likewise, this model has been applied by other authors (e.g. Kang and James, 2004; Lassar et al., 2002), but without any discussion regarding the multidimensional problem.

Eq. (1) represents GM:

$$P_i = \beta_1 P_{i1} + \beta_2 P_{i2} + \beta_3 CI_i + \beta_4 P_{i1} CI_i + \beta_5 P_{i2} CI_i$$
$$SQ_{i(GM)} = P_i - E_i \quad (1)$$

where SQ_i is the GM overall perceived service quality for individual i; P_{i1} and P_{i2} the perception of individual i with respect to attributes or dimensions 1 and 2; P_i the perceived quality of individual i; E_i the expectations; CI_i the corporate Image; β the weighting factors.

This equation is derived from the reasoning in Grönroos' (1984) article, although this author does not formally express this equation. In addition, Grönroos does not specify the possible difference between weighting factors, but only stress the importance of functional quality. However, in the end Grönroos does assert (p. 43):

> In conclusion one should notice that the quality dimensions are interrelated. An acceptable technical quality can be thought a prerequisite for a successful functional quality. On the other hand, it seems as if temporary problems with the technical quality may be excused, if the functional quality is good enough.

This statement clearly contradicts his original proposal because it acknowledges that there could be reciprocal relationships between both quality dimensions or that quality is a higher-order construct that accounts for the communality between technical and functional quality.

Nevertheless, GM is interpreted in applied research as a single algebraic expression (2):

$$SQ_{i(GM)} = \frac{1}{k}\sum_{j=1}^{k} P_{ij} \Rightarrow E(SQ_i) = \overline{SQ}_{(GM)} = \frac{1}{k}\sum_{j=1}^{k} \overline{P}_j \quad (2)$$

where k is the number of dimensions (2 for GM), P_{ij} the perception of individual i with respect to attribute j. This equation is derived from (1), assuming that service quality is an equally weighted composite of both quality dimensions, and expectations and corporate image are not considered. Most importantly, SQ_i is not a distinct variable from dimensions, but a simple algebraic construction.

It should be noted that Grönroos (1993) measures service quality through performance scores only after recognizing the difficulties in making independent measurements of customer's expectations.

Later, Rust and Oliver (1994) offer a three-component model: the service product (similar to technical quality), the service delivery (similar to functional quality), and the service environment (Fig. 1b). Rust and Oliver did not test their conceptualization, but support has been found for analogous models in retail banking (McDougall and Levesque, 1994).

ROM equation is

$$SQ_{i(ROM)} = \beta_1 P_{i1} + \beta_1 P_{i2} + \beta_1 P_{i3} \quad (3)$$

where SQ_i is the ROM overall perceived quality for individual i; P_{i1}, P_{i2} and P_{i3} the perception of individual i with respect to attributes 1, 2 and 3; β the weighting factors.

Again, ROM is interpreted in applied research as a single algebraic expression:

$$SQ_{i(ROM)} = \frac{1}{k}\sum_{j=1}^{k} P_{ij} \Rightarrow E(SQ_i) = \overline{SQ}_{(ROM)} = \frac{1}{k}\sum_{j=1}^{k} \overline{P}_j \quad (4)$$

where k is the number of dimensions (3 for ROM) and P_{ij} the perception of individual i with respect to attribute j.

2.2. Parasuraman, Zeithaml, and Berry's model (SERVQUAL) and Cronin and Taylor's model (SERVPERF)

Based on the disconfirmation paradigm, Parasuraman et al. (1985, 1988) developed the SERVQUAL model, which breaks down the notion of service quality into five constructs, as follows: tangibles, reliability, responsiveness, assurance and empathy. 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. These authors argued that, regardless of the type of service, consumers evaluate service quality using the same generic criteria, which can be grouped into

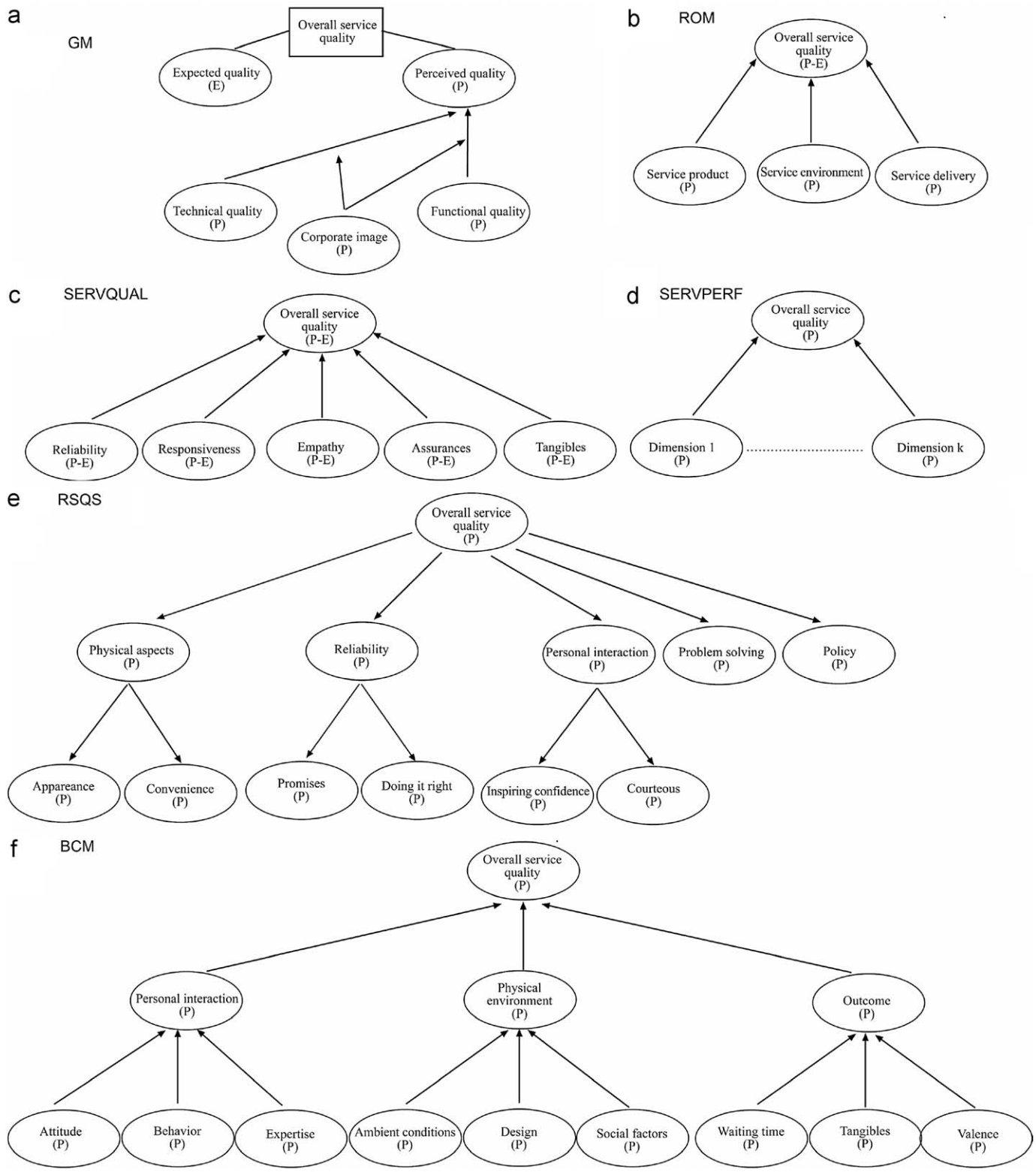

Fig. 1. Service quality models. *Note*: The BCM has been drawn following the empirical analysis achieved by Brady and Cronin (2001). It is important to stress that Brady and Cronin (2001) graphically propose different directions for causal relationships between dimensions and subdimensions. Therefore, in the Brady and Cronin's (2001) original schema, variations in dimensions cause variations in subdimensions. Consequently, the causal direction would be from personal interaction, physical environment and outcome to their respective subdimensions.

five dimensions: tangibles, reliability, responsiveness, assurance and empathy. These five dimensions were derived from 10 overlapping dimensions, which were regarded as essential to service quality by Parasuraman et al.'s (1985) exploratory research. Dimensions of service quality are correlated, and they form the overall service quality perception (Fig. 1c).

