# KOHONEN SELF-ORGANIZING MAPS: A NEURAL APPROACH FOR STUDYING THE LINKS BETWEEN ATTRIBUTES AND OVERALL SATISFACTION IN A SERVICES CONTEXT

## Anne-Françoise Audrain-Pontevia Rouen School of Management, France

## ABSTRACT

This research aims at analyzing and understanding the attributes - overall satisfaction links (A - OSL) for a service. To date, marketing managers tend to assume that these links are linear, even though scholars have for at least two decades pointed out that they can often be nonlinear as well as asymmetric (Kano, Seraku, Takahashi and Tsuji 1984; Anderson and Mittal 2000). Blindly assuming that these links are linear may lead to serious mistakes in estimating the attribute levels which trigger the highest degree of targeted consumers' overall satisfaction (TCOS). In this article, we explore the A - OSL relationship by using a powerful neural network methodology: Kohonen Self-Organizing Maps (KSOM). KSOM have the ability to infer the functions describing A - OSL from data. This methodology also classifies the input data in relation to prototypes on a topological map by using a Euclidian distance criterion.

The analysis of a database for a utility company with the KSOM methodology suggests the existence of five main A - OSL patterns:

- i) Linear A OSL
- ii) attributes having increasing returns on TCOS
- iii) attributes having decreasing returns on TCOS
- iv) attributes having increasing returns on both TCOS and consumers overall dissatisfaction; and
- v) attributes reflecting an assimilationcontrast effect.

## INTRODUCTION

For more than thirty years, consumer satisfaction has been shown to be a key construct in marketing. This is one reason why managers invest heavily in consumer satisfaction programs. They try to identify product or service attributes which lead to high levels of customer overall satisfaction (TCOS). To date, little has been published on how to go about examining the attributes - overall satisfaction's links (A - OSL). even though it has been known for some time that understanding the nature of these links should help managers to optimise customers' overall satisfaction. Most managers consider the A - OSL to be linear in nature. Hence, there may often be an erroneous estimation of a product's or service's attributes leading to high level of TCOS (Anderson and Mittal 2000). Scholars (Swan and Combs 1976; Kano, Seraku, Takahashi and Tsuji 1984; Anderson and Mittal 2000, among others) have suggested that these links may not always be linear, but these suggestions have been based largely on intuition and subjective analysis, and have not resulted as fruits of the labor of empirical research.

The approach described in this article investigates the A - OSL by using an exploratory, yet powerful methodology called Kohonen Self-Organizing Maps (KSOM). KSOM is one of the approaches based on neural networks methods. It can define attributes - overall satisfaction prototypes which characterize the A - OSL. It also classifies the data in relation to these prototypes on the basis of a Euclidian distance criterion. Finally, KSOM provides a topological map (or grid), on which prototypes are related one to another.

Consumer satisfaction has been shown to be crucial for companies because it has been

learned that in general, a satisfied customer is more loyal, buys more (Reichheld and Sasser 1990; Anderson and Sullivan 1993), is less sensitive to product/service prices (Fornell, Johnson, Anderson, Cha, Everitt and Bryant 1996), buys other products/services from the same company (Fornell 1992) and generates positive word-of- mouth (Anderson 1998). Gruca and Rego (2005) reveal that satisfaction plays a key role in building companies shareholder value. Fornell, Mithas, Morgeson and Krishnan (2006) also show that firms which do better in terms of satisfying customers tend to generate a superior return on investment and yield higher profits. Indeed, it is for these reasons that companies seek to develop products or services which maximize customers overall satisfaction.

