Customer satisfaction measurement: comparing four methods of attribute categorisations

Boris Bartikowski, Sylvie Llosa

Customer Satisfaction Measurement:

Comparing Four Methods of Attribute Categorizations

Authors:

Boris Bartikowski   and   *Sylvie Llosa*

IAE d'Aix en Provence, CEROG   IAE d'Aix en Provence, CEROG
Université de Droit, d'Economie et des   Université de Droit, d'Economie et des
Sciences d'Aix-Marseille III, France (Attaché)   Sciences d'Aix-Marseille III, France

Please, write to:

Boris Bartikowski   or to   Professor Sylvie Llosa
Tulpenweg 11   IAE d'Aix en Provence
86316 Friedberg/ Germany   Clos Guiot
Tel.: (49) 179.771.339.5   13540 Puyricard/ France
E-mail: boris.bartikowski@t-online.de   Tel. : (33) 684.071.390
                                               E-mail: Sylvie.llosa@iae-aix.com

Length of the article: 3267 words without references

Customer Satisfaction Measurement:

Comparing Four Methods of Attribute Categorizations

Abstract

The issue of how to weight and categorise service attributes has attracted great attention from academics as well as practitioners. The application of an inappropriate method could lead to misleading interpretations and the implementation of useless and costly actions. We first review several streams of literature concerning the theoretical background of attribute categorizations in relation to customer satisfaction. Then we identify four methods that have been developed to categorise attributes into four classes. In the next step we apply these methods in an empirical study. Criteria for distinguishing the considered approaches conceptually and methodologically are proposed, and implications for future research are discussed.

Key words

Customer satisfaction, Attribute categorization, Service, Insurance

Introduction

Customer satisfaction (CS) has been the object of numerous discussions and recent evolutions indicate that CS remains in the limelight, especially in the service field. CS is typically defined as an overall assessment of the performance of various attributes that constitute a product or a service [e.g. Swan and Combs, 1976; Johnston, 1995; Sampson and Showalter, 1999]. Increased satisfaction increases customer loyalty, positive word of mouth, customer retention and, by extension, a firm's profitability [for a meta analysis see Szymanski and Henard, 2001]. As a consequence, practitioners need to understand how satisfaction is engendered and how it can be influenced. A promising step is to work out (a) which attributes should be improved to increase satisfaction and (b) which attributes should be reduced because high performance on them is costly or offers no increase in satisfaction. To achieve such business intelligence "an important step is to recognise that the links in the satisfaction-profit chain are asymmetric and nonlinear" [Anderson and Mittal, 2000]. Taking this into account delivers attribute categorizations that allow efficiently organizing performance improvements and resource allocations.

This article treats such considerations theoretically and empirically. We first review relevant literature. Then, we present four methodological approaches that are based on the idea of asymmetric and nonlinear links between attribute performance and overall CS. The objective of these methods is to categorise attributes according to their relationship with CS. The empirical part of the article is designed to study the convergence of some of these methods. This allows a better interpretation of the results, brings about theoretical and managerial conclusions and provides suggestions for future research.

Theoretical considerations

The expectation disconfirmation paradigm is probably the best known framework for satisfaction studies [Engel *et al.*, 1968; Oliver, 1977]. It proposes that customers maintain a standard of reference to which they compare perceived performance. Satisfaction results if performance is higher than expected, dissatisfaction results if it is lower. In this sense, one might argue that all attributes have not the same impact on CS when compared to one and the same standard. Two general cases can be distinguished:

- The concept of invariant attribute weights proposes that CS is strongly affected, whether the attribute performs good or bad. For example, the taste of food in a restaurant should always have a strong impact on CS, whether it's good or bad.