During the early 1990s, there was growing argument among academics about service quality conceptualization and measure-

ment, leading to an exchange of critiques and responses between authors that enriched the overall debate. Despite SERVQUAL being the model most widely used and disseminated by academics (Asubonteng et al., 1996; Buttle, 1996; Finn and Kayande, 1998) it was also extensively criticized (Carman, 1990). SERVQUAL's weaknesses led to the development of alternative models to measure customer perceptions of service quality. For example, Cronin and Taylor (1992) developed the SERVPERF model, which is a performance-only measure of service quality (Fig. 1d). Several authors suggested that service quality should be measured considering only consumer perceptions rather than expectations minus perceptions (Carman, 1990; Cronin and Taylor, 1992; McDougall and Levesque, 1994). McDougall and Levesque (1994) considered that including expectation scores on a service quality instrument may be inefficient and unnecessary. This is due to the fact that people tend to give consistently high expectation ratings while their perception scores rarely exceed their expectations (Babakus and Boller, 1992). The performance-only measure obtained 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. In addition, Cronin and Taylor (1992) severely question both the universality of SERVQUAL's five dimensions across service contexts, and the multidimensional nature of SERVQUAL. In fact, they found empirical evidence for unidimensionality.

Equally, Teas (1993, 1994) developed the Evaluated Performance model (EP) in order to overcome some of the problems associated with the gap conceptualization 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.

All these criticism led Parasuraman et al. (1991) to modify SERVQUAL. They revised the concept of expectations and included items to account for importance scores for weighting the dimensions. Later, Parasuraman et al. (1994) again revamped SERVQUAL's structure to embody not only the discordance between perceived service and desired service (labeled as a measure of service superiority, or MSS) but also the discrepancy between perceived service and adequate service (labeled as a measure of service adequacy, or MSA).

The general equation for representing SERVQUAL is as follows:

$$SQ_{i(SERVQUAL)} = \alpha + \beta_1(P_{i1} - E_{i1}) + \beta_1(P_{i1} - E_{i1}) + \beta_2(P_{i2} - E_{i2}) + \beta_3(P_{i3} - E_{i3}) + \beta_4(P_{i4} - E_{i4}) + \beta_5(P_{i5} - E_{i5}) + u_i \quad (5)$$

where SQ_i is SERVQUAL overall perceived quality for individual i; P_{i1}, P_{i2}, P_{i3}, P_{i4}, P_{i5}, the perception of individual i with respect to attributes 1, 2, 3, 4 and 5; E_{i1}, E_{i2}, E_{i3}, E_{i4}, E_{i5} the expectations of individual i with respect to attributes 1, 2, 3, 4 and 5.

Parasuraman et al. (1985) define the determinants of expectations in a very similar form to Grönroos' (1984) model: word of mouth, company marketing communications, personal necessities of consumers and past experience. Hence, there are a number of variables that cause variations in consumer expectations. The difference between GM and SERVQUAL is that SERVQUAL clearly depicts the relationship between quality dimensions and expectations.

Under the same assumptions that applied to (2), the calculation of perceived service quality is reduced to

$$SQ_{i(SERVQUAL)} = \frac{1}{k}\sum_{j=1}^{k} W_j(P_{ij} - E_{ij}) \quad (6)$$

where k is the number of dimensions (5 for SERVQUAL), W_j is the consumer importance assigned to every k distinct dimension. W_j the 1 for original SERVQUAL and is different from 1 for weighted SERVQUAL. Again, SQ_i is a simple algebraic construction, not a distinct variable from dimensions.

However, Parasuraman et al. (1988) admit that one potential application of SERVQUAL is to determine the *relative importance* (italics in the original) of the five dimensions in influencing customer's overall quality perceptions, regressing the overall quality perception scores on the SERVQUAL scores for the individual dimensions. This reasoning led us again to Eq. (5) and raises one of the more controversial issues that we discuss in this article, as we will show later.

Finally Parasuraman et al. (1994) note that intercorrelations of dimensions is an important issue to be taken into account and suggest that future research on service quality should investigate underlying causes and implications of empirical correlations among dimensions of service quality.

On the other hand, the general equation for representing SERVPERF is

$$SQ_{i(SERVPERF)} = \alpha + \sum_{j=1}^{k} \beta_j(P_{ij}) + u_i \quad (7)$$

where SQ_i is the SERVPERF overall perceived quality for individual i; P_{ik}, the perception of individual i with respect to dimension k (k does not have to be necessarily 5 and can vary depending of the analyzed service).

This expression in applied research is again reduced to

$$SQ_{i(SERVPERF)} = \frac{1}{k}\sum_{j=1}^{k} P_{ij} \Rightarrow E(SQ_i) = \overline{SQ}_{(SERVPERF)} = \frac{1}{k}\sum_{j=1}^{k} \overline{P} \quad (8)$$

As can be seen, expressions GM, ROM and SERVPERF are the same, their only difference being the number and labels of the quality dimensions considered.

2.3. Dabholkar, Thorpe and Rentz's model (RSQS)

Dabholkar et al. (1996) proposed a hierarchical model of retail service quality, the Retail Service Quality Scale (RSQS). This scale is regarded as suitable for use in retail businesses which offer a mixture of service and goods, such as department or specialty stores. The RSQS is a multilevel model, where retail service quality is viewed as a higher-order factor defined by two additional levels of attributes (dimensions level and subdimension level). The instrument includes five primary dimensions: physical aspects, reliability, personal interaction, problem solving and policy; and six subdimensions: appearance, convenience, promises, doing it right, inspiring confidence and courteous (Fig. 1e). The procedure used for developing the instrument was based on triangulation of three qualitative research techniques (phenomenological interviews, in-depth interviews and tracking the customers through the store to monitor evaluations of the shopping experience) and literature revision. Dabholkar et al. (1996) used only performance-based measures and found that their scale possessed strong validity and reliability and adequately captured customers' perceptions of retail service quality.

The conceptualization of these authors is radically different from previous models because they consider that service quality is *defined by* and not *formed by* several dimensions. This means that service quality is a higher-order construct with underlying dimensions.

The formal equations of the model are as follows:

$$P_{ij} = \alpha_j + \beta_j SQ_{i(RSQS)} + u_{ij}, \quad j = 1 \ldots k$$
$$P_{ir} = \alpha_r + \beta_r P_{ij} + u_{ir}, \quad r = 1 \ldots s \quad (9)$$

where SQ_i is the RSQS overall perceived quality for individual i; k the number of dimensions (5 for RSQS); s the number of

subdimensions (6 for RSQS); P_{ij} the perception of individual i with respect to attribute j; P_{ir} the perception of individual i with respect to subdimension r.

These equations are equivalent to

$$P_{ij} = \alpha_j + \beta_j SQ_{i(RSQS)} + u_{ij}, \quad j = k - s \ldots k$$

$$P_{ir} = \alpha_r + \alpha_j\beta_r + \beta_r\beta_j SQ_{i(RSQS)} + \beta_r u_{ij} + u_{ir}, \begin{cases} r = 1 \ldots s \\ j = 1 \ldots k - s \end{cases} \quad (10)$$

Under the general assumptions that $\alpha_j = \alpha_r = 0$ and $\beta_j = \beta_r = 1$, and assuming the expected values of the error terms are 0, then

$$SQ_{i(RSQS)} = P_{ij} - u_{ij} \Rightarrow E(SQ_i) = \overline{SQ}_{(RSQS)} = \overline{P}_j, j = k - s \ldots k$$

$$SQ_{i(RSQS)} = P_{ir} - (u_{ij} + u_{ir}) \Rightarrow E(SQ_i) = \overline{SQ}_{(RSQS)} = \overline{P}_r, \quad r = 1 \ldots s \quad (11)$$

Eq. (11) is very interesting because it says that the service quality score of the analyzed service can be obtained from the value of each individual attribute (any of the k–s dimensions or any of the subdimensions). However, this is not what Dabholkar et al. (1996) propose. They indicate that overall service quality can be obtained form the total score of the scale (equivalently, the expected value of the global value is the average score of the k–s dimensions and s subdimensions). Therefore, Dabholkar et al.'s (1996) global service quality score is

$$\overline{SQ}_{(RSQS)} = \frac{1}{2}\left[\frac{1}{k - (k - s)}\sum_{k-s}^{k}\overline{P}_j + \frac{1}{r}\sum_{1}^{s}\overline{P}_r\right] \quad (12)$$

Although Dabholkar et al.'s (1996) study contributed to a greater understanding of service quality in certain retail stores, it was criticized for failing to investigate the relationship between customer perceptions of the quality of the products a retailer carries and customer perceptions of the service quality provided by the retailer (Finn and Kayande, 1997).

2.4. Brady and Cronin's multidimensional and hierarchical model (BCM)

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 multilevel conceptualization of service quality (i.e., Dabholkar et al., 1996). They described a third-order factor model, in which service quality is formed by three primary dimensions, such as interaction quality, physical environment quality and outcome quality. Each of these dimensions is formed by three corresponding subdimensions such as attitude, behavior and experience (interaction quality); ambient conditions, design and social factors (physical environment quality); waiting time, tangibles and valence (outcome quality) (Fig. 1f). Similar to other service quality models, their proposal was based on qualitative research and a literature review.

