A generic instrument to measure customer satisfaction with the controllable elements of the in-store shopping experience

N.S. Terblanche* and C. Boshoff

Department of Business Management, University of Stellenbosch
Stellenbosch 7600, Republic of South Africa
nst@maties.sun.ac.za

In this study an attempt is made to develop a generic instrument that could be used to measure customer satisfaction with the controllable elements of the in-store shopping experience. By closely following the most contemporary guidelines for scale development, and involving 11 063 respondents in four different surveys, the authors emerge with a 22-item instrument to measure satisfaction with the in-store shopping experience. The evidence of the psychometric properties of the proposed ISE instrument offered here is compelling in terms of its uni-dimensionality, with-in-method convergent validity, cross-validation of dimensions in a cross validation sample, reliability of the instrument, its discriminant validity and its nomological validity.

*To whom all correspondence should be addressed.

Introduction

Retailers are under constant pressure to adapt in response to ever-changing and accelerating environmental circumstances (Dabholkar, Thorpe & Rentz, 1996:3). More sophisticated and demanding customers, competition from both domestic and foreign sources and breathtaking new technological developments, are just a few of the variables pressurising retailers to find new ways to differentiate themselves from others. These attempts have ranged from a focus on service delivery to loyalty schemes – all with limited success (Egan, 1999; Sopanen, 1996; Berry, 1986; Hummel & Savitt, 1988; Reichheld & Sasser, 1990; Dabholkar *et al.*, 1996).

Kerin, Jain and Howard (1992:394) earlier contended that it is largely the store shopping experience itself which determines customer perceptions of a store. This school of thought contends that, from a measurement and management perspective, that a comprehensive instrument that captures *all* the dimensions of a shopping experience should be the focus as opposed to just one dimension such as service quality (Parasuraman, Zeithaml & Berry, 1988). Our contention is that in a retail environment where a mix of goods and services is offered, the comprehensive approach would be preferable. When considered in this way service quality, for instance, is only one component of the consumer's in-store shopping experience, as are several other components. If only one component of the in-store shopping experience is considered in isolation, it may be detrimental to our understanding of customers' experiences, and this in turn could lead to strategies that either overemphasise or neglect the importance of one or more retail experience components.

We have focused our investigation on the measurement and management of the controllable elements of the in-store shopping experience (ISE). We thus exclude the uncontrollable variables of a retail experience such as parking facilities, for instance. We have also focused our efforts on in-store retailing which, by definition, excludes retail formats such as catalogue and internet retailing, which do not typically have a significant personal interaction component (between customer and sales staff).

The objectives of this study

This study reports on several phases of a long-term examination of the controllable elements of the in-store shopping experience. The outcome of this stream of research is a generic, multi-item instrument that can be used to measure customer satisfaction with the controllable components of the in-store shopping experience in a variety of different retail environments. Based on the disconfirmation paradigm (Oliver, 1980), the objective of first two phases of this three-phase process was to identify the dimensions of importance to consumers when assessing their satisfaction with an in-store shopping experience. In other words, consumers were asked what their expectations of an in-store shopping experience were. Only once these dimensions had been identified and empirically confirmed, could we proceed to the development of a valid and reliable instrument to measure customer satisfaction with the in-store shopping experience at retailer or shop level.

The scale development process

Churchill (1979) proposed a well-accepted procedure for the development of valid and reliable multi-item instruments. This process consists of the following steps: domain

specification, generation of questionnaire items, empirical surveying, an iterative process of scale purification based on reliability assessment and validity checks, and the development of norms. This process has been enhanced in recent years by the availability of statistical procedures such as confirmatory factor analysis providing additional evidence of construct validity (Tull & Hawkins, 1993: 318). The process suggested by Churchill (1979) has been followed in the first two phases of this study.

The in-store shopping experience

The total retail experience consists of all the elements that encourage or inhibit consumers during their contact with a retailer (Berman & Evans, 1998:19) and can be either non-controllable (e.g. street parking, deliveries from suppliers and consumption taxes) or controllable and include in-store and external elements. In this study the emphasis is on the controllable elements of the in-store shopping experience.

The retail literature suggests that the controllable components of the in-store shopping experience may be grouped under six dimensions, namely:

Service quality
Merchandise quality
Merchandise variety and assortment
Internal store environment
Product prices
Store policies

Two dimensions that some may argue are part of the ISE are excluded from this study, namely store image and store location. In the retailing literature it is often suggested that a favorable store image leads to store loyalty (Hirschman, 1981). Store image, in turn, has been described as consisting of the following three general factors: merchandise-related aspects, service-related aspects, and pleasantness of shopping at a store (Mazursky & Jacoby, 1985). Our approach is that all three of these dimensions are captured in ISE (Figure 1) and that loyalty is an outcome of a positive in-store shopping experience (ISE) rather than an underlying dimension.

Although store location is normally one of the important reasons why customers visit a particular store, it is not addressed in this study because we are of the opinion that over time a favorable location may deteriorate in value because of changes in road patterns, the opening of competitive shops, and changing demographic patterns in the community.

Figure 1 illustrates the theoretical structure of the ISE, and the shaded dimensions are the ones included in the first empirical survey.
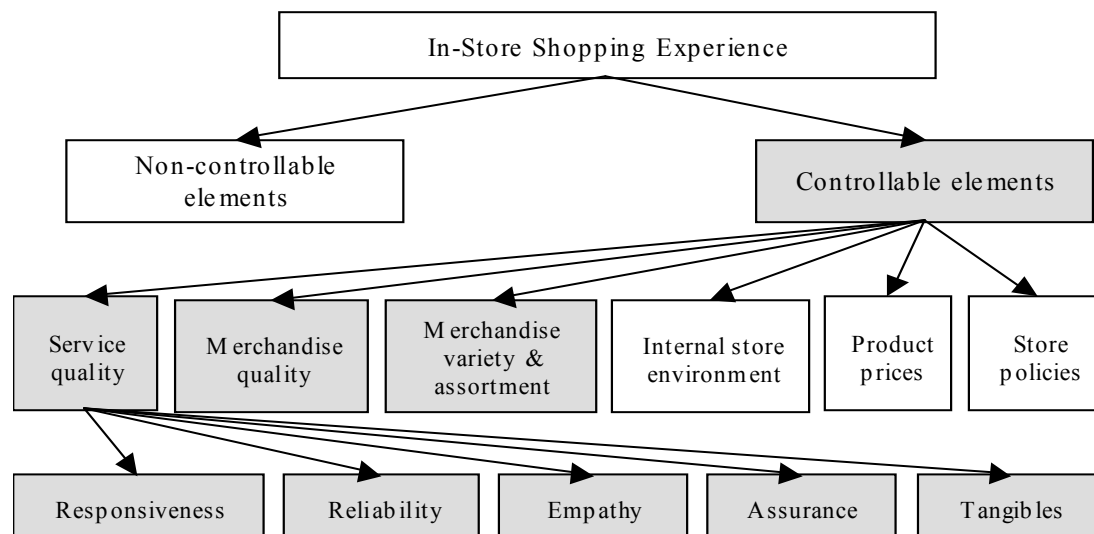

Figure 1: Schematic presentation of the theoretical structure of the in-store shopping experience

Because of the size of the complete ISE model and the large number of items measuring the three dimensions, as well as the very real threat of respondent fatigue, it was decided to include only three of the dimensions in the first empirical phase of the project, namely service quality (responsiveness, reliability, empathy, assurance, tangibles), merchandise quality, and merchandise variety and assortment. For a discussion of these three dimensions, see Terblanche and Boshoff (2001a: 101-103). These three dimensions of the in-store shopping experience (service quality, merchandise quality, and merchandise variety and assortment) and their

associated items were then subjected to an empirical survey to assess their importance to retail shoppers.

