

DIFFERENT COMPARISON STANDARDS AS DETERMINANTS OF SERVICE QUALITY

Veronica Liljander, Swedish School of Economics and
Business Administration, Helsinki, Finland
Tore Strandvik, Swedish School of Economics and
Business Administration, Helsinki, Finland

ABSTRACT

A widely accepted view on how satisfaction and service quality are formed is that they are the result of some type of comparison process, where the perceived performance is evaluated against a comparison standard. Different comparison standards are used in this study on educational services and are related to satisfaction with the service. The authors question the difference between satisfaction and service quality as it has been presented in the service quality literature. Several service quality model specifications are compared. Results show that 1) inferred disconfirmation models are inferior to all other models, 2) performance alone is a strong indicator of satisfaction, and 3) performance and direct disconfirmation of the brand norm explains satisfaction best, 4) attribute-specific measures of comparison standards may be added for diagnostic reasons, but they add little to the explanation of satisfaction and 5) the best brand norm, which most resembles the normative standard used in Servqual, seem to be the least effective comparison standard.

INTRODUCTION

According to the dominating paradigm both satisfaction and service quality are the outcome of a comparison between performance and some standard of comparison or reference point. Within service quality research the standard reference point has been excellent service while the comparison standard in consumer satisfaction research has been the consumer's predictive brand expectations. Although the issue of which comparison standard customers use as their reference point has both theoretical and practical consequences, this question has not been addressed in service quality research until recently. In the consumer satisfaction literature it has been elaborated by Cadotte, Woodruff and Jenkins (Cadotte et al. 1983; Woodruff et al. 1983;

Cadotte et al. 1987; Woodruff et al. 1991).

The way that the standard is defined in the service quality model gives an indication of the scope of the model. The most limited view is based on a dyadic conception of the relation between the customer and the firm. An example of this view is Grönroos' conceptual model of perceived service quality (Grönroos 1982). In his model the customer's evaluation of service quality is based on a comparison between predictive expectations and perceptions of performance. The nature of expectations in the model is not explicitly defined but should be interpreted as predictive expectations according to the author. The evaluation is thus based only on experience with the focal service on a particular occasion. Parasuraman et al.'s model, which prescribes normative expectations as the comparison standard broadens the scope, as it can be argued that these expectations may include experience with the product group and with similar products, or even the whole cultural context. Both are, however, static models and do not explicitly take into account the choice situation. If the customer has a choice, the relevant comparison standard in a model should also include expectations of competitors performance, not only expectations about the performance of the focal firm. In consumer satisfaction studies this type of effect is partly considered in comparison standards like product type, brand and best brand norms. Although any of these may be included in normative expectations, they have not been used as separate constructs in service quality studies.

It has been suggested by some researchers that the difference between service quality and satisfaction lies in the nature of performance evaluations and the nature of expectations, which results in a definition of service quality as the sum of several dis/satisfactory transactions (Parasuraman et al. 1988; Zeithaml et al. 1993). This does not, however, solve the problem with services which are only consumed once (like education and surgical operations) or services that

are consumed for the first time. In these situations it would not be possible to give an evaluation of many transactions, and the only remaining difference between service quality and satisfaction would be the nature of expectations. It would not seem logical that a comparison between predictive expectations and perceptions of the course would yield satisfaction, while a comparison between normative expectations (excellent service) and perceptions of the course would result in service quality. As our study concerns educational services that are consumed only once, we will not make a difference between satisfaction and service quality in this paper, but assume that they are equivalent.

The purpose of our study is to investigate the impact of a number of comparison standards suggested in the consumer satisfaction literature on perceived service quality of an educational service. Different models including predictive expectations and experience-based norms will be presented and compared as to their ability to explain service quality. The contribution of the study is to expand the discussion of service quality by introducing empirical results concerning other comparison standards than what is normally used.

COMPARISON STANDARDS

Many different comparison standards have been proposed but based on different perspectives. No simple classification scheme encompassing the different standards have emerged. Woodruff et al. (1991) suggested a number of dimensions which could be used to classify standards, including type/source, nature of experience, level, and perceptual distinctiveness. They did, however, not combine these to a classification scheme, but consider the development of such a scheme as an area for further research. In our study we will investigate comparison standards that in the literature have been called predictive expectations and experience-based norms, therefore the discussion on past research will be limited to these standards.

The most used comparison standard within consumer satisfaction research has been predictive expectations (Woodruff et al. 1991). Other comparison standards have been suggested but have been only limitedly used. In service quality

research the discussion on standards has been almost non-existent. There are mainly two approaches to the issue. One approach has been to ignore the question of the nature of expectations (Grönroos 1990). The other approach is represented by the research team Parasuraman, Zeithaml and Berry and has evolved from first ignoring the question (Parasuraman et al. 1985) to then defining expectations as desired service, i.e. what the service should be, or service excellence (e.g. Parasuraman et al. 1988), and later as adequate service and desired service (Zeithaml et al. 1993). Several other standards have been suggested and used in satisfaction research, including different experience-based norms, an ideal, a minimum tolerable, a deserved standard and equity. (Woodruff et al. 1991). Expectations in the service quality literature has been used with a different, often implicit, meaning by different authors, thus making it difficult for the reader to evaluate the results.

Predicted service. Most of the studies within the disconfirmation paradigm have defined the reference point as predictions of future product performance (Day 1983; Oliver 1980a; 1980b; Olson and Dover 1979; Swan 1988; 1992). It is usual in satisfaction studies to measure predictive expectations before the consumption of the product, both in field studies (Cadotte et al. 1983; Oliver and Bearden 1983; Swan 1977; 1988; Swan and Trawick 1982) and experiments (Churchill and Surprenant 1982; Tse and Wilton 1988). Several studies have showed a positive relation between inferred disconfirmation of predictive expectations and satisfaction with a service (Cadotte et al. 1983; Liljander and Strandvik 1992a; Swan 1977; 1988). Most satisfaction studies on goods and services have, however, also included a direct isconfirmation measure (Bolfing and Woodruff 1988; Churchill and Surprenant 1982; Cadotte et al. 1983; Tse and Wilton 1988), as it was suggested by Oliver (1980a) that disconfirmation has an independent additive effect on satisfaction. Direct disconfirmation has thus been used as an additional variable to predictive expectations and performance, not as an alternative for inferred disconfirmation. Predictive expectations have been found to have a direct positive effect on satisfaction for goods (Bolfing and Woodruff 1988;

Churchill and Surprenant 1982; Tse and Wilton 1988) and services (Prakash and Lounsbury 1984; Swan 1988). Oliver and Bearden (1983), however, found no relation between expectations and satisfaction.