It is very important to note that Brady and Cronin (2001) propose that subdimensions influence quality dimensions, i.e., subdimensions directly contribute to quality dimensions perception (see the note in Fig. 1). However, their model is operationalized in a different fashion; dimensions are variables that influence subdimensions. This contradiction is not addressed in the Brady and Cronin's article and is a key concern for interpreting their conceptualization.

The general equation for representing BCM, as Brady and Cronin depict in their graphical schema, is the following set of equations:

$$SQ_{i(BCM)} = \alpha + \sum_{j=1}^{k}\beta_j(P_{ij}) + u_i$$

$$P_{ir} = \alpha_r + \beta_r(P_{ij}) + u_{ir}, \begin{cases} r = 1 \ldots 3 \Rightarrow j = 1 \\ r = 4 \ldots 6 \Rightarrow j = 2 \\ r = 7 \ldots 9 \Rightarrow j = 3 \end{cases} \quad (13)$$

where SQ_i is the BCM overall perceived quality for individual i; k the number of dimensions (3 for BCM). R the number of subdimensions (9 for BCM).

It should be recalled that Brady and Cronin (2001) add items to represent primary dimensions and overall service quality, as an implicit assumption of the dimensions as antecedents of service quality and as antecedents of subdimensions. Therefore, assuming the expected values of the error terms are 0, then

$$SQ_{i(BCM)} = \alpha + \sum_{j=1}^{k}\beta_j(P_{ij}) + u_i \Rightarrow E(SQ_i) = \overline{SQ}_{(BCM)} = \alpha + \sum_{j=1}^{k}\beta_j(\overline{P}_j), j = 1 \ldots 3$$

$$\overline{P}_j = \frac{\overline{P}_r - \alpha_r}{\beta_r}, \begin{cases} r = 1 \ldots 3 \Rightarrow j = 1 \\ r = 4 \ldots 6 \Rightarrow j = 2 \\ r = 7 \ldots 9 \Rightarrow j = 3 \end{cases} \quad (14)$$

This means that a difference exists between overall service quality scores and the scores of the aggregate mean of dimensions and the aggregate means of subdimensions. Alpha and beta parameters have to be estimated from the model, and $\overline{P}_j$ can be the mean of any subdimensions pertaining to the corresponding j dimension and pondered by the estimated alpha and beta parameters.

Brady and Cronin (2001, p. 45) clearly assert that dimensions can be used as an effective service quality *proxy*. This means that overall service quality and the composite of quality dimensions are distinct entities. At the same time, the score of each dimension *can not* be directly obtained from the value of each specific subdimensions (as we have depicted in RSQS), because dimensions and subdimensions are also distinct entities. Therefore, each dimension also is a proxy of its corresponding subdimensions.

As we have shown, BCM offers a new perspective for conceptualizing and measuring quality but fails to adequately address some important conceptual contradictions.

To sum up, several worthy models have been proposed to conceptualize and measure perceived service quality. We have derived the equations from each conceptualization showing some important discrepancies among the original schemas, the authors' empirical illustrations, and the form of using service quality models in applied research. Nevertheless, all these contributions have helped to reach a certain consensus on relevant issues in service quality research. (1) It is necessary to develop industry-specific service quality models. There are no universal dimensions/factors/attributes of service quality (Seth et al., 2005), 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[2] (Brady and Cronin, 2001). (2) It is preferable for the model to be performance based measured with respect to the expectations minus perceptions gap measure (Brady and Cronin, 2001).

[2] See, for example, Ekinci (2001) for service quality in hotels, Ko and Pastore (2005) for service quality in the recreational sport industry, or Martínez and Martínez (2007) for service quality in urgent transport service.

3. What is perceived service quality?

Dabholkar et al. (1996, pp. 5–6) wrote:

> Because there is general agreement in the literature about the conceptual definition of service quality, in our qualitative studies we asked general questions about experiences with the retail stores. This is a very common approach in qualitative research. We then selected those aspects of the consumer's experience that were consistent with the generally accepted conceptual understanding of service quality in the literature and used these (along with the literature) to suggest dimensions of retail service quality and items for measuring retail service quality. For example, if customers mentioned courteousness of the service provider and the literature also supported this notion, we included it. If they mentioned price, we excluded it because price is not part of the generally accepted understanding of service quality literature.

This is the traditional perspective adopted by service quality researchers to deal with their models, i.e., there is some degree of consensus among researchers about what service quality is (see the definitions above) and about which aspects, attributes, dimensions or subdimensions are susceptible to be included in a model. This is what Schembri and Sandberg (2002) call the "third-person perspective". This means that those aspects that are elicited in some qualitative research and that do not match with the current existing service quality paradigm are not traditionally considered by researchers (see for example, Brady and Cronin, 2001, p. 36).

Schembri and Sandberg (2002) widely explain this issue in the context of the opposing scientific views of positivism and relativism. These authors explain that to understand the complex nature of service quality, researchers have predominantly, and maybe implicitly, assumed a logical positivist approach. Hunt (1991) explains that positivists rely on observables and the verification principle. That is, positivistic science is objective and discovers the 'true' nature of reality via universal laws that govern the external world where those governing laws can be shown to be true or false and hence cognitively meaningful. However, some researchers (for example, Giorgi, 1983) note broader assumptions associated with positivism, suggesting these metaassumptions parallel the more global philosophy of rationalism. For example, this paradigm maintains that the mental and physical worlds are separate entities with objects existing in the world as a brute reality independent of human experience. In this sense, Schembri and Sanderg argue that traditional service quality research is essentially a dualistic, third-person perspective: the researcher's perspective. Within the traditional rationalistic (inherently positivistic) approach, predefinition of a concept is a necessary part of a rationalistic research approach – hypotheses are formulated, tested and proven, or not.

Positivism and rationalism share the common realist perspective about the world. Classical realism holds that the world exists independently of its being perceived (Hunt, 1990), i.e., theories can be tested against facts that exist independently of theories about them. A less restrictive posture about realism is adopted by critical realism, which says theories can be tested against corresponding facts, despite the fact that they are not fully independent of the theory (Markus, 1998). This critical realism is often identified with (Markus, 1998) or close to (Hunt, 1992) critical rationalism, where Popper's scientific philosophy of falsificationism is developed. Therefore, while this vision regarding science is the common perspective adopted by the majority of researchers, we acknowledge that the debate about the "labels" and the true meaning of these philosophical labels is susceptible to multiple discussions (see Hunt, 1991, 1992).

As Schembri and Sandberg (2002) remind us, reliance on just one methodological approach for any marketing phenomena may create serious limitations. Therefore, Gummeson's (2005) invitation to each individual researcher to consider a multiple research paradigm is not unexpected. The concept of "paradigm" is central to Kuhn's (1970) scientific revolutions argument. A paradigm constitutes the word view of a scientific community, and as Kuhn (1970) explains, this established framework is rarely overturned. However, science progresses through paradigmatic shifts, and critical realism can be a very constraining framework in which to work. Kuhn's advocacy of relativism has influenced marketing philosophy, such as critical relativism (e.g. Anderson, 1983) and the relativistic/constructionist perspective (e.g. Peter and Olson, 1989). Basically, the ideas that support these perspectives are that no interpretation of world can be made independently of human sensations, perceptions, information processing, feelings and actions (Peter, 1992), and there are no independent facts against which theories can be tested (Markus, 1998).

Beyond the discussion about the confronted realist versus relativist/constructionist perspectives (e.g. Hunt, 1991, 1992), we must remember Markus (1998)'s advocacy of constructionist positive contribution: Imagining a world where everything comes with a label, something like a metaphysical supermarket. This world is the realist paradise. In such a world we are trapped inside a fixed and inflexible lexicon where the language is sometimes described as a prison. As Markus asserts, our ability to adapt to emergent social issues stems from our ability to reconfigure the way we connect words to things. This capacity is derived from what constructivism brings to our understanding of language.

Service quality research can profit from a constructionist perspective by extending the current paradigm or rediscovering a new one. Schembri and Sandberg (2002) defend this form of research in the context of service quality (i.e., the "first-person perspective"), recommending phenomenological research to study how consumers conceive service quality and arrive at the widest variation of service quality conceptions. We think that to acknowledge the fact that the concept of quality can have different meanings to different individuals inherently leads to the evaluations of dimensions being distorted, as it could occur that, for several individuals, the dimensions that they are evaluating would not match the dimensions of the concept that they conceive.

A representative example of this approach is the research of Fournier and Mick (1999) on the satisfaction paradigm. These authors found that the current satisfaction paradigm is insufficient for explaining customer satisfaction phenomena; an expanded view has to be considered where satisfaction is a context-dependent process consisting of a multi-model, multi-modal blend of motivations, cognitions, emotions, and meaning, embedded in socio-cultural settings, which is transformed during progressive and regressive consumer–product interactions. In addition, satisfaction is invariably intertwined with life satisfaction and the quality of life itself.