Methodology of the first phase

The first phase of the empirical research was undertaken to assess which dimensions and items measure the three dimensions of ISE shaded in Figure 1 (Service Quality, Merchandise Quality and Merchandise Variety and Assortment). All 22 items of SERVQUAL were used to measure Service Quality. The Merchandise Quality dimension was measured using three Merchandise Quality items suggested by Finn and Kayandé (1997) as well as two

self-generated items. Merchandise Variety and Assortment was measured using a self-generated five-item instrument. All 32 items were linked to a 7-point Likert-type scale. Respondents were asked to rate the importance of the various components of ISE on a 7-point scale where a 7 meant that the aspect under consideration was 'extremely important' and a 1 meant that it was 'not important'.

Customers from two fast food firms, two clothing firms, two supermarkets and two hardware stores constituted the population for the first phase of the research. A total of 2063 questionnaires were completed during personal interviews which were conducted after respondents had completed their shopping. In the scale purification process that followed, a Maximum Likelihood Exploratory Factor Analysis was conducted, specifying a Direct Quartimin oblique rotation (Jennrich & Sampson, 1966) of the original factor matrix, using the computer programme BMDP (Jennrich & Sampson, 1966). This was followed by an assessment of the internal consistency of each dimension as suggested by Churchill (1979).

Results of the first empirical survey

As part of the scale purification process, several different factor solutions were considered. The most interpretable factor structure (factor loading exceeding 0.4 and no cross-loadings) to emerge, was a 3-factor solution. All the factors in the 3-factor solution had Eigen values above 1.00, and a sufficient number of items loading on them to a significant (0.40) extent (Hair, Anderson, Tatham & Black, 1998). The three factors that emerged were named Personal Interaction (measured by 3 assurance, 4 empathy, 2 reliability and 3 responsiveness items of SERVQUAL), Physical Cues (measured by 1 reliability and 2 tangibles items of SERVQUAL and 4 merchandise quality items) and Merchandise Variety and Assortment (represented by 5 variety items). The items that did not load to a significant extent or did not demonstrate sufficient discriminant validity were deleted. The remaining 24 items were then subjected to a reliability analysis to assess the internal consistency of the instrument. All three factors, as well as the overall instrument, returned Cronbach alpha coefficients above the 0.7 level suggested by Peterson (1994). The three remaining empirical factors and their associated items are listed in Table 1.

The outcome of the first survey was that Figure 1 had to be re-configured slightly, as shown in Figure 2, to reflect the results of the scale purification process (Table 1).

Table 1: Empirical factor structure: Survey 1

EMPIRICAL DIMENSION	ITEMS MEASURING THE ISE DIMENSIONS
Personal interaction ($\alpha = 0.937$)	Employees of XYZ should always be willing to help me
	XYZ employees should give me personal attention
	Employees of XYZ should provide me with prompt service
	Employees of XYZ should be consistently courteous
	XYZ should give me individual attention
	Employees of XYZ should never be too busy to respond to my requests
	When I have a problem, XYZ should show a sincere interest in solving it
	I should have confidence in XYZ's employees
	Employees of XYZ should understand my specific needs
	Employees of XYZ should have the knowledge to answer my questions
	XYZ should have my best interests at heart
	XYZ should provide its services at the time it promises to do so
Physical cues ($\alpha = 0.863$)	The products available at XYZ should be of good quality
	XYZ should offer products of good quality
	XYZ should be known for good quality products
	Materials associated with XYZ's service (such as shopping bags, till slips or catalogues) should look good
	XYZ's products should not have defects
	XYZ's physical facilities should look good
	XYZ should provide error-free sales transaction records (till slips, credit card slips, accounts)
Merchandise variety and assortment ($\alpha = 0.857$)	XYZ should offer me a satisfactory choice of products
	XYZ should have all the products that I want to buy
	XYZ should offer a wide variety of products
	The products of XYZ should cater for a wide range of preferences
	XYZ should make products available in a variety of sizes

The second survey

A second empirical survey was then conducted, as suggested by Churchill (1979). The six items with the highest factor loadings that measured personal interaction were retained for the second empirical survey. Three of the items used to measure physical cues measured product quality, and it was decided to retain only one of these three items. The following item was added to measure physical cues: 'Products that function the way they are supposed to'. Two new items, apart from the five that were used to measure merchandise variety and assortment in the first survey, were added for the second empirical survey. These two items were: 'A choice of different brand names' and 'A good selection of well-known brands'. A total of sixteen items remaining from the scale purification process of phase 1, as well the three new items (in total nineteen items measuring Personal Interaction, Physical Cues and Merchandise Variety and Assortment) were combined with items to measure the three remaining (untested) dimensions of ISE, namely Internal Store Environment, Product Prices and Store Policies. The latter three dimensions were discussed in depth in Terblanche and Boshoff (2001b: 12-13). Figure 2 is a schematic presentation of the dimensions of ISE that were subjected to empirical assessment during the second survey.
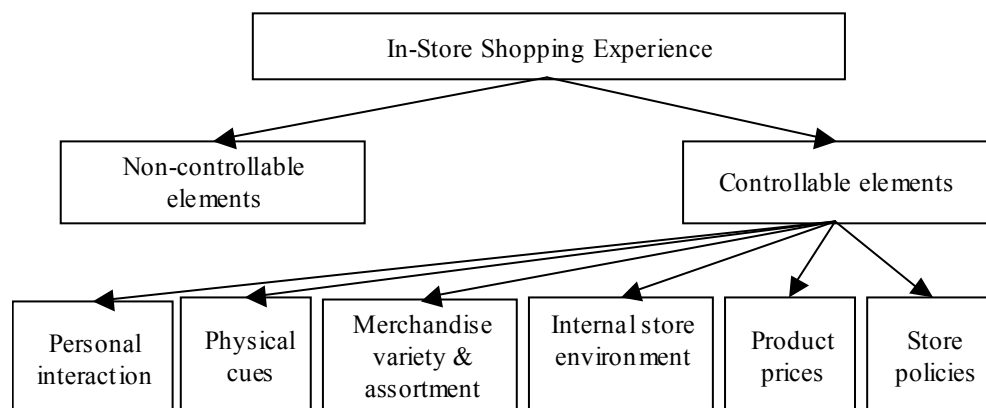

Figure 2: Schematic presentation of the dimensions of the ISE tested in the second phase of the study

Sample and data collection: Second survey

The sampling procedure used in the second survey was a combination of convenience and random sampling. Respondents were visitors to two regional shopping centers. Individual respondents (visitors) were selected on a simple random basis. Personal interviews, using a structured questionnaire, were conducted with visitors to the shopping center when they exited the center. The interviews were conducted over a period of two days to include all the different types of visitors who usually frequent these shopping centers. Respondents were again asked to rate the importance of the various components of ISE on a 7-point scale. A total of 2 504 questionnaires were completed, of which 1 197 were supermarket customers and 1 307 clothing store customers.

The instrument used in the second survey included the 19 items remaining after the scale purification process subsequent to the first survey. The 19 items mentioned above measured the dimensions Personal Interaction, Physical Cues and Merchandise Variety and Assortment in the second survey. The 'new' dimensions in the second survey were Internal Store Environment (9 items), Product Prices (5 items) and Store Policies (7 items).