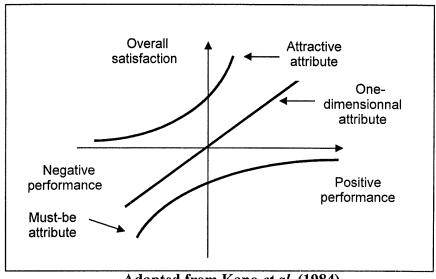
Managers should especially want to identify the product or service attributes which lead to high levels of consumers' satisfaction. To do so, marketing researchers typically measure consumers' satisfaction overall, as well as their evaluations of salient attributes. Then they often estimate attribute-importance using a linear regression model. By using this popular approach. attributes have linearly increasing returns on consumers' overall satisfaction. Indeed, whether it be a conscious view or not, today most empirical considerations of the role of attribute satisfaction on overall satisfaction judgments are considered linear in nature. However, a few scholars have suggested or shown that product or service attributes may not impact overall satisfaction in a linear and symmetric way. Relying on Herzberg, Mausner and Snyderman's bi-factorial theory. Swan and Combs (1976) show for clothes the existence of attributes contributing to an increase in consumer overall satisfaction ('expressive dimension'), while others contribute to a decrease overall dissatisfaction in ("instrumental dimension"). They conclude, in line with Herzberg et al.'s bi-factorial theory, that consumer satisfaction is not a uni-polar construct, but a bipolar one. Satisfaction and dissatisfaction would not compensate since they are two different, independent constructs.

A few scholars in the marketing and quality management fields are currently offering the opinion that A - OSL relationships are not always linear in nature. Swan and Combs' (1976) research has been replicated and confirmed by Gnoth and Hilt (2000) (although admittedly there controversy surrounding this replication (Maddox 1981; Yi 1990). In studying consumer satisfaction with consumer goods, Kano, Tsuji, Seraku and Takahashi (1984) revealed the existence of four main attributes: a) attributes with increasing returns on overall satisfaction (they call them 'attractive' attributes); b) attributes with decreasing returns on overall satisfaction (they call them 'must-be' attributes; c) attributes contributing to consumers satisfaction in a linear way (they call them 'one-dimensional' attributes) and d) attributes hardly contributing to overall satisfaction ('secondary' attributes) (refer to Figure 1 and Table 1, below). To Kano et al., attractive attributes lead to high levels of satisfaction. They can create a competitive advantage. On the other hand, 'must-be' attributes trigger high levels of dissatisfaction. For this reason, managers should first try to reach a minimum level on these attributes before concentrating on one-dimensional and attractive attributes. Kano et al.'s research has also been confirmed by others (Lee and Newcomb, 1997; Brandt, 1988; Vanhoof and Swinnen, 1998).

More recently, relying on Kahneman and Theory (1979, Tverski's Prospect Anderson and Mittal (2000), and Mittal, Ross and Baldasare (1998) suggest and show that the A -OSL linkages are often asymmetric and nonlinear. Prospect theory predicts that a loss weighs more heavily than a gain on people's preferences (see Figure 2 and Table 1, below). It also forecasts that performance has decreasing returns on consumers' preferences. Transposed to consumer satisfaction, one can infer that a negative performance weighs more than a positive one. Prospect theory also suggests that performance has decreasing returns on overall satisfaction (again, see Figure 2 and Table 1), suggesting that the A/OS function is an 'S' shape function.

Figure 1

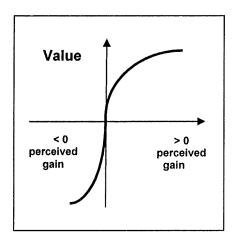
Functions Characterizing Attributes-Overall Satisfaction Links

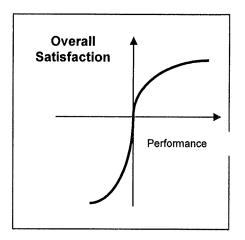


Adapted from Kano et al. (1984)

Figure 2

Transposition of Kahneman and Tversky's (1979) Prospect Theory to the Consumer Satisfaction Construct