- The concept of variant attribute weights proposes that the weights of several attributes are performance related. Consider the following examples: (a) a clean restaurant engenders no satisfaction, but a dirty restaurant strongly attracts negative feelings, (b) receiving a free drink in a restaurant is pleasantly surprising but there's no reason for dissatisfaction if it is not delivered for free. The two-factor theory, developed in the field of job satisfaction, is a famous reference pleading in favor of variant attribute weights: Herzberg *et al.* [1959] conclude that the opposite of job satisfaction is not job dissatisfaction, but no job satisfaction. They furthermore conclude that certain factors operate only to increase satisfaction (motivators) while others only increase dissatisfaction (hygiene factors). This suggests that a negatively (positively) perceived attribute could have a stronger impact on overall CS than if the same attribute has been perceived positively (negatively).

Several authors support indirectly the concept of variant attribute weights by arguing that standards of reference are not always pre-established [e.g. Westbrook and Reilly, 1983; Kahneman and Miller, 1986; Cadotte and Turgeon, 1988; McGill and Iacobucci, 1992; Ngobo, 1997; Llosa, 1997]. Results of Voss *et al.* [1998] point into the same direction: These

authors show that the effect of performance expectations depends on price/performance consistency, which could vary depending on the service exchange. Such reasoning is especially pertinent for services, as attributes contributing to CS could work differently according to how a service encounter takes place. For example, the interior of a hotel's reception hall is more salient if guests have long waiting times, while it is relatively less important if the staff is very courteous. This suggests that each service encounter creates its own standards.

To sum up, the concept of variant and invariant weights proposes four attribute categories: Dissatisfiers influence CS strongly only in case of low performance; Satisfiers only in case of high performance; Criticals impact CS strongly in case of low as well as high performance, and the impact of Neutrals on CS is generally weak. These basic considerations are illustrated in Figure 1.

Take in Figure 1.

The managerial implications of this concept are described as follows: It is quite difficult to satisfy customers through Dissatisfiers, but bad performances strongly diminish satisfaction. For this reason, Dissatisfiers should be as standardised as possible, at the performance level expected by the client. On the other side, firms should provide high performances on several well-chosen Satisfiers. These are satisfaction boosters if they are specifically included into the offer. Firms stand to win bonus points for providing high performance on Criticals, to which clients react very sensitively. They risk demerits if low performance is delivered on them. Finally, improving performance on Neutrals is of no priority. Some of these attributes may be even a source for redirecting resources so as to save money.

Methodological considerations

The above described concept has been treated in numerous theoretical and empirical studies [e.g. Leavitt, 1977; Maddox, 1981; Rust and Zahorik, 1993; Johnston, 1995]. The reported empirical results are mostly based on the assumption that the applied method is a valid measure of the underlying concept of variant and invariant weights (Figure 1). The following sections provide a brief description of four methods.

Direct approaches

There are several techniques for attribute categorization that are said to be "direct approaches", because respondents are directly asked about attributes weights. We now review two direct approaches.

Dual importance mapping (DIM). DIM is an advanced version of the classical importance-performance analysis as proposed by Martilla and James [1977]. The measures used for DIM are: (a) 'stated importance', which corresponds to a respondent's direct assessment of an attributes importance and (b) 'derived importance', which corresponds to the strength of correlation between attribute performance and overall CS [e.g. Vavra, 1997; Oliver, 1997, p. 59]. Plotting these scores on a x/y graph allows categorizing attributes into four classes:

An attribute is Flat if its stated importance is high and if its derived importance is weak. Such attributes correspond to common or expected quality standards that must be reached [Venkitaraman and Jaworski, 1993]. An attribute with both strong stated and strong derived importance is Key, which means that customers react extremely sensitive to its higher or lower performances. An attribute with strong derived but weak stated importance is Value-added. This corresponds to unexpected or pleasantly surprising aspects. Finally, if both importance measures are weak the attribute is said to be of no concern to the customer (Low yield attribute).

Simulation method (SM). SM was developed in the field of tangible products, particularly for the development of new products [Kano, 1984; Berger *et al.*, 1993; Matzler *et al.*, 1996]. As a managerial tool, the approach is often integrated into so-called Six Sigma programs for continual business improvements [e.g. Mazur, 2001]. The proposed DMAIC (Define-Measure-Analyze-Improve-Control) process is based on categories of Critical Quality Characteristics that are obtained through SM. Kano [1984] takes into account that some links between performance and overall CS are nonlinear and asymmetric: Bad performances on Must-be attributes lead to dissatisfaction while good performances on these attributes cannot engender satisfaction. One-dimensional attributes lead to higher dissatisfaction/satisfaction the lower/higher the performance on these attributes is. Attractive attributes engender satisfaction, but dissatisfaction cannot result even if performance is low.