Empirical results: Second survey

The second survey data analyses procedures again closely followed the guidelines for scale development suggested by Churchill (1979). To assess the discriminant validity of the instrument, a Maximum Likelihood Exploratory Factor Analysis was again conducted, specifying a Direct Quartimin oblique rotation (Jennrich & Sampson, 1966) of the original factor matrix.

Although it was expected that a six-factor solution would emerge (in line with Figure 2) several different factor solutions were considered. The most interpretable factor structure (factor loading exceeding 0,4 and no cross-loadings) to emerge, was a 5-factor solution (Table 2). All five factors in Table 2 had Eigen values above or very close to 1.00 and a sufficient number of items loading on them to a significant (0,40) extent (Hair *et al.*, 1998). The three factors retained from the first survey (Personal Interaction, Physical Cues and Product Variety and Assortment) remained fairly stable during the second survey. Of the 'new' dimensions added for the second survey, Store Policies emerged as a separate factor, as expected. Many of the items expected to measure Physical Cues and Product Prices, however, loaded on a common factor that was labeled Merchandise Value (Table 2). As some items expected to measure Physical Cues now loaded on the factor Merchandise Value, the remaining items were in fact measurements of Internal Store Environment, and were therefore labeled as such.

The items that remained, measuring the dimension Store Policies after the scale purification process, all referred to

the narrower concept Complaint Handling, rather than the broader Store Policies and it was re-named as such.

To summarise: the empirical factor structure that emerged after the second phase, suggested that the in-store shopping experience was multi-dimensional and consisted of five dimensions, namely: Merchandise Value, Internal Store Environment, Personal Interaction, Merchandise Variety, and Complaint Handling (see Figure 3) – all measured with scales that demonstrated more than adequate reliability (Table 2).

Convergent validity

Any measuring instrument should be both reliable and valid (Churchill 1979). A variety of different types of validity should be considered before any claims of validity can be made (Tull & Hawkins, 1993). To test the convergent validity of the ISE instrument, the total ISE score (mean 153,52 : SD 18,04) was correlated with scores that were expected to measure the adequacy of the retailer's parking facilities, satisfaction with in-store promotions, and the image of the store. It was expected that consumer perceptions of three retail issues (the adequacy of parking facilities, satisfaction with in-store promotions, and the image of the store) would be positively associated with their ISE scores.

The empirical results reported in Table 3 confirm this contention. The Pearson correlation coefficients shown in Table 3 reveal a consistent pattern of significant positive correlations with the total ISE scores, namely perceptions of adequacy of parking facilities (PARK, mean 6,30, SD 1,22), satisfaction with special in-store promotions (PROM, mean 5,32, SD 1,51) and the image of the company (IMAGE, mean 5,75, SD 1,41), suggesting at least some evidence of the convergent validity of the ISE instrument.

Table 2: Rotated factor loadings

	Factor1	Factor2	Factor3	Factor4	Factor5
Items	**Merchandise value**	**Store environment**	**Personal interaction**	**Merchandise variety**	**Complaint handling**
PRICE1	**0,624**	0,064	0,026	-0,055	-0,026
PHYS2	**0,623**	0,034	0,045	-0,013	0,025
PRICE2	**0,605**	0,108	0,017	-0,056	0,112
PHYS1	**0,601**	-0,118	0,028	0,002	-0,004
PRICE3	**0,512**	0,080	0,064	0,061	0,108
VAR1	**0,504**	0,033	-0,014	0,140	-0,059
VAR2	**0,448**	0,033	0,050	0,167	-0,016
PHYS3	**0,444**	-0,065	0,021	0,018	0,178
ENV8	0,025	**0,739**	-0,044	0,035	0,035
PHYS6	0,106	**0,695**	0,021	-0,082	0,100
ENV5	0,008	**0,655**	0,009	0,118	0,019
PHYS4	-0,056	**0,652**	0,100	0,071	-0,102
ENV6	-0,005	**0,402**	0,063	0,208	0,168
PERS3	-0,139	0,126	**0,654**	0,028	-0,029
PERS2	0,171	-0,092	**0,651**	-0,020	0,001
PERS4	0,034	0,036	**0,623**	0,049	0,090
PERS5	0,058	0,059	**0,489**	-0,017	0,206
PERS6	-0,045	0,012	**0,486**	0,086	0,306
PERS1	0,095	-0,005	**0,458**	0,015	-0,039
VAR3	-0,051	0,042	0,029	**0,707**	-0,036
VAR7	0,028	0,031	0,010	**0,688**	0,035
VAR4	0,054	0,021	-0,001	**0,672**	0,042
VAR6	0,126	0,040	0,019	**0,519**	0,108
POL6	-0,018	0,041	0,050	0,092	**0,737**
POL5	0,074	0,039	0,008	0,039	**0,653**
POL7	0,083	0,025	0,110	0,023	**0,640**
Eigen values	**8,52**	**2,61**	**1,61**	**1,11**	**0,94**
Cronbach Alpha	**0,819**	**0,832**	**0,813**	**0,800**	**0,817**

Table 3: Correlation matrix

	ISE	PARK	PROM	IMAGE
ISE	1,000			
PARK	0,414 0,0001	1,000		
PROM	0,466 0,0001	0,1900 0,0001	1,000	
IMAGE	0,448 0,0001	0,216 0,0001	0,469 0,0001	1,000

Construct validity

To assess the construct validity of the instrument, a confirmatory factor analysis (CFA) was conducted using LISREL 8,30 (Joreskog and Sörbom, 1999). The results summarized in Table 4 and Figure 3 suggest a reasonable fit of the model to the data ($\chi^2 = 2024,5$; df = 289; RMSEA = 0,0504; GFI = 0,939; AGFI = 0,925; and ECVI = 0,900) which provides additional evidence of the construct validity of the proposed instrument.

We realize that using the same data set to perform both an exploratory factor analysis and a CFA is often questioned, but it is certainly not without precedent (Finch & West, 1997). In the end we did proceed because the CFA results do provide evidence of construct validity (Tull & Hawkins, 1993). A second consideration was that this study is a long-term project with several phases and surveys, and that the CFA results of this survey would allow some comparisons with future survey (and particularly CFA) results.

Table 4: Model fit indices

Chi-square	2024,5
df	289
RMSEA	0,0504
ECVI	0,900
Normed fit index (NFI)	0,921
Comparative fit index (CFI)	0,931
Standardised RMR	0,047
GFI	0,939
AGFI	0,925
PGFI	0,773

By closely following the guidelines for multi-item scale development suggested by Churchill (1979), and based on the results of two empirical surveys, we concluded that there are five dimensions of importance to consumers when assessing their expectations/satisfaction with a in-store shopping experience. These dimensions are: Merchandise Value, Internal Store Environment, Personal Interaction, Merchandise Variety and Complaint Handling. These five dimensions and thus the In-store Shopping Experience, can be measured by means of 26 items. The proposed instrument in its current form (at that stage) demonstrated high levels of reliability, and some evidence of discriminant validity, convergent validity, and construct validity. Consistent with the guidelines suggested by Churchill (1979), the instrument in its then current form needed to be subjected to a third empirical assessment and particularly a cross validation assessment to provide conclusive evidence of its construct validity.