Norms based on experience. The consumer may have experience with other brands than the focal brand, and this may affect his/her perceptions of the performance of the focal brand. Woodruff et al. (1983) classify norms into brand-based norms and product-based norms. Cadotte et al. (1983) operationalize three types of norms in a restaurant setting, a best brand norm, a product type norm, and a brand norm. The best brand norm is the expected performance of the brand the consumer believes is the best (it may be the same as the focal brand), the product type norm is the typical performance expected from competing brands of the same product type within a product class, and the brand norm is the typical performance expected from the same or previously used brand. The researchers found these norms to be valid and useful concepts. The results showed that inferred disconfirmation models with the best brand norm or product type norm explain more of the variation in satisfaction than the focal brand as a norm (Cadotte et al. 1987). Prakash and Lounsbury (1984) studied how confirmation of predictive, normative (should) and comparative (two other restaurants) expectations affected satisfaction with a hamburger restaurant. Only predictive expectations seemed to be correlated directly with satisfaction. Inferred disconfirmation of all three standards were, however, significantly correlated with satisfaction. Boling and Woodruff (1988) found that a focal brand model, favorite brand model, and product norm model differed in explaining satisfaction with a wine in high and low involvement situations.

Summary of comparison standards. In Figure 1 we have made an attempt to categorize the different standards of comparison and show how perceived service quality is usually measured (Grönroos 1982; Parasuraman et al. 1988). It is possible to compare also other standards with perceived performance, although this has not been considered earlier within service quality research.

We have categorized the comparison standards

into four types, transaction, relationship, industry standard and ideal. The transactional approach focuses on perceived service quality of a specific transaction. This is the classical way of looking at the interaction between a customer and a service provider. Lately, however, attention has shifted towards a more dynamic approach, where the customer's relationship with the company is considered. If the customer has previous experiences with a service, predictive expectations will include perceptions of the relationship. Grönroos' well known service quality model (Grönroos 1990) separates expectations and image. In our opinion image represents predictive expectations over time. The brand norm can also be compared to the image that a consumer has of a service provider. Both the brand norm and expectations over time are connected to a specific company. The industry standard on the other hand is based on other companies within an industry that the consumer has experienced. The ideal type is not necessarily connected to any existing service. The issue of which standard to use when measuring service quality has been neglected, but is important as different types of standards may lead to different results.

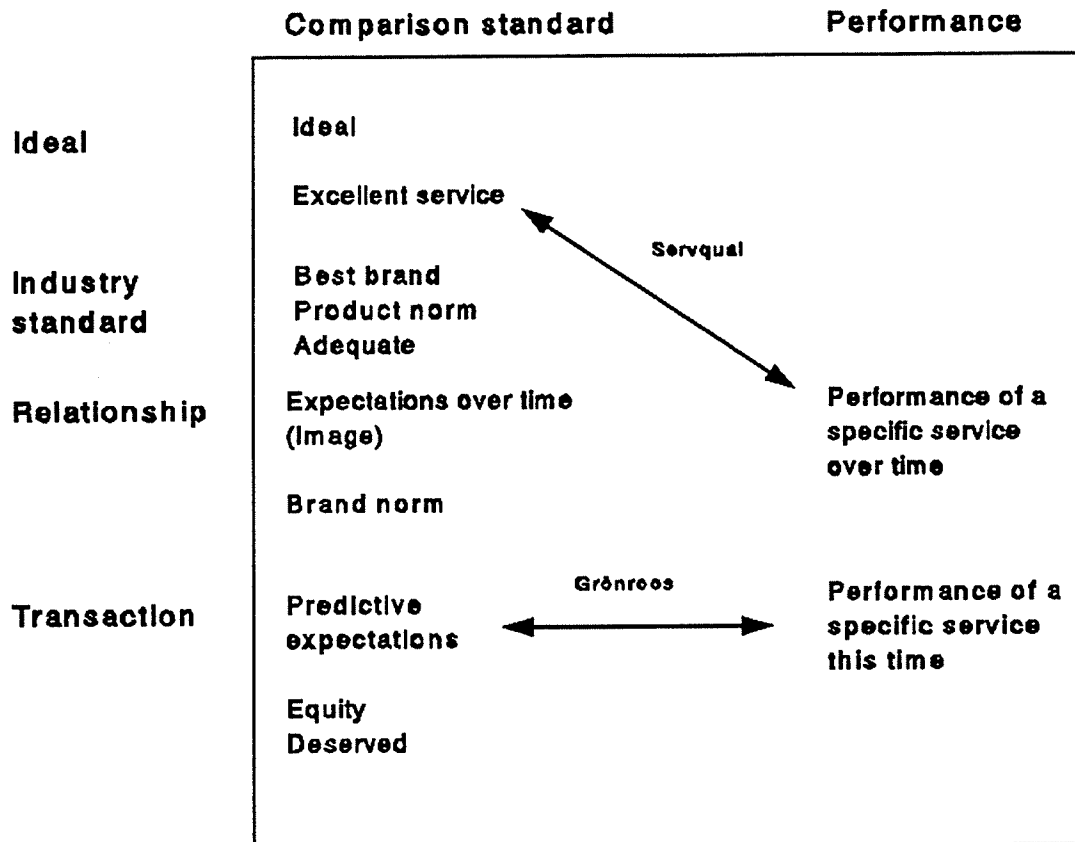
PERFORMANCE

Several studies have showed a connection between performance and satisfaction. Performance was the best predictor of satisfaction for a VCR (Churchill and Surprenant 1982), beer and hamburger restaurant (Prakash and Lounsbury 1984) and a miniature record player (Tse and Wilton 1988). The role of performance has received increasing attention in service quality research as well. Cronin and Taylor (1992), Liljander and Strandvik (1992b) and Parasuraman et al. (1991) found a stronger correlation between the performance of the service and service quality than between inferred disconfirmation of desired service and service quality.