For several years, researchers have debated what the relationship between customer satisfaction and perceived service quality is (e.g. Parasuraman et al., 1985; Bolton and Drew, 1991; Cronin and Taylor, 1992), trying to separate both definitions. However, as Iacobucci et al. (1994) explain, both concepts are attitudes toward the service. They are customer evaluations that can be considered orthogonal concepts, related concepts or even the same concept. As Iacoubucci et al. (1994) indicate, researchers can theoretically distinguish both concepts, but the issue is whether both concepts are substantively different for customers, not for researchers. If quality and satisfaction are viewed as highly related concepts, the

study of Fournier and Mick (1999) should be considered a reference for investigating the meaning of the concept of perceived quality from the customer point of view. Therefore, acknowledging the fact that the concept of quality can have different meanings to different individuals inherently leads to the evaluations of dimensions being distorted, as it could occur that, for several individuals, the dimensions that they are evaluating would not match the dimensions of the concept that they conceive. In addition, rating scales derived from current service quality models would reveal little of how or why quality comes about.

Re-reading the extract from Dabholkar et al. (1996) at the beginning of this section, we see that phenomenological approaches break down with the current, rigid practice of not considering what is not "theoretically" adequate. However, if quality for consumers is dependent on price (here some researchers would say that the proper concept would be "customer perceived value" and not "customer perceived quality"), why do these authors not include price in service quality models? Price has been demonstrated to be a determinant of perceived quality (see for example the research about placebo effect developed by Shiv et al. (2005) or the meta-analysis of Völckner and Hofmann (2007)), so it could be considered a dimension or an indicator of service quality under general assumptions, as we will show further on. Moreover, a paradox emerges from the process of building the depicted service quality models, because they partly rest on qualitative research (and in some cases using phenomenological approaches). However, the creative phase of model building is constrained by the borders of existing literature.

The same occurs with brand image, which it is not included in the aforementioned models (with the exception of GM). Research about cognitive maps of brand associations shows how customers traditionally indicate service quality dimensions to define a brand's image. An example is John et al.'s (2006) study about the brand associations of the Mayo Clinic. After applying Brand Concept Maps' methodology, they found many salient brand associations such as "best patient care available", "best doctors in the world" and "expert in treating serious illness". These statements can be associated, for example, to a functional quality dimension of GM or interaction quality of BCM. At the same time, these core associations are linked to other concepts such as trust, leadership, or innovation that are not traditionally considered in the service quality literature. Can these concepts be considered service quality dimensions/attributes? We think that if those variables are core brand associations directly linked to traditional service quality dimensions, they will probably be taken into consideration when consumers evaluate the excellence or superiority of a service.

A common practice in marketing research consists of separating different areas of marketing as if they were distinct countries with impenetrable border areas. Thus, service quality literature often has not included, in its definition of quality, concepts like price, satisfaction, corporate image, or trust because such concepts correspond with the literature about perceived value, satisfaction, image, and relationship marketing, respectively. However, do we really think that the consumer is capable of making those distinctions? At least we should investigate if we can separate them from the consumer point of view. If social scientists are gradually admitting the necessity of a multidisciplinary perspective in marketing and other social sciences, integrating disciplines such as physics, artificial intelligence, or cognitive psychology (Ball, 2004; Zhang and Zhang, 2007), it seems wrong to put up barriers within marketing science itself.

So, what is perceived service quality and what is the best way to measure it? We need to discuss the multidimensional problem in the context of the philosophy of measurement adopted before attempting to provide a synthesis.

4. Reflective and formative models

We have depicted the equations that represent the multidimensional conceptualization of the most important service quality models. Nevertheless, we have not talked about the relationships between the dimensions and the observed measures because we have supposed that each dimension is perfectly measured by one indicator without error. In practice, however, the aforementioned models propose several indicators for measuring each dimension. This perspective considers the measurement process from a classical test theory perspective. As Borsboom et al. (2003a) explain, classical test theory conceives of measurement in a statistical fashion. A score yielded by an observable indicator is a measure of a theoretical construct if its expected value increases monotonically with that construct. Therefore, the theoretical construct could be taken to be the true score. Measurement error affects the observable variance of the observable indicator but does not bias its expected value.

This perspective is linked to latent variable theory, which is itself compatible with the realist view of science. In addition, realism is associated with causality; theoretical entities are causally responsible for observed phenomena. Therefore, variations in the construct, or theoretical entity, have to be reflected in variations in its observable measures, if these measures are valid for reflecting these variations. The model specifying these relationships is called a reflective model (Edwards and Bagozzi, 2000) and is the most widely used in consumer research.

The reflective model is consistent with realism; latent variable is assumed to exist independent of measurement. However, a constructivist position is possible where latent variable is a construction of the human mind, which need not be ascribed existence independent of measurement. The constructivist position is possible in the most radical view (operationalism), where the latent variable is nothing more that the empirical content it carries, a numerical trick used to simplify the observations (Borsboom et al., 2003b). The model specifying these relationships is called a formative model (Edwards and Bagozzi, 2000), in which a construct is defined in terms of its measures. Because there is one latent variable, both models are considered unidimensional models. There are useful references in marketing about how to distinguish between reflective and formative models that provide recommendations for modeling these two different conceptualizations (e.g. Diamantopoulos and Winklhofer, 2001; Jarvis et al., 2003). Likewise, a deeper theoretical discussion can be found in Edwards and Bagozzi (2000) or Borsboom et al. (2003b).

The philosophical and practical aspects of reflective and formative models is still a matter of debate (see, for example, the different perspectives of Howell et al. (2007), Bagozzi (2007) and Bollen (2007)). However, this controversy is mainly based on how to present formative models within a covariance structure analysis, such as structural equation modeling. When the main objective is to build a scale, we are interested in obtaining an index, a measure of service quality that can be used to profile companies or for further statistical analysis. In this latter case, there is more consensus on how to specify the measurements.

All service quality models have a reflective part because dimensions are measured using several indicators that represent reflections of latent variables. For example, SERVQUAL's 22 indicators (4 for "tangibles", 5 for "reliability", 4 for responsiveness, 4 for assurance and 5 for empathy) reflect those five dimensions, or BCM subdimensions are reflected by 3 indicators

each one. These kinds of models are called spurious models (Edwards and Bagozzi, 2000) and they are respecifications of the formative model that permits a realist interpretation. Some controversy is encountered in GM, because Grönroos indicates that there are several variables that influence both technical and quality dimensions. So, if we interpret the Grönroos words literally, a formative model would be adequate. However, we stress that this author does not explain how to operationalize his model. If we follow some of the GM applications we are following a reflective view (e.g. Kang and James, 2004; Lassar et al., 2002).

Although we remember that the distinction between reflective and formative models can be consulted in the aforementioned references, we want to stress the main differences that serve to discuss the multidimensional problem of service quality models.

Firstly, unidimensional reflective models must meet the principle of local independence, i.e., if a latent variable underlies a number of observed variables, then conditionalizing on that latent variable will render the observed variables statistically independent. This condition means that equally reliable indicators of a latent variable are interchangeable. If we are mainly interested in estimating the mean of the latent, each single indicator should estimate the mean,[3] unexpectedly discarding the requisite that each indicator should not differ in their mean value from the rest. Therefore, high correlations between indicators are not enough to correctly represent the latent because, if indicators are interchangeable, the selection of an indicator with a low score would distort the true mean of the latent if the remaining indicators are highly scored. On the other hand, indicators of the same formative dimension are not interchangeable and do not have to be necessarily correlated; all of them are necessary to determine the meaning of the latent, because they form the latent. In addition, they also require specification of the weights for the construct to be defined. From an operationalist point of view, a concept that it is measured with formative indicators is merely an algebraic construction. There is no causal statement beyond the organizing principle of the representation (Markus, 2004), and there is no distinction between construct and its measures because the construct is defined in terms of its measures.

Secondly, multidimensional reflective and formative models follow the same principles of the unidimensional ones. In the case of multidimensional reflective models the multidimensional construct is manifested by its dimensions and has been labeled as "superordinate construct" (Edwards, 2001). For multidimensional formative models, the multidimensional construct is a composite of its dimensions and has been referred to as an "aggregate construct"[4] (Edwards, 2001). A key issue of multidimensional constructs is that they are conceptualized in terms of their dimensions; they do not exist separately from their dimensions. If a multidimensional construct were replaced by a conceptually analogous construct conceived of as distinct from its dimensions, then the construct and the dimensions would be different entities (Edwards, 2001).