The third survey

Valid measurement is, according to Peter (1979), the *sine qua non* of science. Peter even argues that if a discipline does not use instruments that are valid and reliable, it cannot be regarded as a science. At the most basic level it means that the set of items making up a measuring instrument must measure only one thing in common (Hattie, 1985). Steenkamp and Trijp (1991) recommend that traditional methods of assessing the uni-dimensionality of an instrument (such as item-to-total correlations, exploratory factor analysis, reliability assessment) as described above and recommended by Churchill (1979) are particularly useful to reduce the original number of items, and to provide preliminary scales that can subsequently be tested and refined by means of a confirmatory factor analysis to assess the construct validity of the instrument. Gerbing and Anderson's Monte Carlo study (1992) produced empirical evidence to support this view. In a later study they concluded that ' …exploratory factor analysis can contribute to a useful heuristic strategy for model specification prior to cross validation with confirmatory factor analysis (Gerbing & Anderson, 1996: 62).
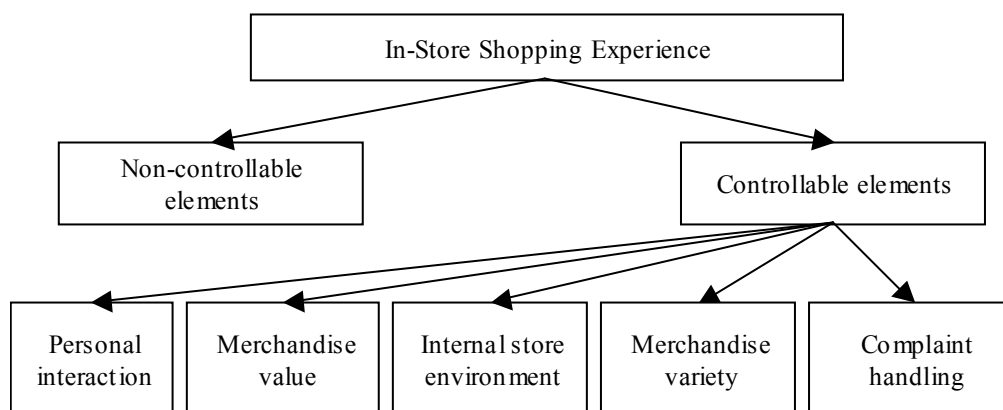

Figure 3: Empirical results of survey 2

The methodology of the third survey

The sampling procedure used for the third survey was again a combination of convenience and random sampling. Respondents were customers of a national clothing retailer and customers of a major national grocery supermarket. Individual respondents (customers) were selected on a simple random basis. Personal interviews, using a structured questionnaire, were conducted with customers of theses stores over a period of two different days to facilitate the cross validation of results. During the first day 1 686 retail shoppers (sample 1) were interviewed, and during day two 1 657 were interviewed (sample 2).

Unlike the first two phases, respondents were asked to rate their satisfaction with the in-store shopping experience of a particular store on a 7-point Likert-type scale. A total of 3 343 questionnaires were completed of which 2 096 were clothing shop customers (1 063 on day 1 and 1 033 on day

2) and 1 247 were supermarket customers (623 on day 1 and 624 on day 2).

The instrument used in the third survey included all 26 items remaining after the scale purification process after the first two surveys, measuring the following five dimensions: Personal Interaction (6 items), Merchandise Value (8 items), Merchandise Variety and Assortment (4 items), Internal Store Environment (5 items), and Complaint Handling (3 items).

Empirical results: Third survey

Confirmatory factor analyses results: Sample 1

In line with the recommendation of Steenkamp and Trijp (1991) and others, a confirmatory factor analysis (Maximum Wishart Likelihood estimation) of the data of sample 1 was conducted as it is able to provide, amongst others, evidence of uni-dimensionality. The kurtosis measures (normalized kurtosis = 291,5; relative kurtosis = 1,744) revealed, however, that the data were not adequately normally distributed. The analysis was then re-run using a more robust Maximum Likelihood by analyzing the asymptotic covariance matrix. Maximum Likelihood parameter estimates of an asymptotic covariance matrix are robust against moderate violations of multivariate normality, provided that the sample size exceeds 100 (Gerbing & Anderson, 1988).

The model fit statistics of the robust maximum likelihood analysis (χ^2 = 2362,2 (p < 0,000), χ^2/df ratio = 8,173; RMSEA = 0,068; GFI = 0,896; AGFI = 0,873; ECVI = 1,590; NFI = 0,910; RMR = 0,096; SRMR = 0,039) suggest that the measurement model closely fits the data in this sample.

Uni-dimensionality

Steenkamp and Trijp (1991) recommend that in the case of poor model fit, one assess the standardized residuals for the potential reasons. Although the measurement model fitted the data very well we nevertheless considered the standardized residuals. An inspection of the standardized residuals did reveal several values higher than the $|2.58|$ cut-off value proposed by Jöreskog and Sörbom (1988), which suggests that that some misspecification may have occurred. However, Steenkamp and Trijp (1991) caution that standardized residuals ought to be treated with circumspection as they are heavily influenced by deviations from multivariate normality, large sample sizes (1 686 in this case) and thus by the power of the test. We nevertheless decided to re-run the model after deleting three Merchandise Value items (MEVAL1, MEVAL6 and MEVAL7) for overfitting and one Personal Interaction item (PERIN6) for underfitting.

The improved model fit statistics (χ^2 = 1179,2 (p < 0,00), χ^2/df ratio = 6,326; RMSEA = 0.038; GFI = 0,936; AGFI = 0,919; ECVI = 0,463; NFI = 0,944; RMR = 0,070; SRMR = 0,029) suggest that the removal of the said four items was well justified.

Although there were still some standardized residuals higher than the $|2.58|$ cut-off value, no pattern was apparent. In other words, the ISE instrument demonstrated sufficient evidence of uni-dimensionality.

Confirmatory factor analyses results: Cross-validation (sample 2)

When developing a new scale, cross-validation is desirable because there is always the possibility that one has capitalized on chance. The ISE instrument was again administered to a sample of 1657 respondents (retail shoppers) similar to the previous sample described above. The offending items (MEVAL1, MEVAL6, MEVAL7 and PERIN6) with the high standardized residuals in the first sample were also included, and the *a priori* proposition that they needed to be removed was again empirically confirmed. The model fit statistics were (χ^2 = 2088,6 (p < 0,00), χ^2/df ratio = 7,225; RMSEA = 0,064; GFI = 0,904; AGFI = 0,884; ECVI = 1,450; NFI = 0,922; RMR = 0,086; SRMR = 0,036), very similar to the first sample. An inspection of the standardized residuals confirmed that the removal of the offending four items would be well justified. The confirmatory factor analysis model was thus re-run without them, resulting in a substantial improvement in model fit statistics (χ^2 = 892,4 (p < 0,00), χ^2/df ratio = 4,48; RMSEA = 0,031; GFI = 0,950; AGFI = 0,937; ECVI = 0,380; NFI = 0,959; RMR = 0,061; SRMR = 0,025).

These results confirm the earlier conclusion that the ISE instrument demonstrates excellent uni-dimensionality.

Within-method convergent validity

Several ways to assess the within-method convergent validity of an item have been proposed. The statistical significance of the regression co-efficient, the correlation of the item with the construct and the overall fit of the model are all indicators of within-method convergent validity. In this model all regression coefficients are strongly significant (p < 0,000) with the lowest t-value being 26,32), all items correlate significantly with each underlying dimension (lowest correlation coefficient is 0,519), whilst the overall model fit has already been alluded to. All these measures point to excellent within-method convergent validity.