It has been suggested that expectations for a continually consumed product (Bolton and Drew 1991a) or a product with which the consumer has large experience (Johnson and Fornell 1991) will equal perceptions of performance of the product. Expectations may here be best interpreted as predictive expectations. When a consumer is

Figure 1
A Categorization of Different Standards of Comparison and Performance

Concepts measured in this study in bold text.



familiar with a service s/he may be able to predict fairly accurately what the service will be like, unless the service varies much from one occasion to another. It is, however, less clear how experience with the service will effect satisfaction models with other standards of comparison than predicted service. Much experience with other brands may well enhance the role of a product norm in a disconfirmation model. On the other hand an ideal or excellent norm will probably not differ between experienced and non-experienced consumers, as showed by Liljander and Strandvik (1992b) where significant differences were only found for one of the expectations attributes and two of the performance attributes, where first time

visitors evaluated the attributes more positively.

DISCONFIRMATION MEASURES

There are basically two methods of measuring disconfirmation. It can be measured as an inferred measure where the score for the measured standard of comparison is subtracted from the score for perceptions, or as a direct measure where the consumer is e.g. asked to compare the performance with his/her expectations on a scale from "worse than expected" to "better than expected". When an inferred measure is employed evaluations on comparison standards may be collected either before or after the consumption of

the product. In satisfaction studies concerning services comparison standards, e.g. brand expectations and experience-based norms have usually been measured before the consumption (Cadotte et al. 1983; Swan 1977; 1988; Swan & Trawick 1982), while the comparison standard in service quality studies, i.e. desired service, has been measured at the same time as performance. Another distinction is that direct and inferred disconfirmation may be measured either on attributes or as holistic measures. This is depicted in Figure 2. In our view one should consider the type of disconfirmation measure used when comparing results from different studies. There is very little evidence of which measure should be used, inferred or direct disconfirmation, and if they should be measured on attributes or holistically. Although several studies have the ingredients to compare the measures, to our knowledge only Tse and Wilton (1988) did some comparisons. They, however, did not consider inferred and direct disconfirmation as interchangeable constructs in themselves, but compared a model of inferred disconfirmation with a model of expectations, performance and direct disconfirmation. They found that of the different standards used, disconfirmation of an ideal product was the best predictor of satisfaction. The three construct model, however, was superior in explaining satisfaction. This was interpreted to mean that a direct measure is a better measure of disconfirmation.

Although the main attention in research has been on disconfirmation and its relation to satisfaction/service quality, the relation between expectations, performance and satisfaction in themselves have been investigated with some interesting results. Different models with the components of expectations, performance and direct disconfirmation of comparison standards have been proposed. Lately special attention has been given to the effect that performance alone seems to have on satisfaction.

Perception of performance is usually assumed to be immediate and connected to the service encounter. To our knowledge no studies have been done where perceived performance would have been measured at different points of time after the encounter. It can be argued that at least some services (like educational services) are

reevaluated over time.

RESEARCH QUESTIONS

The discussion in this section is organized around rather broad research questions, each formalized as different specific models used in our study.

According to many consumer satisfaction studies the customer is satisfied if performance surpasses or equals predictive expectations. Concerning disconfirmation of experience-based norms the relationship is not clear. Although disconfirmation of norms may explain variation in satisfaction, performance does not necessarily have to surpass or equal the norm for the customer to be satisfied. Thus the customer may be satisfied even though e.g. the best brand norm has not been equalled or surpassed. Cadotte et al. (1983) reported that the best brand norm consistently had the highest mean, followed by brand expectations, product norm and brand norm, although the last two changed places for one of the three restaurants studied.

Q1. What is the position of different comparison standards on a scale? What is the relation between different comparison standards?

One specification of the simple disconfirmation model is to consider the comparison standard to be formed and measured before the event. The consumer is according to this model thought to compare these pre-consumption standards with performance.

Satisfaction = Performance - Predictive expectations (1)

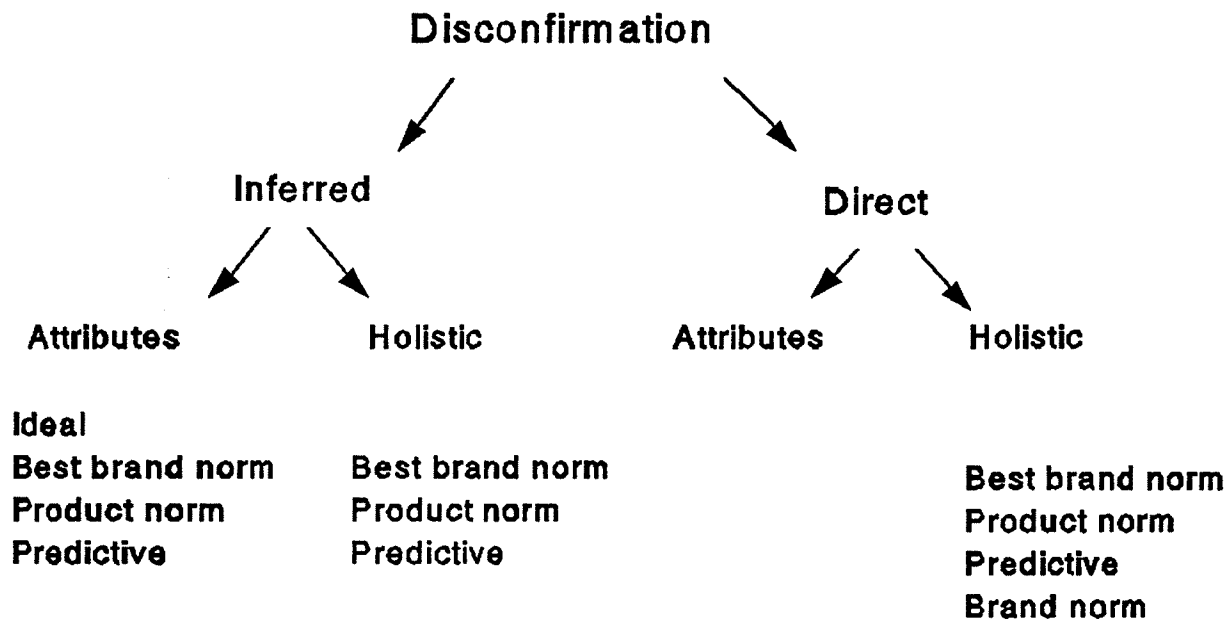
Satisfaction = Performance - Product norm (2)

Satisfaction = Performance - Best brand norm (3)

Q2. Which of the inferred disconfirmation measures explain satisfaction / perceived service quality best? In our study three different inferred disconfirmations measures are employed, disconfirmation of predictive expectations, product norm, and best brand norm.