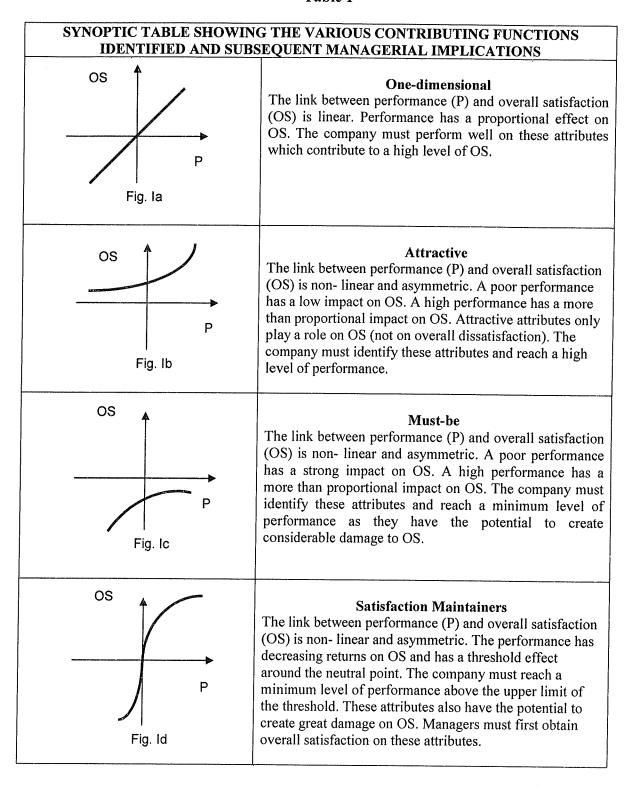




Based on this theory, Anderson and Mittal (2000) have suggested the existence of both 'linear' and 'asymmetric non-linear' A - OSL: 'Satisfaction enhancers' (equivalent to 'attractive' attributes)

and 'satisfaction maintainers', having decreasing returns on overall satisfaction and dissatisfaction (i.e. describing an 'S' shaped, sigmoid function, as illustrated in Figure 2 and ld in Table 1).

Table 1



In conclusion, a method aimed at identifying A – OSL relationships, be they linear or non-linear,

should be of interest to both practitioners and scholars.

## Research Objectives and Methodology

This article summarizes research aimed at exploring the A - OSL linkage and uses a neural network methodology: Kohonen Self-organizing Maps (KSOM). Artificial neural networks can be seen as a paradigm encompassing various sophisticated modelling techniques developed in ways analogous to the structure of the brain (Rojas 1996). What is common to these approaches is that they are based interconnected processing elements, neurons, that work together to produce an output function. Neural network approaches can be differentiated thanks to three elements: their neurons; their architecture, and learning law. The neurons themselves are characterized by the input data, the transfer function, and the output data. The neural network architecture refers to the way the neurons are connected together and to the number of neural layers. The learning law refers to the way the neural network learns from the data. It can either be supervised or unsupervised. When supervised, the system must be trained first: It requires input and output data to calibrate the network. It then 'learns' from the data. When unsupervised, the algorithm develops on the basis of a real data set and it does not require a training phase.

KSOM can legitimately be described as a clustering method relying on a competitive, unsupervised learning algorithm. It identifies a definite number of 'prototypes' characterizing the input data and classifies these data thanks to a Euclidian distance criterion. The input data are simple vectors. The result is a 'topological map' where prototypes are projected and organized so that each prototype is related one to another. For reason, KSOM can be seen as a this 'visualization' segmentation or clustering technique.

In summary, the KSOM algorithm can be described as follows. Once the number of clusters (i.e. the number of lines and columns of the grid) has been defined by the researcher, random vectors (called 'processing elements':  $w_i$ ) are initialized in each cell of the grid. These random vectors represent the first set of cluster prototypes. The competitive, unsupervised algorithm then

compares each input vector (x<sub>i</sub>) on a Euclidian distance criterion to each wi. It selects the closest processing elements and modifies it according to the following learning law':  $w_i^{\text{new}} = w_i^{\text{old}} + \alpha(x$ w<sub>i</sub> old) z<sub>i</sub>. In other words, it moves the processing vector a fraction α from its previous value to the entry vector x. z equals 0, when the processing element is not selected, and 1 when it is selected. The neighboring processing elements of the chosen wi are also modified in a similar fashion, but to a lesser extent. The algorithm iterates the process up to convergence. The final result is the topological map (two-dimensional grid) where adjacent 'prototypes' are related to each other and share common characteristics. KSOM therefore differs substantially from traditional clustering methods, such as k-means, for example, where clusters are mutually exclusive one from another. Besides, KSOM is a 'visualization technique', which is not the case for traditional clustering techniques. It is also robust to non-normality assumptions. Certainly, KSOM can be viewed as an alternative to cluster analysis (Curry, Davies, Evans and Philips 2003). Indeed, like K-means, it classifies data thanks to an unsupervised competitive algorithm that is based on a Euclidian distance criterion.