Confidence

Steenkamp and Trijp (1991) also suggest that confidence in a model can be enhanced by an analysis in which the meaning of the construct is kept invariant by constraining all the parameters in the measurement model in a cross validation sample, to be equal to the measurement model parameters of the first sample. In other words, an assessment is done (in a second, cross validation sample) to determine whether or not the measurement model is identical across the two samples. The null hypothesis (H0) is that the factor loadings, measurement error variance, factor variances and co variances are all identical in both samples. The alternate hypothesis (H1) states that two or more of these parameters are different across the two samples. A Chi-square difference test is used to address H0 and H1. The model under the null hypothesis is fitted to the data by specifying

equality constraints across the two samples, while these equality constraints are relaxed to fit the model under H1. The Chi-square difference test statistic value is then obtained as the difference between the Chi-square test statistic values for the H0 and H1 models. The corresponding degrees of freedom are the difference between the degrees of freedom of the H0 and H1 models. The Chi-square difference test results for the in-store shopping experience measurement model are summarized in Table 5.

Table 5: Overall model cross-validation: Chi-square difference test results

	χ^2-value	df	p-value
Model under H0	2233,7	452	0,0000
Model under H1	2071,6	398	0,0000
Difference	162,1	54	0,0000

The small p–value (p < 0.001) for the Chi-square difference test statistic value in Table 5 suggests that there is sufficient evidence to conclude that the model differs across the two samples. In other words, the overall cross-validation of the measurement model for the total in-store shopping experience instrument is not supported by the data.

The results in Table 6 show that the measurement model did not cross-validate in its entirety (factor loadings, factor variances, factor co-variances, and measurement error variances). However, this does not imply that the cross-validations of the four specific parameter types are not supported. To test this proposition (that any one or all of the factor loadings, factor variances, factor correlations and error terms may be different in the two samples), a series of Chi-square difference tests were conducted.

The next step was then to determine whether or not the factor loadings were identical across the two samples. The null hypothesis (H0) states that all the factor loadings were identical across the two samples, while the alternate hypothesis (H1) states that at least two factor loadings are different across the two samples. As before, a Chi-square difference test is used to test H0 and H1. The results for this Chi-square difference test are listed in Table 6.

Table 6: Factor loading cross-validation: Chi-square difference test results

	χ^2-value	df	p-value
Model under H0	2093,5	415	0,0000
Model under H1	2071,6	398	0,0000
Difference	21,9	17	0,1872

The large p–value (p > 0,01) for the Chi-square difference test statistic value in Table 6 suggests that there is insufficient evidence to reject the null hypothesis (H0) that the factor loadings are identical across both samples. In other words, the data support the cross-validation of the factor loadings, as measured by the in-store shopping experience instrument, across the two samples.

To assess the cross-validation of the factor variances, the null and alternate hypotheses considered were:

H0: The five factor variances are equal across the two samples

H1: The five factor variances differ across the two samples

Table 7: Factor variances cross-validation: Chi-square difference test results

	χ^2-value	df	p-value
Model under H0	2082,6	403	0,0000
Model under H1	2071,6	398	0,0000
Difference	11,0	5	0,0510

The p–value (p > 0.01) for the Chi-square difference test statistic value in Table 7 suggests that there is insufficient evidence to reject the null hypothesis (H0) that the factor variances are identical across both samples. In other words, the data support the cross validation of the factor variances, as measured by the in-store shopping experience instrument, across the two samples.

To assess the cross validation of the factor co-variances, the null and alternate hypotheses considered were:

H0: The ten factor co-variances are equal across the two samples

H1: The ten factor co-variances differ across the two samples

A summary of the corresponding Chi-square difference test results are provided in Table 8.

Table 8: Factor co-variances cross-validation: Chi-square difference test results

	χ^2-value	df	p-value
Model under H0	2096,3	408	0,0000
Model under H1	2071,6	398	0,0000
Difference	24,7	10	0,0059

The small p–value (p < 0,01) for the Chi-square difference test statistic value in Table 8 suggests that there is sufficient evidence to conclude that the factor co-variances differ across the two samples. In other words, the cross-validation of the factor co-variances for the in-store shopping experience instrument is not supported by the data.

Finally, the cross validation of the measurement error variances had to be considered. For this assessment the following null and alternate hypotheses were tested:

H0: The 22 measurement error variances are equal across the two samples

H1: At least 2 of the 22 measurement error variances differ across the two samples

Table 9: Measurement error variances cross-validation: Chi-square difference test results

	χ^2-value	df	p-value
Model under H0	2169,9	420	0,0000
Model under H1	2071,6	398	0,0000
Difference	98,4	22	0,0000

The small p–value (p < 0.001) for the Chi-square difference test statistic value in Table 9 suggests that there is sufficient evidence to conclude that the measurement error variances differ across the two samples. In other words, the overall cross-validation of the measurement error variances for the in-store shopping experience instrument is not supported by the data.

In summary, the cross validation of the entire measurement model across the two samples is not supported by the data. This result is due to the fact that the cross validations of the factor co-variances and the measurement error variances are not supported by the data. However, the cross validations of the factor loadings and the factor variances are indeed supported by the data.

In other words, although the model did not cross validate in its entirety, Table 6 offers evidence that the factor loadings are identical in both samples. We want to argue that the fact that the factor loadings of the ISE measurement model are identical in two separate samples is, from a management perspective, a significant finding.

Reliability

Reliability is generally regarded as a necessary condition for validity (Peter, 1979). Table 10 shows that the internal reliability of the dimensions purported to make up the in-store shopping experience construct. All the Cronbach alpha co-efficients of the underlying dimensions were above the generally accepted cut-off value of 0,7 (Peterson, 1994), and for the whole scale it is 0,951.

Table 10: Reliability results

Personal interaction (PERIN)	0,875
Merchandise value (MEVAL)	0,841
Complaint handling (COHAN)	0,805
Internal store environment (STENV)	0,853
Merchandise variety and assortment (VAROS)	0,853
ISE SCALE	0,951

Discriminant validity

To assess the discriminant validity of the ISE instrument, each underlying dimension of the in-store shopping experience was considered separately. The null hypothesis (H0) in each instance is that one dimension (for example, PERIN: Personal Interaction) is perfectly correlated with all the other dimensions (MEVAL: Merchandise value, COHAN: Complaint Handling, STENV: Internal Store Environment, VAROS: Merchandise Variety and Assortment) in each of five separate chi-square difference tests. The alternate hypothesis (H1) is that at least two of these correlations are not perfect correlations.

The null and alternate hypotheses are tested by means of a Chi-square difference test. The Chi-square test statistic value for the model under H0 is obtained by specifying a model wherein the four correlations are fixed at unity, while that for H1 is obtained by specifying a model in which these four correlations are free parameters to be estimated. The Chi-square difference test statistic value is obtained as the difference between the Minimum Fit Chi-square test statistic values for the models under H0 and H1. The associated degrees of freedom are the difference between the degrees of freedom for the models specified by H0 and H1. In all these Chi-square difference tests, there are 4 degrees of freedom. The results for samples 1 and 2 are shown in Tables 11 and 12 respectively.