The empirical categorization relies on answer combinations to two questions. The functional question refers to a respondents reaction in case of good performance and the dysfunctional question refers to his reactions if the same attribute performs badly. It is proposed that the combination of the resulting responses permits categorizing attributes into six classes, as shown in Table 1.

Take in Table 1.

Besides the three principal categories described above, Table 1 reveals a fourth category called Indifferent. These are attributes that engender neither satisfaction nor dissatisfaction, whether they perform good or bad. So-called Reverse and Questionable are considered as unclear results that must be treated subtly differentiated.

Indirect weights assessments

While the above described techniques for attribute categorizations are largely based on direct questioning, the following approaches are said to be indirect because attribute weights are statistically determined from observed associations with CS. We now review two indirect approaches.

Penalty reward contrast analysis (PRCA). Brandt [1988], and with methodological variations Vanhoof and Swinnen [1998], Mittal *et al.* [1998] or Anderson and Mittal [2000], propose very similar approaches for attribute categorizations. These authors compare relationships between overall CS and attribute performances. For example Brandt and Scharioth [1998] first carry out a recoding of attributes into dummy variables, as shown in Table 2.

Take in table 2

Regression analysis is then performed with the dummies as predictive variables and with global satisfaction as the dependent variable. The mean differences should indicate the impact of negative or positive performances on overall CS. Table 3 summarises rules that are adopted for categorizing attributes into four classes.

Take in table 3

Correspondence Analysis (CA). Llosa [1997, 1999] proposes one more indirect method (Tetraclasse model) for attribute categorizations. The author applies factorial analysis

page 8/24

of correspondences to a contingency table. This table contains the number of high/low attribute performances in the lines and two levels of CS in the columns, as shown in Table 4

Take in Table 4

CA allows figuring both CS and the two modalities of attribute performances on a single factorial axis. This explains obligatory 100% of variance (Figure 2).

Take in Figure 2

The adopted logic for attribute categorizations is that, the nearer an attribute is positioned (row points of factor scores) to the extremities of the axis, the stronger is its influence on CS: An attribute is Basic if it scores high in case of low performance and if it scores low in case of high performance. An attribute is Plus if these facts are turned around. If an attribute scores high, whether the level of performance is low or high, it is said to be Key. It is Secondary if both scores are low. The two factor scores of CS itself (column scores) draw up frontiers that allow distinguishing four attribute categories.

Research Questions

While the conceptual background of the exposed methods is largely based on the same literature, the methodological approaches differ considerably. The attributes categories are given various names, but the proposed interpretations are essentially equal (c.f. Table 5).

Take in Table 5

If marketers are interested in such attribute categorizations for supporting investment decisions and resource allocations, they are certainly eager to obtain meaningful results. But are the attributes always equally classified through the different methods? For example Brandt and Scharioth [1998] report that only 67%, or sixteen of twenty-four considered

attributes, were equally classified for PRCA and for DIM. To expand this assessment to four methods, we conducted the following empirical study.

Research setting

The empirical study was carried out in the field of insurance consultations in Germany. Establishing an efficiently operating sales organization is important for an insurance company. Learning about determinant attributes during the service encounter could be helpful for sales training, for example.

Attributes were identified through 15 in-depth interviews with customers of an insurance company who recently encountered a representative of this company. We asked these individuals about their experiences during the encounter and to point out if they perceived problems or pleasantly surprising aspects. Additional attributes were generated during a round table with five executives of the same company. The exploratory phase produced an exhaustive list of 52 attributes likely to influence CS. We then asked the executive panel to synthesise the list to 20 attributes they considered most relevant (see Appendix). However, because Kano's [1984] method produces long questionnaires, only four attributes were taken into account for this method (see Table 6).