KSOM can also be seen as a substitute to factor analysis, since its main function is to map the input data from an n-dimensional space to a lower dimensions space, while maintaining the original topological space. Kiang and Kumar (2001) find that KSOM provide solutions superior to unrotated factor solutions and that it is more robust than factor analysis when the data are skewed (which is often the case when considering satisfaction data).

Overall, KSOM can be seen as a data reducter or as a classifier, but not as a predictive method such as regression analysis, for example. It presents several advantages to study the A - OSL structure. First, it infers a definite number of functions (called 'prototypes') characterizing the A - OSL from the data. Second it classifies the input data in relation to these prototypes. Third, it provides a managerial visual tool, a topological grid, to understand and analyze these functions. Fourth, KSOM is also able to deal with a great amount of data and is able to deal with skewed data such as satisfaction data.

#### The Data

Having discussed and defended Kohonen Self-Organizing Maps, an example is now presented based on secondary data: A six-year longitudinal data set developed to track consumer satisfaction for a utility (electric heating) company in France. In this survey, consumer overall satisfaction as well as detailed satisfactions were measured on four-point Likert scales. The salient attributes were initially identified by conducting formal content analyses of the transcripts from two different focus group interviews.

Consistent with KSOM requirements, entry vectors must first be defined. The vectors are the followings:  $(x_1^i; x_2^i; x_3^i; x_4^i; x_5^i; x_6^i)$ , where

 $x_j^i$  expresses the average global satisfaction evaluation for people declaring they are not at all satisfied (j=1), slightly dissatisfied (j=2), rather satisfied (j=3) or fully satisfied (j=4) with the attribute 'i'. Data, of course, were first standardized. The two secondary derivatives are also included in order to take into account function concavities.  $x_5^i$  and  $x_6^i$  represent the secondary derivative coordinates of the vector  $(x_1^i; x_2^i; x_3^i; x_4^i)$ . To estimate the classification quality, two criteria are invoked. The first one assesses the average distance between the entry classified vectors and the Kohonen defined prototype (Desmet 2001). It is called Projection Error (P.E.) and is defined as follows:

- EP = 
$$\frac{1}{n} \times \sum_{j=1}^{n} D_j$$
, with *n* the number of input vectors and

where 
$$D_{j} = \sqrt{\sum_{i=1}^{6} (x_{i} - w_{i})^{2}}$$
 and,

 $(x_1^i; x_2^i; x_3^i; x_4^i; x_5^i; x_6^i)$  represent the input vector coordinates,

 $(w_1^i; w_2^i; w_3^i; w_4^i; w_5^i; w_6^i)$  the coordinates of the vector code.

# A second indicator is hereby proposed which measures the average Intra-Class Distance (ICD):

- ICD = 
$$\frac{1}{n} \times \sum_{j=1}^{NVj} D_{Ij}$$
 where n is the number of clusters and,

- 
$$DI_j = \frac{1}{NVj} \sum_{j=1}^{NVj} D^j v^{ij}$$
 is the intra-class distance of vector  $\mathbf{w}^j$  and,

- 
$$D^{j}_{v_{i}} = \sqrt{\sum_{i=1}^{6} (W_{k}^{v_{i}} - W_{jk})^{2}}$$
,  $Nv_{j}$ : number of neighbors to the code vector  $w_{j}$  and,

- $(x_1^i; x_2^i; x_3^i; x_4^i; x_5^i; x_6^i)$  represents the coordinates of the input vector and,
- $(w_1^i; w_2^i; w_3^i; w_4^i; w_5^i; w_6^i)$  the coordinates of the vector code.