The extremely small p-values in Tables 11 and 12 confirm that each dimension is indeed a separate dimension from at

Table 11: Discriminant validity results: Sample 1

	χ^2-value for model under H0	χ^2-value for model under H1	Difference	p-value
Personal Interaction	2624,7	1179,2	1445,5	0,000
Merchandise Value	2160,3	1179,2	927,1	0,000
Complaint Handling	2208,2	1179,2	1029,0	0,000
Internal Store Environment	1481,1	1179,2	301,9	0,000
Merchandise Variety & Assortment	1668,5	1179,2	489,3	0,000

Table 12: Discriminant validity results: Sample 2

	χ^2-value for model under H0	χ^2-value for model under H1	Difference	p-value
Personal Interaction	2396,2	892,4	1503,8	0,000
Merchandise Value	1706,7	892,4	814,3	0,000
Complaint Handling	1662,6	892,4	770,2	0,000
Internal Store Environment	1778,3	892,4	885,9	0,000
Merchandise Variety & Assortment	2351,5	892,4	1459,1	0,000

least one other dimension of the in-store shopping experience, and thus provide further evidence that the discriminant validity of the ISE instrument is supported by the data of the two samples. In other words, the dimensions of the ISE are indeed separable from each other and they do not represent a single dimension.

To confirm these findings, we needed to assess whether or not each dimension of ISE is different from ALL the other dimensions (latent variables). This is accomplished by testing whether or not the correlation between any two dimensions is perfect. The corresponding hypotheses are:

H0: The correlation between the two latent variables (dimensions) is perfect

H1: The correlation between the two latent variables (dimensions) is not perfect

As before, a Chi-square difference test is used for each assessment. In this case, the Chi-square test statistic value for the model under H0 is obtained by fitting a model which specifies the correlation to be unity, to the data. The Chi-square test statistic value for the model under H1 is obtained by fitting a model which specifies the correlation as a free parameter to be estimated, to the data. Each of these Chi-square difference tests has one degree of freedom. The results of the 20 Chi-square difference tests for samples 1 and 2 are shown in Tables 13 and 14 respectively.

Table 13: Discriminant validity between individual dimensions: Sample 1

Dimension 1	Dimension2	χ^2-value for HO model	χ^2-value for H1 model	Difference	p-value
Personal interaction	Merchandise Value	1696,5	1179,2	517,3	0,000
	Complaint Handling	1353,3	1179,2	174,1	0,000
	Store Environment	1870,8	1179,2	691,6	0,000
	Variety & Assortment	2193,7	1179,2	1014,5	0,000
Merchandise Value	Complaint Handling	1725,5	1179,2	546,3	0,000
	Store Environment	1388,7	1179,2	204,5	0,000
	Variety & Assortment	1433,7	1179,2	254,5	0,000
Complaint Handling	Store Environment	1835,8	1179,2	656,6	0,000
	Variety & Assortment	2135,5	1179,2	956,3	0,000
Store Environment	Variety & Assortment	1500,4	1179,2	321,2	0,000

Table 14: Discriminant validity between individual dimensions: Sample 2

Dimension 1	Dimension 2	χ^2-value for HO model	χ^2-value for H1 model	Difference	p-value
Personal interaction	Merchandise Value	1382,5	892,4	490,1	0,000
	Complaint Handling	1040,5	892,4	148,1	0,000
	Store Environment	1295,6	892,4	403,2	0,000
	Variety & Assortment	2099,1	892,4	1206,7	0,000
Merchandise Value	Complaint Handling	1236,0	892,4	343,6	0,000
	Store Environment	1019,9	892,4	127,5	0,000
	Variety & Assortment	1186,8	892,4	294,4	0,000
Complaint Handling	Store Environment	1289,5	892,4	397,1	0,000
	Variety & Assortment	1641,8	892,4	749,4	0,000
Store Environment	Variety & Assortment	1778,3	892,4	885,9	0,000

The extremely small p-values in Tables 13 and 14 show that there is sufficient evidence to reject the perfect correlation hypothesis across both samples. In other words, the data from the two samples support the discriminant validity of the five dimensions of the ISE instrument.

Nomological validity

Nomological validity of a construct is assessed by investigating the relationships of the construct with other constructs in a nomological net. The relationships in the nomological net are based on a theoretical (causal) model for the constructs involved. Although this is often assessed by means of a correlation or regression analysis, these techniques do not allow for formal testing of the nomological net (theory) and they do not incorporate measurement errors for the latent constructs of the

nomological net (Steenkamp & Trijp, 1991). On the other hand, Structural Equation Modeling with Latent Variables does allow for measurement error and a formal test of the nomological net. Consequently, Structural Equation Modeling is a powerful statistical tool to assess the nomological validity of a construct.

In this study the nomological net that was tested was based on industry empirical evidence such as the PIMS studies of the 1990's, academic research (Rust & Zahorik, 1993; Sirohi, Mclaughlin & Wittink, 1998:240) and on the anecdotal evidence provided by theorists (Hoffman & Bateson, 1997: 290; Heskett, Jones, Loveman, Sasser & Schlesinger, 1994; Oliver, 1997; Zeithaml, Parasuraman & Berry, 1990: 9) and depicted in Figure 4. The bulk of evidence from all these sources suggests that satisfaction with the individual components of an in-store shopping

experience will result in customer satisfaction, which will lead to customer retention and loyalty over the long term. It is acknowledged, though, that this relationship is not always linear (Oliva, Oliver & MacMillan, 1992), particularly in a highly competitive industry with limited differentiation potential and low switching costs such as retailing (Jones & Sasser, 1995).

We decided to subject the ISE instrument to one final empirical assessment to assess its nomological validity, by testing the theoretical model depicted in Figure 4.

This time, the ISE instrument was administered to a sample of customers on the database of a retailer selling mainly cosmetics, house ware and gifts. Customer satisfaction was measured with a three-item instrument based on Anderson, Fornell and Lehmann (1994) and Macintosh and Lockshin (1997); loyalty was measured with a four item instrument based on Zeithaml, Berry and Parasuraman (1996); Sirohi *et al*., (1998) and East, Hammond, Harris and Lomax (2000). In total 34 000 questionnaires were mailed and 3 153 received back for an effective response rate of 9,27%.

Besides altering the type of retailer, we also used a mail survey as opposed to the mall intercept interviews used earlier. We believed that using a different type of retailer as well as a different type of data collection technique would add additional insight into the robustness of the ISE instrument.

As proposed by Steenkamp and Trijp (1991), LISREL 8.51 for Windows (Jöreskog & Sörbom ,2001) was used to fit the model depicted in Figure 4 to the data, to avoid the limitations associated with correlation and regression analyses. More specifically, the Weighted Least Squares (WLS) method for Polychoric Correlation Matrices was used. The WLS estimates of the path coefficients are shown in Table 15 (χ^2 = 2634,1; df = 334; RMSEA = 0,047; GFI = 0,978; AGFI = 0,974; NFI = 0,923). Table 15 shows that all the estimates are in the predicted direction and all, but the influence of Customer Complaint Handling (COHAN) on Customer Satisfaction, are statistically significant. In other words, the results summarized in Table 15 provide strong empirical support for the relationships of the dimensions of the ISE instrument with other constructs in a nomological net as predicted by theory.

In addition we re-assessed the reliability of each construct in the model. Not only did the ISE instrument return a high Cronbach alpha of 0,945, but each dimension's Cronbach alpha comfortably exceeded 0,8 (PERIN: 0,933; MEVAL: 0,851; COHAN: 0,885; STENV: 0,882; VAROS: 0,851). The instruments used to measure Customer satisfaction (α = 0,850) and Loyalty (α = 0,790) also proved to be reliable.