Figure 2
Different Types of Disconfirmation Measures

Comparison standards used in the empirical study are shown in the figure.



Satisfaction may also be determined by direct disconfirmation alone. Although direct disconfirmation in earlier studies has been used only in connection with other constructs it is possible that this construct in itself may determine satisfaction. Direct disconfirmation of predictive expectations, product norm, best brand norm, and brand norm are investigated.

Satisfaction = Direct disconfirmation of predictive expectations (4)

Satisfaction = Direct disconfirmation of product norm (5)

Satisfaction = Direct disconfirmation of best brand norm (6)

Satisfaction = Direct disconfirmation of brand norm (7)

Q3. Which of the direct disconfirmation measures explain satisfaction / service quality best?

As we mentioned earlier there is to our knowledge very little research on the difference between using inferred and direct disconfirmation measures. To our knowledge no one has compared inferred and direct disconfirmation alone as predictors of satisfaction although this is a very interesting research question.

Q4. Are inferred disconfirmation measures or direct disconfirmation measures better at explaining satisfaction?

Performance alone has been found to be a strong predictor of satisfaction (Churchill and

Surprenant 1982; Cronin and Taylor 1992; Liljander and Strandvik 1992b; Tse and Wilton 1988)

Satisfaction = Performance (8)

Q5. Is experience with the service a better, or as good a, predictor of satisfaction / service quality as inferred or direct disconfirmation?

A model with predictive expectations, performance and direct disconfirmation of predictive expectations has been reported as a good model of satisfaction (Churchill and Surprenant 1982; Tse and Wilton 1988). Liljander and Strandvik (1992b), however, found that a model with performance and direct disconfirmation of brand expectations explained as much of the variation in satisfaction as a three construct model with "should"-expectations added. As a comparison we will also test models with experience-based norms, performance and direct disconfirmation of the norms.

Satisfaction = Performance + Direct disconfirmation of predictive expectations (9)

Satisfaction = Performance + Direct disconfirmation of product norm (10)

Satisfaction = Performance + Direct disconfirmation of best brand norm (11)

Satisfaction = Performance + Direct disconfirmation of brand norm (12)

Satisfaction = Predictive expectations + Performance + Direct disconfirmation of predictive expectations (13)

Satisfaction = Product norm + Performance + Direct disconfirmation of product norm (14)

Satisfaction = Best brand norm + Performance + Direct disconfirmation of best brand norm (15)

Q6) Does performance together with direct disconfirmation explain more of the variation in satisfaction than either inferred or direct disconfirmation measures alone? Are the models 9 to 12 better than, or as good, as the models 13 to 15 at explaining satisfaction with the service, as found by Liljander and Strandvik (1992b)?

Q7) Which of the models 13 to 15 explain satisfaction best? Are any of these models a better predictor of satisfaction than inferred disconfirmation alone (models 1 to 3), as found by Tse and Wilton (1988)?

DATA COLLECTION

Educational services were chosen as the subject of the study and two questionnaires were designed to test different models of service quality. The self-administered questionnaires were administered at the beginning and end of three different marketing courses at a business school in spring 1992. Participation was voluntary. In the first part of the study students evaluated attribute-specific and global questions on predictive expectations, product type norm and best brand norm, and weights of attributes. Global questions of expected satisfaction with the course were also included. In the second part about three months later, they evaluated course performance on attributes, direct disconfirmation of predictive expectations, product type norm, best brand norm and brand norm. Evaluations on overall satisfaction with the course were also collected. The measures used in the study are presented in Appendix 1. 181 students participated in the first part of the study in February 1992 and 140 in the second part in May 1992. Only those who filled out both questionnaires (n=82) are used in our analysis.

A 7-grade scale was used anchored in subjective perceptions of the "worst course the person had experienced" and an "ideal course". All ratings are thus relative to these anchors. Different respondents then might have different objective definitions of worst and ideal based on their experience and other personal characteristics. In this study we were not interested in connecting the relative scales to actual conditions, but only in studying the formation of quality and satisfaction based on the students socially constructed reality. The standards of comparison and performance were measured on 11 attributes (Appendix 2). Several holistic measures were also used. The attributes were mainly chosen from an earlier study on course quality (Liljander and Strandvik 1992a), selecting those that had been most relevant in explaining service quality in that study. In

addition some attributes that students often consider important were included. Variables used in the analysis are listed in Table 1.

Table 1
Direct and Derived Measures Used in the Analysis
(Abbreviations as they are used in the text)

<u>Measure</u>	<u>Measurement before course</u>	<u>Measurement after course</u>	<u>Derived measures</u>
Predictive expectations on attributes			
PE			
Product norm on attributes			
PN			
Best brand norm on attributes			
BBN			
Holistic predictive expectations			
HOL(PE)			
Holistic product norm			
HOL(PN)			
Holistic best brand norm			
HOL(BBN)			
Performance on attributes		PER	
Holistic performance measure		HOL(PER)	
Direct disconfirmation of predictive expectations			DD(PE)
Direct disconfirmation of product norm			DD(PN)
Direct disconfirmation of best brand norm			DD(BBN)
Direct disconfirmation of brand norm			DD(BN)
Average of inferred disconfirmation of predictive expectations			AVG(PER-PE)
Average of inferred disconfirmation of product norm			AVG(PER-PN)
Average of inferred disconfirmation of best brand norm			AVG(PER-BBN)
Average of performance			AVG(PER)
Average of predictive expectations			AVG(PE)
Average of product norm			AVG(PN)
Average of best brand norm			AVG(BBN)
Graphics satisfaction measure		Satisfaction	

Predictive expectations were defined as what the students thought the course would be like based on a written description of the course, an introduction lecture to the course by the professor and word-of-mouth. The product norm was defined as the typical performance of courses at the school, the best brand norm as the best course that the student had experienced and the brand norm as a typical marketing course. Satisfaction was measured as a graphical total satisfaction measure where the respondents are asked to

evaluate their satisfaction on a commonly used (e.g. Oliver and Westbrook 1982) 7-point scale where the scale values are faces with expressions displaying different degrees of satisfaction,

RESULTS

Scale reliabilities were tested with Cronbach's alpha. All direct attribute scales (PE, PN, BBN, PER) and attribute-specific inferred disconfirmation scales (PER-PE, PER-PN, PER-BBN) had reliability coefficients between 0.78 and 0.82. We also tested if there were any significant differences between those who only answered one of the questionnaires (before or after the course) and those that took part in both parts of the study. No differences could be found. The discussion of the results are organized around the seven research questions although some of them are discussed together into the same section.