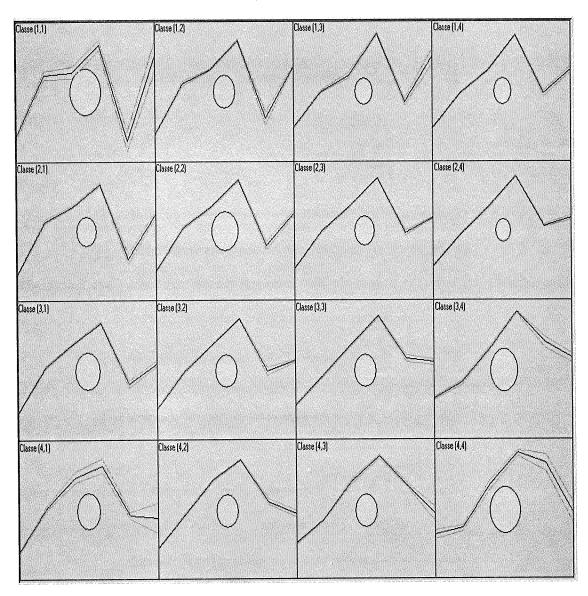
It is standard practice to consider the classification to be of good quality when the E.P. value is smaller than the I.C.D.

#### Results

We ran the KSOM algorithm on the above defined entry vectors using the 'Courboscope' software, designed by E.D.F. (Electricité de France) to classify consumption patterns. A Kohonen Self-Organizing Maps Procedure is also available on SAS as well as on

other software products. Analyses were run for  $2 \times 2$ ;  $3 \times 3$ ;  $4 \times 4$  and  $5 \times 5$  maps and the solution with the optimum Projection Error (P.E.) and Inter-Class Distance (I.C.D.) indicators was chosen in each instance. The P.E. and I.C.D. indices are the best for the  $4 \times 4$  map: The P.E. equals .209 and the I.C.D. .360. This reveals that, for the  $4 \times 4$  map, the intra-class distances are smaller than the average inter-class distances, which is not the case for the other maps. The results are portrayed in Figure 3, below

Figure 3

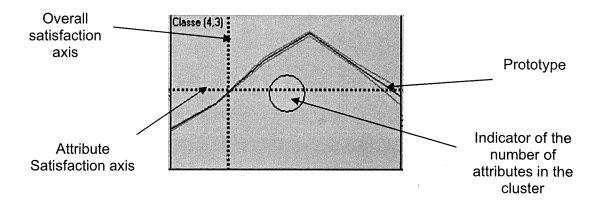


The above 4 × 4 Kohonen map suggests the existence of both linear and non linear attribute / overall satisfaction links (A - OSL).

Let us consider a definite cluster (cluster '4,3' from Figure 3). The horizontal axis represents the average attribute satisfaction level, and the vertical axis, the overall satisfaction level (Figure 4). The four first points define the cluster's prototype, (the third and fourth points account for secondary derivatives). The curve in

the middle represents the cluster's prototype. Around each prototype, two curves are representing +/- 1.5 the standard deviation. Like E.P. and I.C.D., they provide 'visual' information on the classification's quality: When these curves are close to the prototype, the classification is considered homogenous. Information on the cluster size (i.e. number of attributes classified) is also given by circle's size.

## FIGURE 4



Consistent with results summarized in Table 1, the prototype '4,3' and the related attributes are 'satisfaction maintainers'.

Overall, this Kohonen map reveals the existence of both linear and non-linear / asymmetric A - OSL. More precisely, it identifies:
a) Linear or quasi-linear prototypes, (accounting for 'one-dimensional' attributes; cf. cluster '3,3'; or cluster '2,4'); b) prototypes with increasing returns on overall satisfaction ('attractive attributes'; cf. cluster '3,4' for example); c) prototypes with decreasing returns on overall satisfaction ('must-be attributes'; cf. cluster '4,1' for example); d) prototypes with increasing returns on overall satisfaction and dissatisfaction ('satisfaction maintainers', cf. cluster '4,4' for example). The results also suggest the existence of attributes with increasing returns on overall

satisfaction and a threshold effect (cf. cluster '1,1'). These attributes present an indifference zone around the mean. Within this zone, the performance does not impact significantly the overall satisfaction. Outside, the performance strongly impacts customers overall satisfaction. In line with the marketing literature, we infer that this expresses an 'assimilation-contrast' effect of the performance on overall satisfaction. This is consistent with the research of Anderson (1973) and that of Woodruff, Cadotte and Jenkins (1983). Though these attributes were not taken into account in the typologies derived in this article, the results nevertheless underline the importance of analyzing A/OS links in order to optimize overall customer satisfaction.