Table 15: Nomological assessment: Parameter estimates for model in Figure 4

Path	Parameter estimate (t-value)
CSAT → LOYALTY	0,981 (86,67)***
PERIN → CSAT	0,191 (2,65)**
MEVAL → CSAT	0,315 (3,79)**
COHAN → CSAT	0,108 (1,56)
STENV → CSAT	0,196 (3,90)**
VAROS → CSAT	0,194** (2,59)

***	= p < 0,001
**	= p < 0,01
*	= p < 0,05

Conclusion

Retailing consists of varied and diverse activities. There is more than anecdotal evidence that Monday morning shoppers are different from Wednesday afternoon shoppers, who are all different from weekend shoppers. Retail shopping is also varied in terms of the types of products bought. Marketing and retail textbooks point to the differences between convenience goods, shopping goods and specialty goods - not because marketers need to memorize product classifications but because we know that consumers behave differently when buying and shopping for different types of products. And the retail environment is further complicated by the diversity of retail formats such as catalogue retailing and more recently, electronic retailing.

Against this background an attempt to develop a generic instrument to measure satisfaction with in-store shopping experiences across such a diversity of retail circumstances is indeed an ambitious one. This is particularly true if one considers the problems experienced by others (such as SERVQUAL) who have attempted to develop generic instruments.

Despite the daunting challenge, this study attempted to develop a generic instrument that could be used to measure customer satisfaction with the controllable elements of the in-store shopping experience. The final questionnaire consists of 22 items measuring 5 dimensions of the in-store shopping experience that we believe are generic across all in-store retail environments, based on customer expectations (see Appendix A).
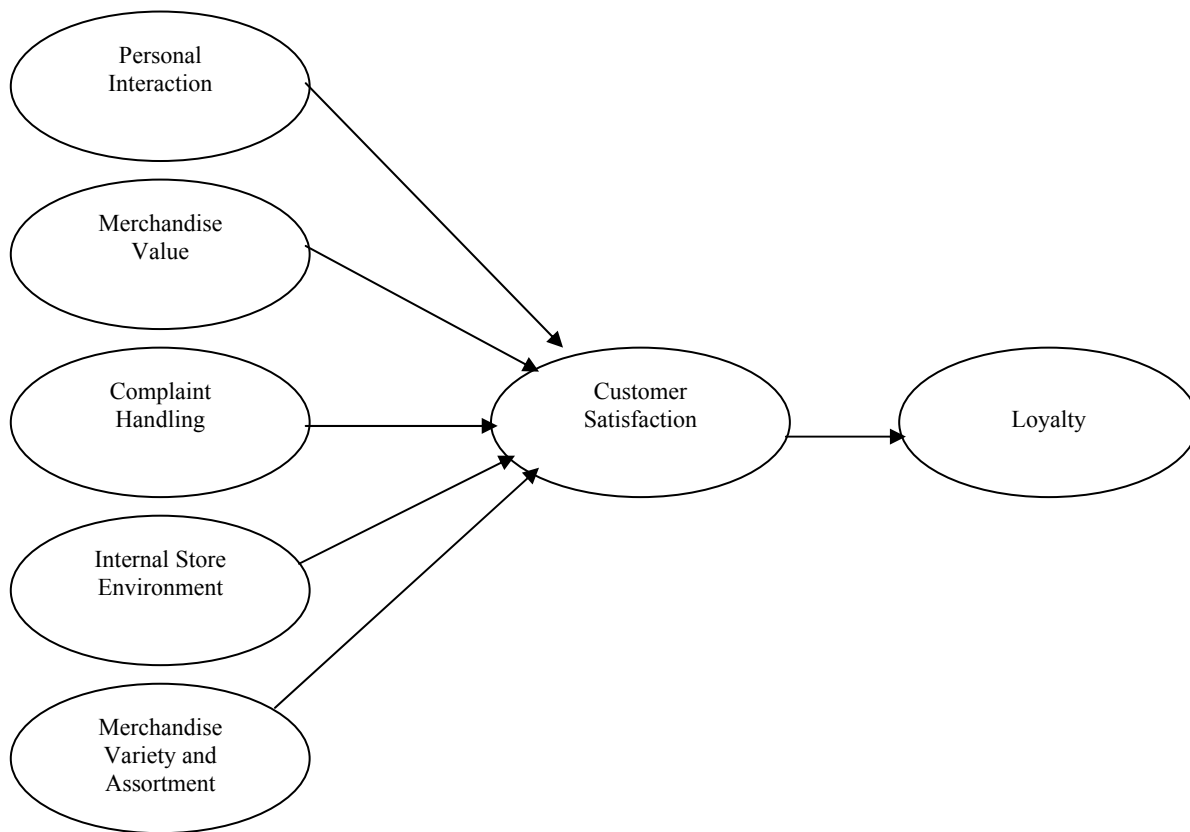
Figure 4: Theoretical model to assess nomological validity

We have closely followed the most contemporary guidelines for scale development involving 11 063 respondents in four different surveys. The evidence of the psychometric properties of the proposed ISE instrument offered here is compelling in terms of its unidimensionality, with-in method convergent validity, cross-validation of dimensions in a cross-validation sample (Tables 5-9), the reliability of the instrument (Table 10), its discriminant validity (Tables 11-14), and its nomological validity (Table 15).

We also would like to point out that the instrument has been administered to customers in a variety of different retail industries (fast food, clothing, grocery, hardware, cosmetics, gifts and house ware). Also it has proved to be robust across different data collection techniques (mall intercept and mail surveying).

The only area where the instrument did not succeed completely was in terms of the cross-validation of the entire model. From a management perspective, however, it is important that the factor loadings (or ISE dimensions) did indeed cross-validate successfully in the cross-validation sample.

Some critics may point to the fact that one of the five ISE dimensions (Customer Complaint Handling or COHAN) is not a significant predictor of Customer satisfaction in the nomological net depicted in Figure 4. We must point out that Customer complaint handling has consistently been shown to be an important issue to respondents. However, we have to acknowledge that most of our respondents have probably not lodged a complaint to the retailer about whom we asked them, and this may be the reason why there is no significant relationship between Customer Complaint Handling and Customer Satisfaction.

We suggest, however, that any measure of customer satisfaction at store level in a retail environment needs to include all five dimensions of the ISE to ensure that the uniqueness of any retail situation can be adequately captured.

References

Anderson, E.W., Fornell, C. & Lehmann, D.R. 1994. 'Customer satisfaction, market share, and profitability : Findings from Sweden', *Journal of Marketing*, **58**(3): 53-66.

Berman, B. & Evans, J. R. 1998. *Retail management.* 7th Edition. Upper Saddle River : Prentice Hall.

Berry, L. L. 1986. 'Retail businesses are service businesses', *Journal of Retailing*, **62**(1): 3-6.

Churchill, G.A. (Jr). 1979. 'A paradigm for developing better measures of marketing constructs', *Journal of Marketing Research*, **16**(1):64-73.

Dabholkar, P. A., Thorpe, D. I. & Rentz, J. O. 1996. 'A measure of service quality for retail stores: Scale development and validation', *Journal of the Academy of Marketing Science*, **24**(1): 3-16.

East, R., Hammond, K., Harris, P. & Lomax, W. 2000. 'First-store loyalty and retention', *Journal of Marketing Management*, **16**: 307-325.

Egan, J. 1999. 'Relationship marketing : A retailing perspective'. Paper presented at the 28th Annual EMAC Conference, Berlin, Germany.