Take in Table 6

Insurance customers may hold different expectations when insuring a car or when signing a long-term life insurance policy. It was therefore concluded that this type of service encounter is particularly heterogeneous. In order to homogenise the sample, respondents were required having encountered a sales representative who proposed certain types of low budget policies (accident insurance, supplementary health insurance). As customers' perceptions may considerably change by time, it was furthermore required that the encounter took place within the last 14 days. The company delivered a list of 225 customers who applied to these criteria. In order to maximise the return rate, these individuals were first contacted by

telephone and asked to take part in the study. The 195 persons who agreed received a questionnaire together with a stamped and addressed return envelope. As insurance services contain sensible personal data and in order to avoid response bias, respondents were asked to send back their questionnaire anonymously. Our sample consists of 123 persons who sent back their questionnaires within a delay of 10 days (response rate = 63%).

The following measures were used (see Appendix for details): Satisfaction was measured with four five-point rating scales that record emotional and intentional reactions [e.g. Westbrook and Oliver, 1981; Hausknecht, 1990]. These built a cumulative measure of overall CS (explained variance 1st factor = 80.5%; α=0.92.). We also used five-point rating scales for measuring the 20 attribute performances. For measuring declared importance, we first asked respondents to choose the five most important attributes and then to rank them according to their importance. This procedure facilitates the respondent's task, as it is less difficult than ranking 20 attributes in one go. Finally, for SM, eight additional questions were formulated.

Data analysis

DIM: Derived importance was defined as the strength of correlation (R^2) between CS and perceived attribute performances. Declared importance (DI) was defined as the probability that an attribute is the most, the second … the fifth important. This was obtained as follows: DI = $P(x/1^{st}$ rank)*5 + $P(x/2^{nd}$ rank)*4 + …+ $P(x/5^{th}$ rank). The mean scores of declared and the mean scores of stated importance separate four attributes categories.

SM: This analysis requires cumulating the responses of the combinations of functional and dysfunctional questions and then examining the resulting distribution. In this study, ambiguous categorizations were produced. Results suggest that several customer segments yield different preferences. For example, the attribute Gift could be interpreted as Attractive as well as Indifferent.

PRCA: For each attribute, perceived performance was divided into three groups that were interpreted as high, normal and low. As proposed by Brandt and Scharioth [1998], a dummy variable regression, with overall satisfaction as the dependent variable, was conducted. This allows determining if CS in case of low or high performance is statistically different from CS in case of normal performance. For example for the attribute punctuality it was found that CS is significantly higher in case of high (D1), than in case of normal performance. There was however no significant difference in CS from normal to low performance (D2). Punctuality was therefore classified a Satisfier.

CA: Correspondence Analysis is a non-parametric method that requires choosing the type of distance among the rows and the columns of the correspondence table and to set up a kind of normalization. In this study chi-square distances and symmetrical normalization were used. The two factorial row scores for each attribute are compared to the two column scores for CS. For example the attribute Information affects CS strongly if it performs low, while it has little impact if it performs well.

Summary of Results

Table 7 summarises how the four central attributes were categorised through the four methods.

Take in Table 7

It is evident that these results are far from converging, in spite of the methods' largely identical theoretical background. This is not too surprising as the measures and the applied data analysis differ considerably. But, which method should then be used for supporting investment decisions and resource allocations? Each method produces results that lead to different conclusions. The next section clears up this ambiguity and develops arguments for using one or another method.

Conclusions and Tracks for Future Research

An important theoretical advantage of the two indirect approaches (PRCA and CA) is that their results are derived from actually perceived performances instead of using hypothetical experiences. For example Brandt and Scharioth [1998] state that SM is "based on how people say they would respond to a given level of attribute performance." These attitudinal reactions could be different from satisfaction judgments, which are per definition post experience (one cannot be satisfied with a hotel room without having stayed there, for example). For these reasons, it is questionable from a conceptual perspective, whether the results obtained from SM are meaningful in a satisfaction context. SM might nevertheless be used to gain initial insights into the meaning of new product or service features.