The Different Comparison Standards and Their Relation to Performance

In this section results concerning Q1 are presented. The correlations of comparison standards and performance are presented in Table 2. The results differ slightly depending on whether or not the measure is an average of attributes or a holistic measure. Of the average measures of comparison standards the product norm and best brand norm are not correlated, while of the holistic measures the predicted expectations and product norm are not correlated. The corresponding average and holistic comparison standard measures are all significantly correlated, although the predictive measures show the highest correlation (0.60). The average of performance on attributes and the holistic measure of performance are also well correlated (0.78). This may indicate that compared to predictive expectations and performance, additional factors beside the 11 attributes are considered when evaluating the product norm and best brand norm holistically.

Of the comparison standards only predictive expectations are correlated with performance. The correlation is interesting considering that the consumers had no previous experience of the courses and there was a three month period between the evaluations before and after the

course. The correlation may also be due to cognitive dissonance or assimilation effects when evaluating the performance of the course. Although the number of respondents per course was limited we performed an exploratory correlation analysis for each course. This analysis showed that predictive expectations were correlated with performance only for the course in the second year of studies. This might mean that the information given about the course matched the performance of it better than was the case for the other courses.

Table 2
Pearson Correlation Matrix of Comparison Standards and Performance

	AVG(PE)	AVG(PN)	AVG(BBN)	HOL(PE)	HOL(PN)	HOL(BBN)	AVG(PER)
AVG(PN)		0.41**					
AVG(BBN)		0.25*	0.09				
HOL(PE)		0.60**	0.39**	0.05			
HOL(PN)		-0.01	0.28**	0.11	0.08		
HOL(BBN)		0.19	0.07	0.37**	0.34**	-0.24*	
AVG(PER)		0.31**	0.13	0.11	0.10	-0.07	0.10
HOL(PER)		0.26*	0.07	0.10	0.14	0.01	0.04
							0.78**

* = $p < 0.05$, ** = $p < 0.01$ Profiles with attribute-specific

information for the different comparison standards and performance are presented in Figure 3. Predictive expectations (PE) are significantly ($p < 0.05$) different from the product norm (PN) on all attributes except attribute number 4. PE also differs significantly from the best brand norm (BBN) on all attributes except number 7 and 8, and from performance (PER) on all attributes except 2, 3, 4 and 5. The BBN differs from PN on all attributes except number 4 and all differences between BBN and PER are significant. For PN and PER no significant differences can be found for attributes 1, 3, 4, 8, 9, 10 and 11. Thus on an aggregate level the only comparison standard that performance surpasses, on at least a few attributes, is the product norm.

Satisfaction was regressed on the performance attributes and the inferred disconfirmation of each

comparison standard. The results will not be presented here in detail but it can be concluded that coefficients for attributes 1-5 were consistently the lowest while attribute 6 had the highest coefficient. The importance weights inferred by simple regression analysis also differed remarkably from the directly asked importance of attributes.

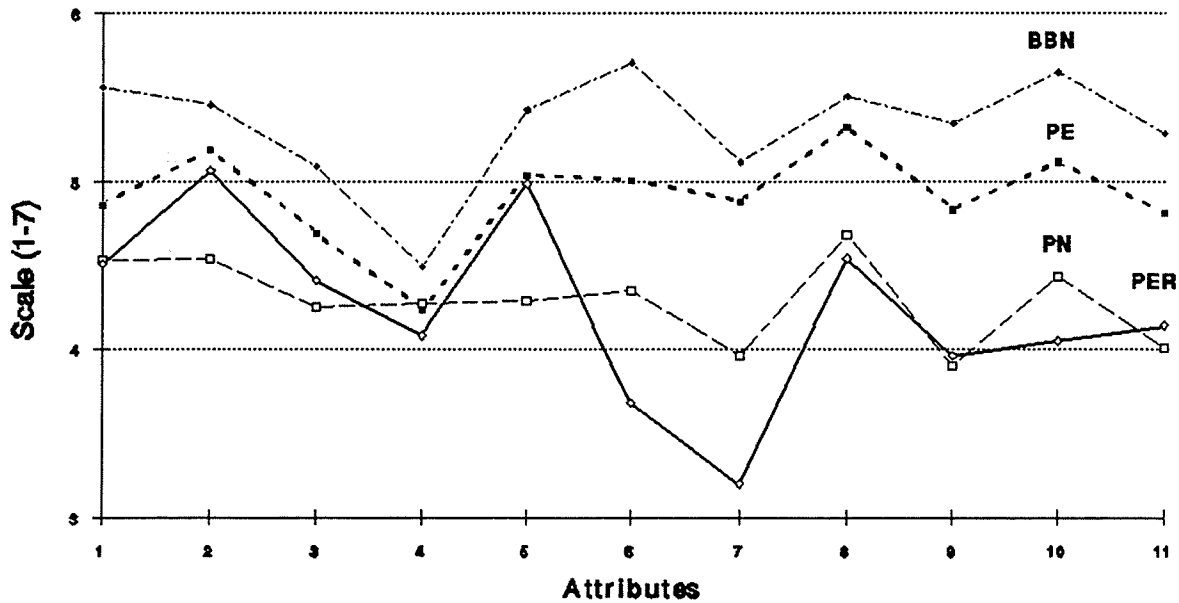
In Table 3 both the means of the 11 attributes and of the holistically measured comparison standards and performance are shown. The best brand norm has the highest mean followed by predicted expectations. Both these standards are significantly different from each other and from the product norm and performance. The product norm and performance are, however, not significantly different from each other. These results are similar to Cadotte et al. (1983) where the best brand norm was followed by predicted expectations and then product norm. The position of performance on the scale was not reported by Cadotte et al.