5. The classification of service quality models

We can classify the service quality models in three distinct groups, showing the limitations of each conceptualization:

1. Multidimensional reflective model: RSQS is a multidimensional reflective model. This model assumes that the proposed dimensions are different forms manifested by perceived service quality. Likewise, these dimensions cause variations in their respective subdimensions, and, finally, the indicators are observable representations of each subdimension.
2. Multidimensional formative models: GM, ROM, SERVQUAL and SERVPERF are conceptually multidimensional formative models. The service quality construct is formed by its dimensions and this does not exist separately from its dimensions. Perceived service quality is an algebraic construction derived from dimensions, so service quality and dimensions are not separate concepts.
3. Multidimensional formative–reflective model. BCM combines both the formative and reflective approaches. The service quality construct is formed by the primary dimensions. At the same time these dimensions are reflected by several subdimensions that act as manifestations of the dimensions. However, BCM breaks with the formative–reflective label when it is operationalized by adding items to measure dimensions and subdimensions. Therefore, its conceptualization does not correspond with the form of operationalizing the model.

5.1. Problems with the multidimensional models

5.1.1. Multidimensional reflective model

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, or if they are perceived in a disparate form, the reflective view is open to debate. Could the customer's evaluations of personal interaction be, for example, low correlated with physical aspects in a specific service? In a study about customer evaluation of service performance in the public sport service, Martínez (2006) found that there was a non-significant correlation between 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. In addition, as equally reliable dimensions should be interchangeable, RSQS could be operationalized with only a sample of its five dimensions. However, if we assume that this action would modify the meaning of service quality construct we would be making a formative, rather than reflective, reasoning.

Moreover, the utility of reflective multidimensional constructs has been deeply questioned (see Hayduk et al., 1995; Edwards 2001), 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 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 behavior of the observed items. Certainly, this fact is observed in the

[3] Under the assumption that the latent and the measure share the same origin and have the same scale.

[4] Law et al. (1998) made the same distinction, although they labelled them as "latent model" and "aggregate model", respectively. We will maintain Edwards' (2001) labels in order to avoid the confusion of identifying multidimensional reflective models with "latent models", because all models we consider are models with latent variables.

studies of Martínez and Martínez (2007, 2008b) where standardized total effects of service quality on some indicators exist 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/subdimensions/latents in the lower-order level were analyzed in a more detailed form. For example, the "tangibles" subdimension in the study of Martínez and Martínez (2008b) could be divided in more disaggregated factors like "ambient conditions" and "equipment", and these new factors could be considered the lower-order latents (we can labeled them as "sub-subdimensions"). Consequently, service quality would be a fourth-order concept manifested by dimensions (third-order level), subdimensions (second-order level), sub-subdimensions (first-order level) and observable indicators. It is quickly noted that this procedure could be repeated and that the difficulties of both the conceptualization and the statistical procedure for empirically testing the model would increase dramatically.

5.1.2. Multidimensional formative models

Two main problems arise with these kinds of models:

Firstly, as the service quality concept is formed by its dimensions, all relevant dimensions should be included in the model. Some authors argue that a census of components (dimensions) should be used (Bollen and Lennox, 1991; Diamantopoulos and Winklhofer, 2001). This means that in order to define the meaning of service quality we need to include all relevant facets that form the concept. Although, for example, Rossiter (2002) proposes to relax this restrictive condition using a panel of experts for selecting the most relevant dimensions, it is fairly clear that service quality score is totally dependent on the proposed dimensions.

Secondly, are dimensions of quality distinct from overall service quality? We mean to distinguish the attribute-level performance from global judgments about the service as disparate entities, with divergent meaning, antecedents and consequences. Although this is a very controversial issue that also affects reflective models, it is much more problematic for formative models, because these kinds of models tend to be operationalized adding reflective indicators to the construct, in order to identify the parameters in the widely used structural equation modeling approach (see Edwards, 2001; Kline, 2006). If we can measure service quality in a reflective fashion we are implicitly admitting that service quality and dimensions are distinct entities, because dimensions do not form service quality, they are a proxy of service quality, i.e., service quality score is not yet a weighted combination of its dimensions but the expected value of the reflective indicators. Therefore, service quality would be unidimensional rather than multidimensional, i.e., service quality would exist independently of its dimensions. In this case, it would be difficult to discuss the advisability of moving from operationalism to realism when looking at causality and measurement because the non-literal operationalist view of causal models admits that people can reason in a fashion understood as causal without assuming any projection of causal properties beyond their reasoning process onto the thing about which they reason (Markus, 2004). However, this is not the common interpretation of the majority of marketing researchers about the model represented in Fig. 2c. Rather service quality is regressed on its dimensions, and these dimensions *determine* the variation of service quality concept with certain degree or error. This is an implicit causal statement. Finally, what score of service quality should we consider? The score derived from the reflective items or from dimensions? It seems clear that both scores can be dissimilar and that theoretically the score derived from the reflective items would be used.

5.1.3. Multidimensional formative–reflective model and Brady and Cronin's (2001) operationalization

This kind of model encompasses the problems of reflective and formative models, so it seems weightier. However, as we have mentioned, the operationalization of Brady and Cronin's (2001) model is somewhat distinct from its formative–reflective conceptualization, adding even more controversy. Therefore, we have to discuss this operationalization. Brady and Cronin (2001) add items to represent the primary dimensions and the overall service quality as an implicit assumption of the dimensions as antecedents of service quality and subdimensions. This means that service quality dimensions cause variation in overall service quality and the more disaggregated subdimensions of service quality. This is a very risky statement because they are admitting that service quality, dimensions and subdimensions are distinct entities. In addition, this means that an overall attitude toward the service (service quality) and several service quality attributes (subdimensions) share the same common causes (dimensions).

Obviously, this form of operationalizing the model is subject to the difficulties of inferring causality in cross-sectional designs (Kaplan et al., 2000; Kline, 2006). Therefore, causal relationships among the variables measured at the same time (cross-sectional designs), representing attitudes or subjective perceptions, are cyclic or non-recursive rather than asymmetric. This fact invalidates a great part of applied research that proposes causal modeling of relationships among attitude variables (under a realism point of view) and, specifically in the context of service quality models, it severely questions the BCM. It is not difficult to note that if we assume that service quality and quality dimensions are distinct constructs it would be practically impossible to decide if customers overall perception of service quality is caused by the individual evaluations of dimensions or if customer evaluation of dimensions are caused by a prior global judgment of service quality. Counterfactual theory of causation (Hitchcock, 2002; Pearl, 2000) is useful in explaining this problem.[5]

We suspect that Brady and Cronin's (2001) operationalization was constructed to avoid identification problems for testing their model using the structural equation modeling approach.[6] However, we defend that these kinds of issues related with methodologies for model testing do not have to be prioritized over the corresponding theoretical conceptualization.

[5] Based on the Pearl's (2000) work, Hitchcock (2002) explains how to derive the asymmetry of causation from a corresponding asymmetry in the truth values of counterfactuals. For example, it may be true that if Mary had not smoked, she would have been less likely to develop lung cancer, but we would not normally agree that if Mary had not developed lung cancer, she would have been less likely to smoke. Using a similar reasoning, if the good quality perception of a dimension such as physical quality is considered an antecedent of good global service quality, then it may be true that if physical quality had not been good, then global service quality would have been less likely to be good. However, using counterfactuals we could also agree with the sentence: If global service quality had not been good, it would have been less likely physical quality to be good. Therefore, the asymmetry in causal relationships using attitudinal variables can be analyzed using counterfactuals. Counterfactual theory is one of the basis of Martínez and Martínez's (2008a) proposal regarding how to build causal models in cross-sectional research when using attitudes.

[6] For example, Ko and Pastore (2005) or Martínez and Martínez (2008) implemented third-order hierarchical multidimensional models of service quality, based on Brady and Cronin's (2001) work, but using the Dabholkar et al.'s (1996) approach. These third-order models were reflective, and a series of constraints were specified in order to identify the model. The problem of non-convergence or improper solutions increases when testing this type of models using structural equation modeling.