PRCA and SM allow categorizing each single attribute independently from other attributes (absolute categorization) while CA-categorizations (as well as DIM) depend among other things on the type and the number of the considered attributes (relative categorization). An advantage of SM is that it allows predicting even on an individual level to which category an attribute belongs, while DIM, PCRA as well as CA are based on the hypothesis of homogeneity of the sample. To strengthen their meaningfulness, these three methods should be applied as prognostic tools to clearly pre-determined customer segments. Such segments are nowadays often accessible through Data Warehouse applications within Customer Relationship Programs. Attribute categorizations could be worked out for well-chosen segments and be transmitted into actions. But also bench marking customer segments through attribute categorizations provides fruitful implications.

Some criticism can be addressed to SM as collecting the required data is cumbersome if many attributes are considered. In contrast, collecting data for DIM, PRCA as well as for CA is easily implemented, even for numerous attributes. A further critical point is that the proposed rules for SM-categorizations (Table 1) are arbitrary and lack solid theoretical

reasoning. But also DIM-categorizations are somewhat arbitrary, as the applied rules are defined on the basis of the results' distribution and not on theoretically developed arguments (in this study mean scores of 'stated importance' and 'derived importance' were used to separate four categories). PRCA, as a parametric method, suggests clarity for attribute categorizations because levels of significance are considered; however, the likelihood of finding significant differences is directly related to sample size. In contrast, CA is based on the idea that an attribute belongs to a certain category if its factor scores are lower or higher than the scores of each of the two satisfaction levels. This brings about some clarity and comparability of the results because the frontiers between attribute categories are determined a priori and independent of any sample size.

PRCA and CA include somewhat arbitrary decisions associated with coding: PRCA requires three modalities of performance and CA requires two modalities of performance and two modalities of CS. This study used rating scales and the median of the obtained answer distributions was used as a statistical criterion for recoding performance into two classes. It is obvious that attribute categorizations could change if performance and CS were coded differently. One might even argue that almost any result could be produced as a function of the applied coding scheme. This problem could be solved by replacing interval scales for performance and CS (in the case of CA) with ordinal scales that measure the modalities directly. The advantage is then that the resulting attribute categorizations depend clearly on theoretical reasons that can be developed a priori.

These arguments suggest six criteria for evaluating how the four considered methods work and how their results can be interpreted. These criteria are summed up in Table 8.

Take in Table 8

This article addresses the concept of attributes variant and invariant weight in relation to overall customer satisfaction. The empirical study shows that similar approaches for

attribute categorizations lead to different results. This brings to light qualities and possible applications of the considered techniques. Further efforts should take into account that these methods work fundamentally different. Developing suggestions as to how the construct validity of the methods can be assessed would be a fruitful direction for future research.

References

Anderson, E.W. and V. Mittal, 2000, 'Strengthening the Satisfaction-Profit Chain', *Journal of Service Research*, Vol. 3, No. 2, November, pp. 107-120.

Berger, C., R. Blauth, D. Boger, C. Bolster, G. Burchill, W. Du Muchel, F. Pouliout, R. Richter, A. Rubinoff, D. Shen, M. Timko, D. Walden, 1993, 'Kano's method for understanding customer-defined quality', *The Journal of the Japanese Society for Quality Control*, Fall, pp. 3-35.

Brandt, D.R., 1988, 'How service marketers can identify value-enhancing service elements', *The Journal of Services Marketing*, Vol. 2, No. 3, pp. 35-41.

Brandt, D.R. and J. Scharioth, 1998, 'Attribute life cycle analysis. Alternatives to the Kano-method', in *51ᵗʰ ESOMAR-Congress*, pp. 413-429.

Cadotte, E.R. and N. Turgeon, 1988, 'Dissatisfiers and satisfiers. Suggestions from consumer complaints and compliments', *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Vol. 1, pp. 74-79.

Engel, J.F., D.T. Kollat, R.D. Blackwell, 1968, *Consumer Behavior*, New York.