Table 3
Means of Different Comparison Standards and Performance
(N=82)

	Best brand norm	Predicted expectations	Product norm	Performance
Average on attributes	5.3	4.9	4.3	4.2
Holistic measure	5.4	5.0	4.1	4.0

We also divided the data into satisfied and dissatisfied consumers to investigate any differences in how they evaluated the comparison standards. The data is presented in Table 4. As could be expected the biggest difference can be found for the means on performance. The only significant difference between satisfied and dissatisfied consumers for a comparison standard is between the attribute means for the product norm where satisfied consumer have evaluated a typical course slightly higher than those who expressed dissatisfaction. All comparison standards are significantly different from each other ($p < 0.05$) and all except predictive expectations for those who were satisfied, are

Figure 3
Different Comparison Standards and
Perceived Performance on Attributes
N=82



significantly different from performance. The relative position of the standards remain the same as in Table 3. The best brand norm represents the highest level of comparison standard on the seven-graded scale. Predictive expectations exceed the average product norm. For the dissatisfied consumers perceived performance is lower than all comparison standards, while for the satisfied consumers performance exceeds the product norm.

The order of the comparison standards in both our and Cadotte et al.'s study seem to indicate that consumers choose a product which they believe will perform better than what is typical within the same product class. Table 4 shows that consumers are satisfied even though the best brand norm or

predictive expectations are not surpassed. The product norm can, however, here be seen as a critical line resembling what is acceptable for the service. If performance is above the product norm the customer is satisfied and if it is below it s/he is dissatisfied.

Inferred and Direct Disconfirmation Compared with Performance

Results for research questions Q2 - Q5 are reported in this section. The service quality models 1-8 represented predictions of satisfaction by either disconfirmation or performance. Results of the models are presented in Table 5.

Table 4
Means of Different Comparison Standards
and Performance for Satisfied and
Dissatisfied Customers
Satisfied = 5-7 on a 7-point scale,
Dissatisfied = 1-4 on scale values
N = 8

<u>Best brand norm</u>	<u>Predicted expectations</u>	<u>Product norm</u>	<u>Performance</u>
Attributes/satisfied (N = 41)	5.35	4.95	4.45* 4.74*
Attributes/dissatisfied (N = 41)	5.31	4.87	4.15* 3.69*
Holistic/satisfied (N = 41)	5.39	5.03	4.05 4.85*
Holistic/dissatisfied (N = 41)	5.34	5.05	4.07 3.20*

* = $p < 0.05$ between satisfied and dissatisfied consumers

Table 5
Satisfaction Regressed on Inferred and
Direct Disconfirmation

	<u>Std coef</u>	<u>R²</u>
AVG(PER-PE)	0.50**	0.25
AVG(PER-PN)	0.52**	0.27
AVG(PER-BBN)	0.56**	0.31
DD(PE)	0.78**	0.60
DD(PN)	0.71**	0.51
DD(BBN)	0.60**	0.36
DD(BN)	0.76**	0.58
HOL(PER)	0.82**	0.67
AVG(PER)	0.72**	0.52

** = $p < 0.01$

Results are mixed and seem to be dependent on when the comparison standard is measured. Standards measured long before the evaluation of performance clearly seem to be inferior to perceived disconfirmation at the time when performance is evaluated. What probably happens is that through a learning process the beliefs about the service change during the process of using the service. What customers expect before using a unique service might influence what they choose, but not how satisfied they are with the service.

Our results indicate that performance is a

strong predictor of satisfaction. Inferred disconfirmation explains clearly less of the variation in satisfaction than direct disconfirmation and performance. Direct disconfirmation of predictive expectations and direct disconfirmation of the brand norm are better predictors of satisfaction than the two other direct disconfirmation measures. The effect seems to be related to the respondent's scope, i.e. how close the norm is to the specific service. Direct disconfirmation of the best brand norm which includes a wider scope of experiences as the base for the norm, explains considerably less than the norm with the most narrow scope, direct disconfirmation of predictive expectations.

A problem with holistic direct disconfirmation measures is that they give limited diagnostic information about reasons for dis/satisfaction. It is, however, also possible to use direct disconfirmation on attributes. From a practical point of view the employment of direct disconfirmation measures is easier as only one measurement is needed.

Another view on the relation between responses on different measures is given in table 6. Different disconfirmation measures are crosstabulated with an overall satisfaction measure. In the table direct and inferred disconfirmation regarding a specific comparison standard is presented. Performance has been classified as either worse than the comparison standard, as the comparison standard and better than the comparison standard. In the table shaded cells represent expected situations. In table 6 it can be noted that even if performance is worse than predictive expectations (inferred disconfirmation) 27 percent of the respondents are satisfied. The results showed that regardless of which disconfirmation type or comparison standard was used, respondents are more satisfied according to the overall satisfaction measure, than what would be predicted based on the disconfirmation measures. Brand norm represents the only exception, with a more balanced fit.

Experience and Direct Disconfirmation, Including and Excluding Comparison Standards, as Predictors of Satisfaction

To answer research questions Q6 and Q7

Table 6
Inferred and Direct Disconfirmation of
Predictive Expectations Compared to Overall
Satisfaction on the Graphics Scale
Dissatisfied = Scale Values 1-3,
Indifferent = Scale Value 4,
Satisfied = Scale Value 5-7
Percents of Total (N = 82)

	Dissatisfied	Indifferent	Satisfied
Worse than comparison standard			
Inferred	29	18	27
Direct	29	13	9
As comparison standard			
Inferred	2	0	4
Direct	2	2	13
Better than comparison standard			
Inferred	0	0	20
Direct	0	2	28

satisfaction was first regressed on performance and direct disconfirmation (models 9-12) and then on these constructs with comparison standards added (models 13-15). The results are presented in Tables 7 and 8 respectively.

There were two measures of performance in the study, an overall measure and an evaluation on 11 attributes, while direct disconfirmation was only measured holistically. Although a holistic measure of performance yields the best R²; it has the managerial disadvantage of not revealing which parts of the service customers evaluated positively and which negatively. As it is unlikely that only holistic measures would be used in practice we have chosen to use the attribute-based construct of performance in the analysis.