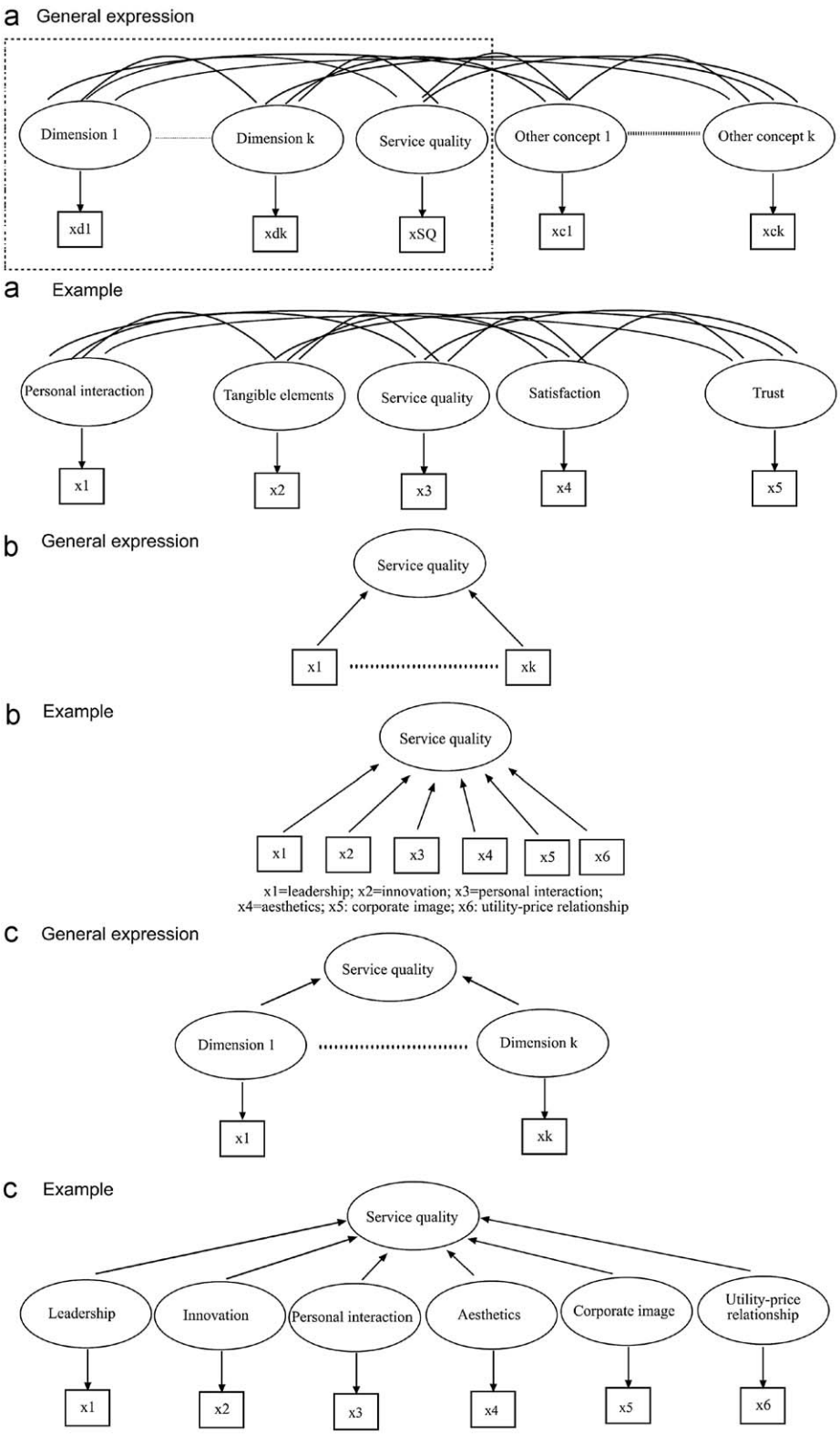

Fig. 2. Options for conceptualizing and measuring service quality.

6. Towards more creative models of service quality

We have made a concise review of the most relevant service quality models and have pointed out several important contradictions and shortcomings of each conceptualization and the way each is operationalized. Considering the key concerns that we have previously debated, we will discuss some thoughts that could serve academics and practitioners for stimulating their perspectives about what is service quality and how to measure it. Recall that there is a great demand for service quality models from companies because of 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.

First of all, we strongly recommend the use of qualitative techniques to find out more about the meaning of service quality for consumers. We think that the literature of service quality needs updating with studies that use phenomegraphic techniques and ethnographics in order to "re-discover" the meaning of quality. Again, we wish to call attention to Fournier and Mick's study (1999) about the satisfaction paradigm as an excellent basis for questioning current service quality research practice. In addition, methodologies for studying brand associations, such as Zaltman's Metaphor Elicitation Technique – ZMET (Zaltman, 1997) or Brand Concept Maps (John et al., 2006), could be used for understanding consumer thoughts about the service quality concept (treating the concept of service quality similar to a brand). This would permit eliciting statements and concepts directly linked with the focused evaluated entity, i.e., the service quality concept. These associations could serve as a starting point for building models to represent the most proper concepts or dimensions related to service quality and to think in a more creative way about a decidedly mature research area. The study of Ekinci (2001) illustrates a fresh perspective on service quality research, using the Q-Sort method to validate service quality dimensions. This method, similar to ZMET of Brand Concept Maps, also relies on consensus to guarantee content validity of the proposed dimensions. However, Ekinci (2001) validated specific dimensions that came from literature review, not from a qualitative study. We stress that prioritizing the consumer voice using the first-person approach, findings can be more creative than when researchers only relay on literature review. Under our view, future research on service quality must move in this direction.

Researchers also have to pay attention to the necessity 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; Ueltschy and Krampf, 2001). 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 due to involvement of customer contact and interaction with employees while a service is delivered. This circumstance increases the importance of the first-person approach in the early stages of research about service quality, because each specific study should adapt to the specific cultural context. There is ample evidence in psychology and sociology supporting that cultural differences may yield divergences in the response patterns (e.g. Hofmans et al., 2009). In addition, as Windschitl and Wells (1996) indicate, interpretation of verbal expressions differ depending on the context or cultural setting to which expressions was referring.

Secondly, quantitative analysis of service quality requires that researchers and practitioners be clearly positioned regarding the issue of considering quality as a composite of several dimensions, as a higher-order factor underlying dimensions or as a distinct entity from dimensions. We have showed several disadvantages of considering each one of these three options. It is very difficult to offer any incontestable argument for choosing one over the others. Recall that this choice depends greatly on the researcher's philosophical view. However, we propose that the best options are the following:

Option 1. To consider service quality as a distinct entity from dimensions without specifying any causal relationships between them (Fig. 2a).

This perspective implicitly acknowledges that customers think differently when evaluating specific attributes than when making an overall judgment, i.e., overall judgments are more susceptible to being "inflated" or "deflated" by general feelings toward the service (e.g. corporate image, emotions, satisfaction, trust, commitment, etc.) and ostensibly go beyond a mere evaluation of specific factors. In this case, it would be appropriate to consider attributes/dimensions as those very specific elements that could be handled by the company (e.g. attributes related with personal interaction, or physical environment) and ignore those attributes that may be outside the company's control and which can not be easily handled by management policies (e.g. valence). The meaning of valence makes this distinction easily understandable. Valence captures attributes that determine whether customers believe the service outcome is good or bad, regardless of their evaluation of any other aspect of the experience (Brady and 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 unfavorable service experience. Therefore, some factors that shape the valence of the outcome are outside the direct control of service management. 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. 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 specific service attributes. This distinction is simply 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 fulfill his/her purchasing goals for this specific purchase situation.

Concepts that are global evaluations of the company (e.g. trust, corporate image, commitment, satisfaction) can be susceptible to measurement if qualitative research shows that they are related with the service quality concept. However, this perspective assumes that service quality is a concept that is also different from these other global attitudes.

Realist researchers would agree with this reasoning. However, it seems clear that this way of proceeding permits a less creative service quality conceptualization because dimensions are expected to be somewhat constrained to attributes that have been traditionally considered in the discussed service quality models.

Option 2. To consider service quality as a composite of several elicited variables (Fig. 2b).

This perspective agrees with formative models of service quality, so the concept is defined in terms of its measures. There is room for obtaining more creative service quality measures, because service quality is simply an algebraic construction derived from what customers think about the meaning of quality for them. The first qualitative stage of the research would be determinant here because there is much more flexibility for building a composite of different variables. For example, if a customer of a specific service associates quality with concepts such as innovation, price–utility relationship, advertisement investment, leadership, satisfaction or trust, there would be no problem in creating an index formed by these variables. Obviously, constructivist researchers would agree with this reasoning.

This formative model would be similar to a widely known technique for evaluation of customer's attitudes: Importance-Performance Analysis (IPA) (see Abalo et al., 2007; Ekinci, 2002). Consumers judge the importance and performance of each relevant attribute and a weighted combination of these attributes forms the service quality index.

Option 3. To consider service quality as a composite of several elicited dimensions in a spurious model (Fig. 2c).

Spurious models are respecifications of formative models where the construct of interest is a composite of several latent variables that are measured in a reflective way. This would permit accounting for the measurement error in the observable indicators of each dimension. Likewise, the researcher is free to choose if he/she wants to implement several measures of each dimension, like the classical multi-item scales used in several service quality models. However, as we have commented, unidimensionality is a fundamental requisite for each dimension.

Again, realist researchers would agree with this view. However, the distinction between a realist and a constructivist perspective in these kinds of models is difficult to make because service quality is formed by its dimensions and, in the special case of dimensions measured without error, the resulting model would be statistically equivalent to the model proposed as Option 2, i.e., a formative model. This fact is addressed in Edwards and Bagozzi (2000) and Borsboom et al. (2003b), but these authors do not provide a clear explanation about the philosophical interpretation.