# RESEARCH AND MANAGERIAL IMPLICATIONS

The first lesson for managers to learn is that they must not consider the functions relating consumer attribute evaluation to consumer overall satisfaction as always or even necessarily linear. Indeed, our results confirm that these relationships are often more complex. This research also confirms findings from previous studies relying either on Herzberg et al.'s bi-factorial theory or on Kahneman and Tversky's prospect theory. The typology that resulted from our methodology supports the existence of attributes ('must-be' or 'satisfaction maintainers') which mostly impact consumer's overall dissatisfaction, while others ('attractive' or 'assimilation-contrast') have more potential to trigger off very high levels of consumer overall satisfaction. The research also revealed the existence of "assimilation-contrast" attributes having a neutral zone around the mean point, in line with Anderson (1973) and Woodruff et al.'s (1983) results. This finding suggests that a performance around the mean (perhaps within 1 standard deviation) does not significantly impact consumers' overall satisfaction but that high or low performances (greater than plus or minus 1 standard deviation) dramatically impact overall satisfaction / dissatisfaction.

In general, the typology uncovered suggests three main managerial actions. First, it encourages managers to prevent dissatisfaction by focusing on 'must-be', 'assimilation-contrast' and 'satisfaction maintainers', each of which has the potential to cause great damage to overall satisfaction. A certain minimum level performance has to be reached on these attributes. Managers should then pay attention to 'one dimensional' and 'attractive' attributes which can contribute to high levels of overall satisfaction. In particular, 'attractive' attributes can create very high levels of satisfaction. That is why managers should try to detect these attributes and strongly invest in them. Presence of these attributes (and the emphasis placed on them as connected to the company's brand or service marketing communication tactics) can create a competitive advantage. It also should be noticed that assimilation-contrast attributes have increasing returns on overall satisfaction. It is therefore

important to detect the minimum required level that enables performance to strongly impact satisfaction.

This article also points out the usefulness of KSOM as a tool to characterize A - OSL and to classify the contribution of attributes to overall satisfaction. This neural network approach defines prototypes of the A - OSL on the basis of satisfaction data measured on Likert or other scales that have interval scale properties, and provides managers with a typology diagram where prototypes are related. Also of interest is the fact that KSOM, while defining clusters with neighbouring properties, enables the tracking of the attributes / overall satisfaction functions over time. Moreover, since it is possible to project new entry vectors on a previously defined grid, KSOM enables benchmarks over time or across products / services. Finally, KSOM is a very powerful tool to visualise the A - OSL linkage, as it is not sensitive to data non-normality and it can deal with very large databases. It therefore should be very useful to managers seeking to identify the levers of consumer satisfaction.

## **CONCLUSION**

This article demonstrates how the KSOM neural network approach can be used to explore. analyse and track consumer satisfaction. This methodology provides managers with tools to understand and identify product / service attributes which lead to high levels of overall consumer satisfaction. More precisely, this article discussed an application of KSOM for a service and it revealed the existence of six general types or categories of attributes: 'Attractive'; 'onedimensional'; 'must-be'; 'assimilation-contrasts'; 'satisfaction maintainers' and 'secondary'. More work is now needed to replicate these findings for others product types, for both the consumer goods and B2B sectors. More work is also needed to understand why a particular attribute belongs to a specific attribute category type.