Hausknecht, D.R., 1990, 'Measurement scales in consumer satisfaction/dissatisfaction', *Jounal of Consumer Satisfaction/Dissatisfaction and Complaining Behavior*, Vol. 3, pp. 1-11.

Herzberg, F., B. Mausner, B.B. Snyderman, 1959, *The motivation to work*, 2ⁿᵈ ed., New York.

Johnston, R., 1995, 'The determinants of service quality. Satisfiers and dissatisfiers', *International Journal of Service Industry Management*, Vol. 6, No. 5, pp. 53-71.

Kahneman, D. and D.T. Miller, 1986, 'Norm theory. Comparing reality to its alternatives', *Pychological Review*, Vol. 93, No. 2, pp. 136-153.

Kano, N., 1984, 'Attractive quality and must-be quality', *The Journal of the Japanese Society for Quality Control*, April, pp. 39-48.

Leavitt, C., 1977, 'Consumer satisfaction and dissatisfaction. Bi-polar or independent?', in H.K. Hunt, (ed.), *Conceptualization and measurement of consumer satisfaction and dissatisfaction*, Cambridge, pp. 132-149.

Llosa, S., 1997, 'L'analyse de la contribution des éléments du service à la satisfaction: Un modèle tétraclasse', *Décisions Marketing* n°10, pp. 81-88.

Llosa, S., 1999, 'Contributions to the study of satisfaction in services', *AMA SERVSIG Service Research Conference* 10-12 April, New Orleans, pp.121-123

Maddox, R.N., 1981, 'Two-factor theory and consumer satisfaction. Replication and extension', Journal *of Consumer Research*, Vol. 8, pp. 97-102.

Martilla, J. A., J.C. James, 1977, 'Importance Performance Analysis', *Journal of Marketing,* Vol. 41 January, pp. 77-79.

Matzler, K., H.H. Hinterhuber, F. Bailom, E. Sauerwein, 1996, 'How to delight your customers', *Journal of Product and Brand Management*, Vol. 2, pp. 6–17.

Mazur, G.H. 2001, 'QFD and design for Six Sigma: Customer driven DFSS', *45th Congress of the European Organization for Quality (EOQ)*, Istanbul.

McGill, A., Iacobucci, D. 1992, 'The role of post-experience comparison standards in the evaluation of unfamiliar services', in J. Sherry, B. Sternthal, (eds.), *Advances in Consumer Research*, Vol. 19, pp. 570-578.

Mittal, V., W.T. Ross, P.M. Baldasare, 1998, 'The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions', *Journal of Marketing*, Vol. 62, January, pp. 33-47.

Ngobo, P.V., 1997, 'The Standards Issue: An accessibility-Diagnosticity Perspective', *The Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Vol. 10, pp. 61-79.

Oliver, R.L., 1977, 'The effects of expectations and disconfirmation on post-exposure product evaluations', *Journal of Applied Psychology*, pp. 246-250.

Oliver, R.L., 1997, *Satisfaction. A behavioral perspective on the consumer*, McGraw-Hill International Editions.

Rust, R.T. and A.J. Zahorik, 1993, 'Customer satisfaction, customer retention, and market share', *Journal of Retailing*, Vol. 69, pp. 192-215.

Sampson, S. E. and M.J. Showalter, 1999, 'The performance-importance response function: Observations and implications', *The Service Industries Journal*, Vol. 19, pp. 1-25.

Swan, J. and J. Combs, 1976, 'Product Performance and Consumer Satisfaction. A new concept', *Journal of Marketing*, Vol. 40, April pp. 25-33.

Szymanski, D.M., and D.H. Henard, 2001, 'Customer satisfaction: A meta-analysis of the empirical evidence', *Journal of the Academy of Marketing Science*, Vol. 29, No. 1, pp. 16-35.

Vanhoof, K. and G. Swinnen, 1998, 'Attribute importance - assessing nonlinear patterns of factors contributing to customer satisfaction', *ESOMAR Publication Series*, Vol. 204, pp. 160-171.

Vavra, T.G., 1997, *Improving your measurement of customer satisfaction. A guide to creating, conducting, analyzing, and reporting customer satisfaction measurement programs*, Milwaukee.