Table 7 shows that adding direct disconfirmation to performance improves the explanation of satisfaction compared to models with either of the constructs alone. Inferred disconfirmation is clearly an inferior measure compared to any of these models. When we compare the results of Table 7 to those in Table 8 where the comparison standards are added we can see that the differences are rather small. For predictive expectations there is no difference in the R² and for the models with the product norm and best brand norm the R² are improved by 0.02. The managerial implications would be that it is not necessary to measure comparison standards directly, but rather measure different overall direct

disconfirmation measures. These results resemble those obtained by Liljander and Strandvik (1992) although here there was at least a small improvement in R² when the comparison standards were added.

The findings by Tse and Wilton (1988) are also confirmed in that all the models in Table 8 are better than any model of inferred disconfirmation. The best model seems to be the brand norm added to performance. The order of the other comparison standard models regarding their ability to explain satisfaction remains the same in both tables, i.e. predictive expectations are followed by the product norm and the best brand norm in that order. What seems to be important is the closeness of the reference point to the service in question. The brand norm and predictive expectation that are directly connected to the focal brand may be more realistic reference points than the product norm and best brand norm.

Table 7
Satisfaction Regressed on Performance
Together with Direct Disconfirmation
Measures

	<u>Std coef (1)</u>	<u>Std coef (2)</u>	<u>R²</u>
AVG(PER) + DD(PE)	0.35**	0.53**	0.67**
AVG(PER) + DD(PN)	0.47**	0.41**	0.63**
AVG(PER) + DD(BBN)	0.58**	0.28**	0.59**
AVG(PER) + DD(BN)	0.52**	0.44**	0.75**

** = p < 0.01

According to Green, Tull and Albaum (1988) one test for multicollinearity is that the correlations among the predictors should not exceed the correlation of the criterion variable with the predictors. The data passed this test for all except one of the constructs. Many of the correlations were, however, only slightly lower within the predicting variables than with satisfaction. There are many different views on where the line should be drawn for what is an acceptable level for correlation. We feel that the high correlations

should be kept in mind when looking at the results of the regression analysis. This might be one reason for accepting performance alone as the best measure of satisfaction, unless the other constructs are needed by the management for diagnostic purposes.

Table 8
Satisfaction Regressed on Expectations, Performance and Direct Disconfirmation

	<u>Std coef (1)</u>	<u>Std coef (2)</u>	<u>Std coef (3)</u>	<u>R²</u>
AVG(PE) + AVG(PER) + DD(PE)	0.10	0.30**	0.55**	0.67**
AVG(PN) + AVG(PER) + DD(PN)	0.14**	0.43**	0.45**	0.65**
AVG(BBN) + AVG(PER) + DD(BBN)	0.05	0.49**	0.36**	0.61**

** = $p < 0.01$

DISCUSSION

The results showed that the assumption that customers rely on prepurchase expectations as the comparison standard is not effective. We have used the term inferred disconfirmation to represent this standard. It proved to be the weakest in predicting satisfaction with the service. From a practical point of view this is fortunate as it reduces measurement difficulties and costs dramatically, as measurement only has to be made in one point in time. In our study we lost a considerable amount of the respondents as two measurements had to be matched to get the inferred measure.

It has to be remembered that the service in our case was an educational service with a long duration, about three months, where it seems natural that the whole interaction process influences the perceptions of both comparison standards and performance. In this case the service is also in some respect unique, the customer experiences only once this particular service. What according to our results seems to be the most effective comparison standard is the brand norm, which in our case represent courses given by the same department. This could in another context be interpreted as the consistency of a particular service providers' services. Students

seem to be satisfied if the performance exceeds the brand norm. It is not necessary that performance lives up to what students believed the course would be like (predictive expectations) or the best course (best brand norm) they had experienced. Regardless of how Parasuraman et al.'s normative comparison standard is interpreted, the alternatives seems to be either as our ideal or the best brand norm, it is clear that the standard used in Servqual would not be effective.

There has been a growing interest among practitioners in customer satisfaction measurement. A new aspect from this perspective is the need for diagnostic information. Management is interested in understanding the reasons for customer satisfaction or dissatisfaction. One issue in understanding what drives customer satisfaction is the choice of comparison standards. In service quality models one particular standard has been assumed. If it can be shown that customers in fact use another standard or maybe several other standards simultaneously this would be important information. The normal service quality model suggests that customers are satisfied or dissatisfied according to whether the service lives up to their prepurchase expectations. Experience-based norms suggested in the literature have partly a broader scope, by suggesting that satisfaction might be dependent on what customers know and believe about competing services. The analysis could be taken one step further by including explicit comparisons with competing services by specifying a best competitor norm. In this way the customers choice situation would be included in the service quality model. This would mean a shift from a static dyadic perspective, which is implicit in service quality models, to a more competition oriented choice perspective which might be called for in many practical situations. In this study we, however, used only "traditional" experience-based norms. We have not looked at individual and situational differences in the use of comparison standards. This would be an important extension of this line of research as indicated already by Woodruff et al. (1987, 1991).

REFERENCES

- Bolting, Clair and Robert B. Woodruff (1988), "Effects of Situational Involvement on Consumers' use of Standards in Satisfaction/Dissatisfaction Processes,"

- Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 16-24
- Bolton, Ruth N. and James H. Drew (1991a), "A Longitudinal Analysis of the Impacts of Service Changes on Customer Attitudes," *Journal of Marketing*, 55, 1-9.
- Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins (1983), "Norms and Expectations: How Different are the Measures?," in Ralph L. Day and H. Keith Hunt (eds.), *International Fare In Consumer Satisfaction and Complaining Behavior*, Bloomington, IN, Indiana University, 49-56.
- Cadotte, Ernest R., Robert B. Woodruff and Roger L. Jenkins (1987), "Expectations and Norms in Models of Consumer Satisfaction," *Journal of Marketing Research*, 24, (August), 305-314.
- Churchill and Surprenant (1982), "An Investigation Into the Determinants of Customer Satisfaction," *Journal of Marketing Research*, 19 (November), 491-504.
- Cronin, Joseph J. and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56, (July), 55-68.
- Day, Ralph H. (1983), "The Next Step: Commonly Accepted Constructs for Satisfaction Research," in Ralph L. Day and H. Keith Hunt (eds.), *International Fare In Consumer Satisfaction and Complaining Behavior*, Bloomington, IN, Indiana University, 113-117.
- Green, Paul E., Donald S. Tull and Gerald Albaum (1988), *Research for Marketing Decisions*, 5th ed., Englewood Cliffs, Prentice Hall Int.
- Grönroos, Christian (1982), "Strategic Management and Marketing in the Service Sector," Research Reports No 8., Helsinki, Finland, Swedish School of Economics and Business Administration.
- Grönroos, Christian (1990), *Service Management and Marketing*, Toronto, Lexington Books.
- Johnson, Michael D. and Claes Fornell (1991), "A Framework for Comparing Customer Satisfaction Across Individuals and Product Categories," *Journal of Economic Psychology*, 12, 267-286.
- Liljander, Veronica and Tore Strandvik (1992a), "Är tjänstekvalitet skillnaden mellan förväntningar och upplevelse?" (translated: Is Service Quality the Difference Between Expectations And Performance?) Working Papers No. 233, Swedish School of Economics and Business Administration, Helsinki.
- Liljander, Veronica and Tore Strandvik (1992b), "The relation between service quality, satisfaction and intentions," in Working Papers No. 243, Swedish School of Economics and Business Administration, Helsinki.
- Oliver, Richard L. (1980a), "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research*, 17, (November), 460-469.
- Oliver, Richard L. (1980b), "Conceptualization and Measurement of Disconfirmation Perceptions in the Prediction of Consumer Satisfaction", in Keith Hunt and Ralph L. Day (eds.), *Refining Concepts and Measures of Consumer Satisfaction and Complaining Behavior*, Bloomington, IN, Indiana University,
- Oliver, Richard L. and William O. Bearden (1983), "The Role of Involvement in Satisfaction Processes," in Richard P. Bagozzi and Alice M. Tybout (eds.), *Advances in Consumer Research*, 10, Ann Arbor, Association for Consumer Research, 250-255.
- Oliver, Richard L. and Robert A. Westbrook (1982): "The Factor Structure of Satisfaction and Related Postpurchase Measures," in Ralph L. Day and H. Keith Hunt (eds.): *New Findings on Consumer Satisfaction and Complaining Behavior*, Bloomington, IN: Indiana University, 11-14.
- Olson, Jerry C. and Philip A. Dover (1979), "Disconfirmation of Consumer Expectations Through Product Trial," *Journal of Applied Psychology*, 64, (April), 179-189.
- Parasuraman, A., Valarie A. Zeithaml and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49, (Fall), 41-50.
- Parasuraman, A. and Valarie A. Zeithaml and Leonard L. Berry (1988), SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, 64, 1 (Spring), 12- 40.
- Parasuraman, A. and Valarie A. Zeithaml and Leonard L. Berry (1991), "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing*, 67, 4, Winter 1991, 420-450.
- Prakash, Ved and John W. Lounsbury (1984), "The Role of Expectations in The Determination of Consumer Satisfaction," *Journal of the Academy of Marketing Science*, 12, 4 (Fall), 1-17.
- Swan, John E. (1977), "Consumer Satisfaction with a Retail Store Related to the Fulfillment of Expectations on an Initial Shopping Trip," in Ralph L. Day (ed.), *Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, Bloomington, IN, Indiana University, 10-17.
- Swan, John E. (1988), "Consumer Satisfaction Related to Disconfirmation of Expectations and Product Performance," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 1, 40-47.
- Swan, John E. (1992), "Satisfaction Work: The Joint Production of Patient Satisfaction by Health Care Providers and Patients," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 5, 69-80.
- Swan, John E. and I Fredrick Trawick (1982), "Satisfaction, Disconfirmation, and Comparison of Alternatives," in H. Keith Hunt and Ralph L. Day (eds.), *Conceptual and Empirical Contributions To Consumer Satisfaction and Complaining Behavior*, 17-24.

- Tse, David K. and Peter C. Wilton (1988), "Models of Consumer Satisfaction Formation: An Extension," *Journal of Marketing Research*, 25, (May), 204-12.
- Woodruff, Robert B., Ernest R. Cadotte and Roger L. Jenkins (1983), "Modeling Consumer Satisfaction Processes Using Experience-Based Norms," *Journal of Marketing Research*, 20, (August), 296-304.
- Woodruff, Robert B., D. Scott Clemons, David W. Schumann, Sarah F. Gardial and Mary Jane Burns (1991), "The Standards Issue in CS/D Research: A Historical Perspective," *Journal of Consumer Satisfaction, Dissatisfaction and Complaining Behavior*, 4, 103-110.
- Zeithaml, V., L. Berry and A. Parasuraman (1993), "The Nature and Determinants of Customer Expectations of Service," *Journal of the Academy of Marketing Science*, 21, 1, (Winter), 1-12.

APPENDIX 1

Description of the Direct Measures Used in the Study

<u>Construct</u>	<u>Type</u>	<u>No</u>	<u>Scale</u>
<u>Standards of comparison</u>			
Predictive expectations			
Attribute-specific	11	worst course	ideal course
Global	1	worst course	ideal course
Product type norm			
Attribute-specific	11	worst course	ideal course
Global	1	worst course	ideal course
Best brand norm			
Attribute-specific	11	worst course	ideal course
Global	1	worst course	ideal course
<u>Performance</u>			
Attribute-specific	11	worst course	ideal course
Global	1	worst course	ideal course
<u>Direct disconfirmation</u> The performance was			
Predictive expectations			
Global	1	much worse	much better
		than	than
Product type norm			
Global	1	much worse	much better
		than	than
Best brand norm			
Global	1	much worse	much better
		than	than
Brand norm			
Global	1	much worse	much better
		than	than
<u>Satisfaction</u>			
Satisfaction			
Global	1	Not at all	very
		satisfied	satisfied

APPENDIX 2

Attributes (translations from Swedish)

1. How relevant is the course for your future profession?
2. How up-to-date is the content of the course?
3. How demanding is the course?
4. How difficult is the course literature?
5. How relevant is the course literature?
6. Does the teaching stimulate your interest in the subject?
7. How useful are the lectures?
8. How useful are the exercises?
9. How well-structured is the course?
10. How competent are the teachers?
11. Does the teaching stimulate your own critical thinking?

Send correspondence regarding this article to:

Veronica Liljander or Tore Strandvik
 Swedish School of Economics and Business Administration
 Department of Marketing
 Arkadigatan 22
 SF-00100 Helsinki, Finland