## REFERENCES

- Anderson, Eugene W. and Mary W. Sullivan (1993), "The antecedents and consequences of customer satisfaction for firms," *Marketing Science*, 12 (2), 125-143.
- Anderson, Eugene W. and Vikas Mittal (2000), "Strengthening the satisfaction-profit-chain," *Journal of Service Research*, vol. 3, 107-120.
- Anderson, Eugene W. (1998), "Customer satisfaction and word of mouth," *Journal of Service Research*, 1 (1), 5-17.
- Anderson, Rolph E. (1973), "Consumer dissatisfaction: the effects of disconfirmed expectancy on perceived product performance," *Journal of Marketing Research*, 10 (1), 38-44.
- Brandt, Dominique R. (1988), "How service marketers can identify value-enhancing service elements," *The Journal of Services Marketing*, 2 (3), 35-41.
- Curry, Bruce, Davies, Fiona, Evans, Martin, Moutinho, Luiz and Philips, Paul (2003), "The Kohonen Selforganizing map as an alternative to cluster analysis: An application to direct marketing," The Market Research Society, 45 (2), 191-211.
- Desmet, Pierre (2001), "Buying behavior study with basket analysis: Pre-clustering with a Kohonen map," European Journal of Economics and Social Systems, 15 (2), 17-30.
- Fornell, Claes, Sunil Mithas, Forrest V. Morgeson and M.S. Krishnan (2006), "Customer satisfaction and stock prices: High returns, low risk." *Journal of Marketing*, 70 (1), 3-14.
- Fornell, Claes, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha and Barbara Everitt Bryant (1996), "The American Customer Satisfaction Index: Nature, purpose and findings," *Journal of Marketing*, 60 (4), 7-18.
- Gnoth, Juergen and S. Hilt (2000), "The hierarchy of cognitive and affective effects on satisfaction formation," 29<sup>th</sup>. EMAC Conference, Rotterdam, CD Rom.
- Gruca, Thomas S. and Lopo L. Rego (2005), "Customer satisfaction, cash flow and shareholder value," *Journal of Marketing*, 69 (3), 115-130.
- Herzberg, Frederik, Bernard Mausner and Barbara B. Snyderman (1959), *The Motivation to Work* New York (NY): John Wiley.

Send Correspondence Regarding This Article To:

Anne-Françoise Audrain-Pontevia Rouen School of Management Bd. André Siegfried - BP 188 76825 Mont-Saint-Aignan cedex - France Tel: 33 2 32 82 5 69 / Fax: 33 2 32 82 5 69 E-mail: anne-francoise.audrain@groupe-esc-rouen.fr

- Kano, Noriaki Nobuhiko Seraku, Fumio Takahashi and Shinichi Tsuji (1984), "Attractive quality and must-be quality," The best on quality, ASQC Quality Press.
- Kahneman, Daniel and Amos Tversky (1979), "Prospect theory: An analysis of decision under risk," Econometrica, 47 (2), 313-327.
- Kahneman, Daniel and Amos Tversky (1984), "Choices, values and frames". *American Psychologist*, 39, 341-350.
- Kiang Melody Y. and Ajith Kumar (2001), "An evaluation of Self-Organizing Map Networks as a robust alternative to factor analysis in data mining application," *Information Systems Research*, 12 (2), 177-194.
- Kohonen, Teuvo (1995), Self-Organizing Maps, Springer-Verlag, Berlin,
- Lee, M.C. and Newcomb J.F. (1997), "Applying the Kano methodology to meet customer requirements: Nasa's microgravity science program," *Quality Management Journal*, 4 (3), 95-109.
- Maddox, R Neil (1981), "Two-factor theory and consumer satisfaction: Replication and extension," Journal of Consumer Research, 8 (2), 97-102.
- Mittal, Vikas, W.T. Ross and P.M. Baldasare (1998), "The Asymmetric impact of negative and positive attribute level performance on overall satisfaction and repurchase intentions," *Journal of Marketing*, 62 (1), 33-47.
- Rojas, Raul (1996), Neural Networks: A systematic introduction, Springer-Verlag, Berlin.
- Rust, Roland.T., Anthony J. Zahorik and Timothy L. Keiningham (1995), "Return on quality (ROQ): Making service quality financially accountable," *Journal of Marketing*, 59 (2), 58-70.
- Swan, John E. and Linda J. Combs (1976), "Product Performance and Consumer Satisfaction: A New Concept," *Journal of Marketing*, 40 (2), 49 67.
- Vanhoof, Koen and Gilbert Swinnen (1998), "Attribute performance: Assessing non linear patterns of factors contributing to customer satisfaction," Proceedings of the EKSOMAR Annual Conference.
- Yi, Youjae (1990), "A critical review of customer satisfaction," In *Review of Marketing*, ed. V.A. Zeithaml, Chicago, American Marketing Association.