Finn, A. & Kayandé, U. 1997. 'Consistency of the relationship between retailer product and service quality'. Paper read at the 4th Recent Advances in Retailing and Services Science Conference, Scottsdale, Arizona.

Finch, J.F. & West, S.G. 1997. 'The investigation of personality structure: Statistical models', *Journal of Research in Personality*, **31(4):** 439-485.

Gerbing, D. W. & Anderson, J.C. 1988. 'An updated paradigm for scale development incorporating unidimensionality and its assessment', *Journal of Marketing Research*, **25(**2**):** 186-192.

Gerbing, D. W. & Anderson, J.C. 1992. 'Monte Carlo evaluations of goodness of fit indices for structural equation modelling', *Sociological Methods and Research*, **21**(2): 132-160.

Gerbing, D. W. & Anderson, J.C. 1996. 'Viability of exploratory factor analysis as a precursor to confirmatory factor analysis', *Structural Equation Modelling*, **3**(1): 62-72.

Hair, J. F. (Jr), Anderson, R.E., Tatham, R. L. & Black, W. C. 1998. *Multivariate data analysis.* 5th Edition. Upper Saddle River : Prentice Hall.

Heskett, J.L., Jones, T.O., Loveman, G.W., Sasser, W.E. (Jr.) & Schlesinger, L.A. 1994. 'Putting the service-profit chain to work', *Harvard Business Review*, **72**(2): 164-174.

Hattie, J. 1985. 'Methodology review: Assessing the unidimensionality of tests and items', *Applied Psychological Measurement*, **9**(June): 139-164.

Hirschmann, E.C. 1981. 'Retail research and theory'. **In** Enis, B.E. & Roering, K.J.(Eds.). *Review of marketing.* Chicago : American Marketing Association.

Hoffmann, K.D. & Bateson, J.E.G. 1997. *Essentials of services marketing*, 4th Edition. Fort Worth : Dryden.

Hummel, J.W. & Savitt, R. 1988. 'Integrated customer service and retail strategy', *International Journal of Retailing*, **3**(2): 5-21.

Jennrich, R.I. & Sampson, P.F. 1966. 'Rotation for simple loadings', *Psychometrika*, **31**:313-323.

Jones, T.O. & Sasser, W.E. (Jr.) 1995. 'Why satisfied customers defect', *Harvard Business Review*, 73(6):88-99.

Jöreskog, K.G. & Sörbom, D. 1988. *LISREL*. Lincolnwood, IL: Scientific Software International, Inc.

Jöreskog, K.G. & Sörbom, D. 2001. *LISREL 8.51 for Windows* [Computer Software]. Lincolnwood, IL: Scientific Software International, Inc.

Kerin, R.A., Jain, A. & Howard, D.J. 1992. 'Store shopping experience and consumer price-quality-value perceptions', *Journal of Retailing*, **68**(4): 376-397.

Macintosh, G. & Lockshin, L.S. 1997. 'Retail relationships and store loyalty: A multi-level perspective', *International Journal of Research in Marketing*, **14**(5): 487-497.

Mazursky, D., & Jacoby, J. 1985. 'Forming impressions of merchandise and service quality'. **In** Jacoby, J. & Olsen, J. (Eds.). *Perceived quality: How consumers view stores and merchandise*. Lexington, M.A. : Lexington Books.

Oliva, T.A., Oliver, R.L. & MacMillan, I.C. 1992. 'A catastrophe model for developing service satisfaction strategies', *Journal of Marketing*, **56**(3): 83-95.

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

Oliver, R.L. 1980. 'A cognitive model of the antecedents and consequences of satisfaction decisions', *Journal of Marketing Research*, **17**(4): 460-469.

Parasuraman, A., Zeithaml, V. A. & Berry, L. L. 1988. 'SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality', *Journal of Retailing*, **64**(1): 12-40.

Peter, J.P. 1979. 'Reliability: A review of psychometric basics and recent marketing practices', *Journal of Marketing Research*, **16**(1): 6-17.

Peterson, R.A. 1994. 'A meta-analysis of Cronbach's coefficient alpha', *Journal of Consumer Research*, **21**(2)**:** 381-391.

Reichheld, F.F. & Sasser, W.E. 1990. 'Zero defections: Quality comes to service', *Harvard Business Review*, **68**(5): 105-111.

Rust, R.T. & Zahorik, A.J. 1993. 'Customer satisfaction, customer retention, and market share', *Journal of Retailing*, **69**(2): 193-215.

Sirohi, N., McLaughlin, E.W. & Wittink, D.R. 1998. A model of consumer perceptions and store loyalty intentions for a supermarket retailer, *Journal of Retailing*, **74**(2): 223-245.

Sopanen, S. 1996. *Customer loyalty schemes in retailing across Europe*. Oxford: Oxford Institute of Retail Management. Oxford Reports on Retailing.

Steenkamp, J-B.E.M. & Baumgartner, H. 2000. 'On the use of structural equation models for marketing modelling', *International Journal of Research in Marketing*, **17**(2/3): 195-202.

Steenkamp, J-B.E.M. & Trijp, H.C.M. 1991. 'The use of LISREL in validating marketing constructs', *International Journal of Research in Marketing*, **8**(4):283-299.

Terblanche, N.S. & Boshoff, C. 2001a. 'Measuring customer satisfaction with selected elements of the total retail experience : An exploratory study', *South African Journal of Economic and Management Sciences*, **4**(1):99-117.

Terblanche, N.S. & Boshoff, C. 2001b. 'The measurement of customer satisfaction with some of the controllable elements of the total retail experience : An exploratory study', *South African Journal of Business Management*, **32**(4):11-19.

Tull, D.S. & Hawkins, D.I. 1993. *Marketing research : Measurement and method.* 6th Edition. New York: Macmillan.

Zeithaml, V.A., Parasuraman, A. & Berry, L.L. 1990. *Delivering quality service: Balancing customer perceptions and expectations*. New York: The Free Press.

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

APPENDIX A

ITEMS TO MEASURE THE IN-STORE SHOPPING EXPERIENCE OF RETAIL CUSTOMERS[1]

Merchandise value	
MEVAL1	XYZ's products are of good quality
MEVAL2	Prices at XYZ store offer value for money
MEVAL3	XYZ's products function the way they are supposed to
MEVAL4	XYZ's product prices represent good value
MEVAL5	XYZ's products are free from defects and flaws
Internal Store Environment	
STENV1	XYZ has attractive décor
STENV2	XYZ has attractive physical facilities (check-out counters, shelves, etc)
STENV3	XYZ has attractive product and promotional displays
STENV4	XYZ has attractive materials associated with their service (shopping bags, catalogues, etc)
STENV5	XYZ has well-spaced product displays
Personal interaction	
PERIN1	XYZ's staff give me personal attention
PERIN2	XYZ's staff are always willing to help me
PERIN3	XYZ's staff provide me with prompt service
PERIN4	XYZ's staff are courteous
PERIN5	XYZ's staff are never too busy to assist me
Merchandise variety	
VAROS1	XYZ offers a choice of different brand names
VAROS2	XYZ offers a good selection of well-known brands
VAROS3	XYZ offers a variety of brand names that are available in many different sizes
VAROS4	XYZ offers a wide variety of products
Complaint handling	
COHAN1	XYZ has an effective means of dealing with customer complaints
COHAN2	XYZ has a fair system for the handling of complaints
COHAN3	XYZ staff efficiently deal with customer complaints

1)   All items have been re-coded