Venkitaraman, R.K and C. Jaworski, 1993, 'Restructuring customer satisfaction measurement for better resource allocation decisions: an integrated approach', *Fourth Annual Advanced Research Techniques Forum of the American Marketing Association*, Chicago, June.

Voss, G.B., A. Parasuraman, D. Grewal, 1998, 'The Roles of Price, Performance, and Expectation in Determining Satisfaction in Service Exchanges', *Journal of Marketing*, Vol. 62, pp. 46-61.

Westbrook, R.A. and R.L. Oliver, 1981, 'Developing better measures of consumer satisfaction. Some preliminary results', Monroe, K.B. (ed.), *Advances in Consumer Research*, Vol. 8, pp. 94-99.

Westbrook, R.A. and M.D. Reilly, 1983, 'Value-percept disparity. An alternative to the disconfirmation of expectations theory of satisfaction', in R.P. Bagozzi, A.M. Tybout, (eds.), *Advances in Consumer Research*, Vol. 10, pp. 256-261.

Appendix

A) Satisfaction measures:

1. How satisfied are you with the service of the insurance consultant you met recently? (very satisfied/ satisfied / neither satisfied nor dissatisfied/ very dissatisfied);

2. How would you describe your emotions towards this service encounter (very pleasant/ pleasant/ neither pleasant nor unpleasant/ very unpleasant);

3. If a good friend of yours needed insurance consultations, would you recommend that consultant? (surely/ maybe yes/ don't know/ maybe no/ certainly not);

4. Would you again engage the same insurance consultant you recently met? (surely/ maybe yes/ don't know/ maybe no/ certainly not)

B) Performance measures:

5-point rating scales: "entirely true/ mostly true/ neither, nor/ mostly wrong/ entirely wrong", used for 20 attributes formulated as assertions:

1. (Time) The appointment fit in my time schedule;

2. (Punctuality) The consultant was punctual;

3. (Reputation) I know him for a long time;

4. (Clothes) He was appropriately dressed;

5. (Expression) He expressed himself appropriately;

6. (Politeness) He was polite and kind;

7. (Presentation) He presented things clearly;

8. (Passion) He took time for listening and was patient;

9. (Situation) He took into account my personal situation (partner, family etc.);

10. (Proposition) He made meaningful suggestions to me;

11. (Explication) He explained the details clearly;

12. (Preparation) He was well-prepared well the meeting;

13. (Objectivity) The advice proceeded objectively;

14. (Process) He informed me about the process in case of a claim;

15. (Duration) The conversation didn't last too long;

16. (Utility) He clearly showed how to use the products;

17. (Obtrusiveness) He wasn't obtrusive;

18. (Question) He could answer all my questions to my satisfaction;

19. (Information) He left meaningful and informative material to me;

20. (Gift) He gave me a little present (e.g. ballpoint pen).

Tables and figures:

Figure 1.

Variant and Invariant Impact of Attribute Performances on Overall CS*

*Adapted from Llosa [1999]

Figure 2.
Attributes scores on one factorial axis obtained through Correspondence Analysis

Table 1.
Kano's [1984] Evaluation Rules

		Dysfunctional question: If x performs badly, how do you feel?				
		1	2	3	4	5
	1	Q	A	A	A	O
	2	R	I	I	I	M
Functional question:	3	R	I	I	I	M
If x performs well, how do you feel?	4	R	I	I	I	M
	5	R	R	R	R	Q
Answers: 1 = I like it that way; 2 = It must be that way; 3 = I am neutral; 4 = I can live with that way; 5 = I dislike it that way						
Interpretation of answer combinations: A = Attractive; O = One dimensional; M = Must-be; I = Indifferent; R = Reverse; Q = Questionable						

Table 2.
Recoding Attribute Performance into dummy variables

Performance of attribute x	D1	D2
P: positive, better than expected	1	0
E: equal, expectations are met	0	0
N: negative, worse than expected	0	1