Finally, we believe that the distinction between realism and constructionism should not be another border area to separate two disparate forms of conceiving service quality research. Instead, this distinction is a reference framework for understanding different positions. Can a researcher navigate a path between these two positions? We think it is possible. In fact, as Peter (1992) asserts, some researchers are knowledgeable about multiple research paradigms and may apply them at different times to different problems in fruitful ways.

What is important for advancing in this research area is to improve procedures that yield useful results and minimize their limitations. As Laudan (1977) argues, the objective of science is to solve problems, providing acceptable answers to interesting questions. The three research options that we have just depicted can be three valid ways to conceptualize and measure service quality that outperform the current service quality models.

7. Conclusion

We have discussed past research on the conceptualization and measurement of perceived service quality, outlining the most important models proposed in the last 25 years. We have pointed out several shortcomings of these models and summarized their points of intersection.

The realist/constructivist framework within which we have situated the service quality research practice should help academics and practitioners to decide the best way to carry out their studies on the perception of quality in services. We have stressed the necessity of building more creative service quality models on the basis of a well-developed qualitative stage of research. For quantitative analysis it is necessary to decide between a formative or a reflective perspective about measurement and adopt one of the three options that we summarize in this paper in order to minimize the various limitations that characterize the most widely used models.

We have deliberately avoided the widespread structure equation modeling notation to deduce the models' equations because this methodology is mainly interested in modeling the covariance structure, where construct and measures are normally expressed as deviations from their means. However, applied research about service quality is very interested in knowing the service quality score because this value acts as a proxy of company performance. The systematic use of tools for measuring quality is obviously focused on comparing the service quality score over time, in order to obtain a dynamic view of customer perceptions. We direct readers to Edwards (2001) or Williams et al. (2003) to find out more on dealing with multidimensional reflective and formative models in more comprehensive nomological networks.

Once researchers know what service quality is for customers with regard to the specific service considered and which of the three options for achieving quantitative analysis is more in accord with their view, they must pay attention to other methodological issues that will be important for a successful research. For example, for options 1 and 3, researchers must decide to use one or multiple indicators per latent variable (see for example Bergkvist and Rossiter (2007) or Hayduk (1996), and the contrasting view of Diamantopoulos (2005) or Mulaik and Millsap (2000)). Other techniques for assessing unidimensionality (distinct from classical test theory), such as Guttman scaling, could be used (Ekinci, 2001). In addition, for any of the options considered, researchers must be cautious about the presence of the halo effect and other method biases (see Podsakoff et al., 2003). Specifically the halo effect is a permanent threat for service quality research because sometimes customers do not have sufficient information to correctly evaluate a large number attributes or quality dimensions and can only judge a few of them. This could be the main reason underlying the high correlations encountered in some studies between quality dimensions (see Cronin and Taylor, 1992). Finally, unobservable heterogeneity in customer's attitudes might be considered in order to obtain a profile of different customer segments with different service quality perceptions (e.g. Allenby et al., 1998; Vermunt and Magidson, 2002).

The distinction between: (1) conceptual models, which are presumed to represent what is happening within an individual consumer for a specific service, and (2) what is done empirically to supposedly provide answers to managerial questions (usually by assuming some model for relationships to be estimated for a population of combinations of services and customers), is habitually made in practice. In fact, all service quality models we have reviewed are conceptual models. The dilemma arises when these models are used to empirically quantify service quality, as we have explained. The philosophical and methodological problems thus occur, because model implementations fail to address some basic concerns regarding measurement. The first-person approach that we suggest, together with the implementation of any of the aforementioned three simple options, would avoid all these problems. These three options are perfectly compatible with more comprehensive conceptual models integrating causes and consequences of perceived quality. Therefore, causes such as company actions (investment in advertising, adopting a quality management system, adopting a CRM system, etc.) and consequences such as future behavior (word of mouth communication, repurchase intentions, etc.), could be integrated in comprehensive causal-effect models (Martínez and Martínez, 2008a).

In conclusion, we have shown how to more properly answer the questions, *what is perceived service quality*? and *how to measure it*? We hope that researchers and practitioners interested in this mature area of marketing research reconsider some of their old paradigms when embarking on new studies.

Acknowledgement

We are in debt with the anonymous reviewers and the editor for their helpful comments.

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 (February), 115–121.