Table 3.
Categorization rules for PRCA

An attribute is	Basic if $(CS|N) < (CS|E)$ and $(CS|E) \approx (CS|P)$
	One-dimensional if $(CS|N) < (CS|E)$ and $(CS|E) < (CS|P)$
	Attractive if $(CS|N) \approx (CS|E)$ and $(CS|N) < (CS|P)$
	Low impact if $(CS|N) \approx (CS|E)$ and $(CS|E) \approx (CS|P)$

Table 4.
Contingency Table

	Low CS	High CS	
attribute 1: low performance	$n1_{ij}$	$n1_{ij}$	$\sum=n$
attribute 1: high performance	$n1_{ij}$	$n1_{ij}$	
attribute 2: low performance	$n2_{ij}$	$n2_{ij}$	$\sum=n$
attribute 2: high performance	$n2_{ij}$	$n2_{ij}$	
…	…	…	$\sum=n$
Overall satisfaction	n_{ij}	n_{ij}	$\sum=n$

Table 5.
Synthesis of Attribute Categories

	Impact on CS			
	Variant		Invariant	
Author(s)	Strong if performance is low, low if performance is high	Weak if performance is low, high if performance is high	Always strong	Always weak
Herzberg *et al.* [1959]	Hygiene	Motivator	-	-
Kano [1984]	Must-be	Attractive	One-dimensional	Indifferent
Cadotte and Turgeon [1988]	Dissatisfier	Satisfier	Critical	Neutral
Brandt [1988]	Minimum requirement	Value enhancing	Hybrid	Unimportant as determinant
Brandt and Scharioth [1998]	Basic	Attractive	One-dimensional	Low impact
Venkitaraman and Jaworski [1993]	Flat	Value-added	Key	Low
Llosa [1997, 1999]	Basic	Plus	Key	Secondary

Table 6.
Four Relevant Attributes

Attribute	Description
Situation	Consultant takes into account client's personal situation.
Information	Consultant left information (leaflet etc.)
Presentation	Consultant presents clearly
Gift	Consultant left a gift (pen, diary etc.)

Table 7.
Results

Attribute	SM	PRCA	DIM	CA
Presentation	Dissatisfier	Critical	Critical	Critical
Situation	Dissatisfier	Critical	Critical	Critical
Information	Dissatisfier	Critical	Neutral	Dissatisfier
Gift	Satisfier	-	Neutral	Satisfier
Politeness	-	-	Satisfier	Dissatisfier
Punctuality	-	Satisfier	Neutral	Neutral
Time	-	Satisfier	Dissatisfier	Satisfier
Fame	-	Dissatisfier	Neutral	Neutral
Clothes	-	Satisfier	Neutral	Satisfier
Expression	-	Satisfier	Neutral	Satisfier
Passion	-	Critical	Satisfier	Critical
Proposition	-	Critical	Satisfier	Dissatisfier
Explications	-	Critical	Critical	Dissatisfier
Preparation	-	Critical	Satisfier	Critical
Objectivity	-	Critical	Critical	Satisfier
Process	-	Satisfier	Satisfier	Dissatisfier
Duration	-	Satisfier	Neutral	Neutral
Utility	-	Critical	Satisfier	Dissatisfier
Obtrusiveness	-	Satisfier	Critical	Critical
Questions	-	Critical	Critical	Critical

Table 8.
Criteria for Distinguishing between four Methods for Attribute Categorizations

	DIM	SM	PRCA	CA
1. Respects the nature of CS as a post experience evaluation (no performance simulations)	Yes	No	Yes	Yes
2. Permits categorizing each single attribute without taking into account other attributes ('absolute categorization')	No	Yes	Yes	No
3. Works on an individual level (not only on an aggregated level)	No	Yes	No	No
4. Data can be easily collected (large numbers of attributes can be studied)	Yes	No	Yes	Yes
5. Rules for attribute categorizations are theoretically developed (not arbitrarily chosen)	No	No	Yes	Yes
6. Clearly established rules for attribute categorizations (always the same rules can be applied, independent of the obtained results or the sample size)	No	Yes	No	Yes