Allenby, G.M., Arora, N., Ginter, J.L., 1998. On the heterogeneity of demand. Journal of Marketing Research 35, 384–389.
Anderson, P.F., 1983. Marketing, scientific progress, and scientific method. Journal of Marketing 47 (Fall), 18–31.
Asubonteng, P., McCleary, K.J., Swan, J.E., 1996. SERVQUAL revisited: a critical review of service quality. Journal of Services Marketing 10 (6), 62–81.
Babakus, E., Boller, G.W., 1992. An empirical assessment of the SERVQUAL scale. Journal of Business Research 24, 235–268.
Bagozzi, R.P., 2007. On the meaning of formative measurement and how it differs from reflective measurement: comment on Howell, Breivik, and Wilcox (2007). Psychological Methods 12 (2), 229–237.
Ball, P., 2004. Critical Mass: The Physics of Society. William Heinemann, London.
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 (2), 175–184.
Bitner, M.J., Hubbert, A.R., 1994. Encounter satisfaction versus overall satisfaction versus quality. In: Rust, R.T., Oliver, R.L. (Eds.), Service Quality: New Directions in Theory and Practice. Sage Publications, Inc., London, pp. 72–94..
Bollen, K.A., Richard Lennox, R., 1991. Conventional wisdom on measurement: a structural equation perspective. Psychological Bulletin 110, 305–314.
Bollen, K.A., 2007. Interpretational confounding is due to misspecification, not to type of indicator: comment on Howell, Breivik, and Wilcox (2007). Psychological Methods 12 (2), 219–228.
Bolton, R.N., Drew, J.H., 1991. A longitudinal analysis of the impact of service changes on customer attitudes. Journal of Marketing 55 (January), 1–9.
Borsboom, D., van Heerden, J., Mellenbergh, G.J., 2003a. Validity and truth. In: Yanai, H., Okada, A., Shigemasu, K., Kano, Y., Meulman, J.J. (Eds.), New Developments in Psychometrics. Proceedings of the International Meeting of the Psychometric Society 2001, pp. 321–328.
Borsboom, D., Mellenbergh, G.J., van Heerden, J., 2003b. The theoretical status of latent variables. Psychological Review 110 (2), 203–219.
Brady, M.K., Cronin, J.J., 2001. Some new thoughts on conceptualizing perceived service quality: a hierarchical approach. Journal of Marketing 65 (July), 34–49.
Buttle, F., 1996. SERVQUAL: review, critique, research agenda. European Journal of Marketing 30 (1), 8–31.
Carman, J.M., 1990. Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions. Journal of Retailing 66, 33–55.
Cronin, J.J., Taylor, S.A., 1992. Measuring service quality: a reexamination and extension. Journal of Marketing 56, 55–68.
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 (1), 3–16.
Diamantopoulos, A., 2005. The C-OAR-SE procedure for scale development in marketing: a comment. International Journal of Research in Marketing 22, 1–9.
Diamantopoulos, A., Winklhofer, H., 2001. Index construction with formative indicators: an alternative to scale development. Journal of Marketing Research 38 (2), 269–277.
Edwards, J.R., Bagozzi, R.P., 2000. On the nature and direction of relationships between constructs and measures. Psychological Methods 5, 155–174.
Edwards, J.R., 2001. Multidimensional constructs in organizational behavior research: towards an integrative and analytical framework. Organizational Research Methods 4 (2), 144–192.
Ekinci, Y., 2001. The validation of the generic service quality dimensions: an alternative approach. Journal of Retailing and Consumer Services 8 (6), 311–324.
Ekinci, Y., 2002. A review of theoretical debates on the measurement of service quality: implications for hospitality research. Journal of Hospitality and Tourism Research 26 (3), 199–216.
Finn, A., Kayande, U., 1997. Consistency of the relationship between retailer product and service quality. In: Fourth Recent Advances in Retailing and Services Science Conference, Scottsdale, Arizona.
Finn, A., Kayande, U., 1998. The service quality literature: a generalizability perspective. Advances in Services Marketing and Management 7, 97–129.
Fournier, S., Mick, D.G., 1999. Rediscovering satisfaction. Journal of Marketing 64 (4), 5–23.
Giorgi, A., 1983. Concerning the possibility of phenomenological research. Journal of Phenomenological Psychology 14 (Fall), 129–170.
Grace, D., O'Cass, A., 2005. Service branding: consumer veredicts on service brands. Journal of Retailing and Consumer Services 12 (2), 125–139.
Grönroos, C., 1982. Strategic Management and Marketing in the Service Sector. Swedish School of Economies and Business Administration, Helsinki.
Grönroos, C., 1984. A service quality model and its marketing implications. European Journal of Marketing 18 (4), 36–44.
Grönroos, C., 1993. toward a third phase in service quality research: challenges and future directions. Advances in Services Marketing and Management 2, 49–64.
Gummeson, E., 2005. Qualitative research in marketing. Road-map for a wilderness of complexity and unpredictability. European Journal of Marketing 39 (3/4), 309–327.
Hayduk, L.A., 1996. LISREL Issues, Debates and Strategies. Johns Hopkins University Press, Baltimore, MD.
Hayduk, L.A., Ratner, P.A., Johnson, J.L., Bottorff, J.L., 1995. Attitudes, ideology and the factor model. Political Psychology 16 (3), 479–507.
Howell, R.D., Breivik, E., Wilcox, J.B., 2007. Reconsidering formative measurement. Psychological Methods 12 (2), 205–218.
Hitchcock, C., 2002. Probabilistic causation. Stanford Encyclopedia of Philosophy.
Hofmans, J., Theuns, P., Van Acker, F., 2009. Combining quality and quantity. A psychometric evaluation of the self-anchoring scale. Quality & Quantity 43 (5), 703–716.
Hunt, S.D., 1990. Truth in marketing theory and research. Journal of Marketing 54 (July), 1–15.
Hunt, S.D., 1991. Positivism and paradigm dominance in consumer research. Journal of Consumer Research 18 (June), 32–44.
Hunt, S.D., 1992. For reason and realism in marketing. Journal of Marketing 56 (April), 89–102.
Iacobucci, D., Grayson, K., Ostrom, A., 1994. The calculus of service quality and customer satisfaction: theoretical and empirical differentiation and integration. In: Schwartz, T.A., Bowen, D.E., Brown, S.W. (Eds.), Advances in Services Marketing and Management, vol. 3. JAI Press, Greenwich, CT, pp. 1–67.
Jarvis, C.B., MacKenzie, S.B., Podsakoff, P.M., 2003. A critical review of construct indicators and measurement model misspecification in marketing and consumer research. Journal of Consumer Research 30, 199–218.
John, D.R., Loken, B., Kim, K.-H., Monga, A.B., 2006. Brand concept maps: a methodology for identifying brand association networks. Journal of Marketing Research 43 (4), 549–563.
Kang, G.-D., James, J., 2004. Service quality dimensions: an examination of Grönroos's service quality model. Managing Service Quality 14 (4), 266–277.
Kaplan, D., Harik, P., Hotchkiss, L., 2000. Cross-sectional estimation of dynamic structural equation models in disequilibrium. In: Cudeck, Stephen H.C. Du Toit, Dag Sorbom (Eds.), Structural Equation Modeling: Present and Future. A Festschrift in Honor of Karl G. Joreskog. Robert. Lincolnville, Scientific Software International, pp. 315–339.
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.
Kline, R.B., 2006. Reverse arrow dynamics. In: Gregory, R. Hancock, Ralph, O. Mueller (Eds.), A Second Course in Structural Equation Modelling. CT, Information Age Publishing, Greenwich, pp. 43–68.
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.
Kuhn, T.S., 1970. The Structure of Scientific Revolutions, second ed. University of Chicago Press, Chicago.
Lassar, W.M., Manolis, C., Windsor, R., 2002. Service quality perspective and satisfaction in private banking. Journal of Services Marketing 14 (3), 244–271.
Laudan, L., 1977. In: Progress and Its Problems. University of California Press, Berkeley, CA.
Law, K.S., Wong, C.-S., Mobley, W.H., 1998. Towards a taxonomy of multidimensional constructs. Academy of Management Review 23 (4), 741–755.
Markus, K.A., 1998. Science, measurement, and validity: is completion of Samuel Messick's synthesis possible?. Social Indicators Research 45, 7–34.
Markus, K.A., 2004. Varieties of causal modeling: how optimal research design varies by explanatory strategy. In: van Montfort, K., Oud, J., Satorra, A. (Eds.), Recent Developments on Structural Equation Models: Theory and Application. Kluwer Academic Publishers, Dordrecht, pp. 175–196.
Martínez, J.A., 2006. La calidad percibida en las piscinas de verano del Ayuntamiento de Cartagena. Report for the City Council of Cartagena, Spain.
Martínez, L., Martínez, J.A., 2007. Measuring perceived service quality in urgent transport service. Journal of Retailing and Consumer Services 14, 60–72.
Martínez, J.A., Martínez, L., 2008a. Building better causal models to measure the relationship between attitudes and customer loyalty. International Journal of Market Research 50 (4), 437–447.
Martínez, L., Martínez, J.A., 2008b. Developing a multidimensional and hierarchical service quality model for the travel agencies industry. Tourism Management 29 (4), 706–720.
Martínez-Costa, M., Martínez, J.A., Choi, T., Martínez-Lorente, A.R., in press. ISO 9000/1994, ISO 9001/2000 and TQM: the performance debate revisited. Journal of Operations Management, doi:10.1016/j.jom.2009.04.002.
Mattila, A.S., 1999. The role of culture in the service evaluation process. Journal of Service Research 1 (3), 250–261.
McDougall, G.H.G., Levesque, T.J., 1994. A revised view of service quality dimensions: an empirical investigation. Journal of Professional Services Marketing 11 (1), 189–209.
Mittal, V., Ross, W.T., Baldasare, P.M., 1998. The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions. Journal of Marketing 62 (January), 33–47.
Mulaik, S.A., Millsap, R.E., 2000. Doing the four step right. Structural Equation Modelling: A Multidisciplinary Journal 7, 36–73.
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 (Fall), 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 (1), 12–40.
Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1991. Refinement and reassessment of the SERVQUAL scale. Journal of Retailing 67 (4), 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 (3), 193–194.
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.

Pearl, J., 2000. In: Causality: Models, Reasoning, and Inference.. Cambridge University Press, Cambridge, UK.
Peter, J.P., 1992. Realism or relativism for marketing theory and research: a comment on Hunt's "Scientific Realism". Journal of Marketing 56 (2), 72–79.
Peter, J.P., Olson, J.C., 1989. The relativistic/constructionist perspective on scientific knowledge and consumer research. In: Elizabeth, C. Hirschman (Ed.), Interpretive Consumer Research. Association for Consumer Research, Provo, UT, pp. 24–28.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method variance in behavioral research: a critical review of the literature and recommended remedies. Journal of Applied Psychology 88, 879–903.
Rossiter, J.R., 2002. The Coarse procedure for scale development in marketing. International Journal of Research in Marketing 19, 305–335.
Rust, R.T., Oliver, R.L., 1994. Service quality: insights and managerial implications from the frontier. In: Rust, R.T., Oliver, R.L. (Eds.), Service Quality: New Directions in Theory and Practice. Sage Publications, London, pp. 1–19.
Schembri, S., Sandberg, J., 2002. Service quality and the consumer's experience: towards an interpretive approach. Marketing Theory 2 (2), 189–205.
Seth, N., Deshmukh, S.G., Vrat, P., 2005. Service quality models: a review. International Journal of Quality and Reliability Management 22 (9), 913–949.
Shiv, B., Carmon, Z., Arieli, D., 2005. Placebo effects of marketing actions: consumers may get what they pay for. Journal of Marketing Research 42 (November), 383–393.
Teas, K.R., 1993. Expectations, performance evaluation, and consumer's perceptions of quality. Journal of Marketing 57 (October), 18–34.
Teas, K.R., 1994. Expectations as a comparison standard in measuring service quality: an assessment of a reassessment. Journal of Marketing 58 (January), 132–139.
Ueltschy, L.C., Krampf, R.F., 2001. Cultural sensitivity to satisfaction and service quality measures. Journal of Marketing Theory and Practice 9, 12–31.
Vermunt, J., Magidson, J., 2002. Latent class cluster analysis. In: Jacques, A. Hagenaars, Allan, L. McCutcheon (Eds.), Advances in Latent Class Analysis. Cambridge University Press, Cambridge, pp. 89–106.
Völckner, F., Hofmann, J., 2007. Perceived price-quality relationship. A meta-analytic review and assessment of its determinants. Marketing Letters 18, 181–196.
Williams, L.J., Edwards, J.R., Vandenberg, R.J., 2003. Recent advances in causal modeling methods for organizational and management research. Journal of Management 29 (6), 903–936.
Windschitl, P.D., Wells, G.L., 1996. Measuring psychological uncertainty: verbal versus numeric methods. Journal of Experimental Psychology: Applied 2, 343–364.
Zaltman, G., 1997. Rethinking market research: putting people back in. Journal of Marketing Research 34 (November), 424–437.
Zeithaml, V.A., 1988. Consumer perceptions of price, quality, and value: a Means-end model and synthesis of evidence. Journal of Marketing 52 (July), 2–22.
Zhang, T., Zhang, D., 2007. Agent-based simulation of consumer purchase decision-making and the decoy effect. Journal of Business Research 60 (8), 912–922.