Market Research

Developing a Scale to Measure Service Quality: An Exploratory Study

Nusser Raajpoot, Khoon Koh, Anita Jackson

Introduction

Declining attendance, shrinking budgets, increasing public accountability and indirect competition from leisure services are some of the challenges facing museum managers (Kelly, 2004). It has been argued that the first step in responding to these challenges is to adopt a marketing and customer orientation by shifting from being about something to being for somebody (Weil, 1999). In implementing a customer orientation, museum managers are looking to develop strategies for retaining and increasing their customer bases. One strategy that has proved helpful in retaining customers at other types of service organizations, similar to museums, is to deliver high-quality experiences that result in higher levels of loyalty and profitability. Any improvement in service quality, however, would require a precise measurement tool.

The concept of quality is generally considered difficult to define and measure. Measurement becomes even more difficult when we consider the intangible nature of services. Although significant progress has been made in the last 20 years in developing service-specific quality measurement tools – in terms of both the development of core quality dimensions and adaptation/extension to various service contexts – little has been done in the area of cultural management, including museums and art galleries. The lack of a measurement scale suitable for capturing the unique nature of the museum experience, although understandable in the context of very recent adaptation of marketing practices by museums, is a major hurdle in improving service quality.

In order to understand the need for a museum-specific quality scale, one must appreciate the uniqueness of the museum visit. Museum services differ from other services in at least four ways. First, museums services can be seen as infrequently used. Infrequency of use may imply that expectations are not well formed (Arnould and Price, 1993), which in turn could mean that visitors will not be evaluating the museum experience at an attribute level but will instead be making a holistic evaluation that is not attribute-specific.

Second, a museum visit is an extended service encounter involving a sequence of related events occurring at different points in time. It is generally assumed that the visitor will make an overall assessment of the experience instead of separately evaluating each event during the visit. Although the average service performance during a visit is important, so too is the management of peaks and troughs of satisfaction at different time points (Verhoef, Antonides and Hoog, 2004). Services research has also focused on the contribution of various events during service delivery (Szymanski and Hise, 2000), emphasizing the need to manage both primacy and recency effects.

Third, the hedonic nature of a museum visit must be recognized. Hedonic experiences are

Nusser Raajpoot (PhD, University of Sydney, Australia) is Associate Professor of Marketing at Central Connecticut State University in the United States. His research interests are in the area of service quality measurement, design, evaluation and recovery, and his work has been published in leading academic journals.

Khoon Koh is Director of the Institute of Tourism and Hospitality Research at Central Connecticut State University. His research interests are in the areas of tourism entrepreneurship, destination marketing, service quality and higher education issues.

Anita Jackson (PhD, University of Arkansas) is Professor of Marketing at Central Connecticut State University, having previously taught at Louisiana State University. She teaches Fundamentals of Marketing, Consumer Behaviour and Marketing Research.

intense, positive and intrinsically enjoyable (Arnould and Price, 1993). They are emotionally intense, and emotions cannot be planned in advance. The spontaneous nature of the hedonic experience reflects a lack of specific expectations and the resultant difficulty of applying a perception-minus-expectation framework.

Fourth, motivation to visit museums is positive in nature and involves all aspects of positive consumption: sensory gratification (e.g., visual pleasure of looking at an art work); intellectual stimulation (e.g., viewing a rare object of an artistic or historical nature); and/or social approval (e.g., sensing the approval of friends and family). Motivations underlying consumption situations have been classified as either positive, where the goal is to achieve an enhanced positive emotional state, or negative, where the goal is to remove or avoid a negative experience and return to "normal" (Rossiter and Percy, 1997). The nature of the consumption motivation is an important consideration, as it largely determines the process of quality evaluation. While the services that are created to satisfy negative motives, such as insurance and health care, are likely to be evaluated at the attribute level, services that are created to achieve a positive state are likely to be evaluated at higher levels of abstraction.

There are many services that would, to varying degrees, illustrate one or more of these four differences. For example, most leisure services are hedonic, positively motivated, used infrequently and extended in nature. However, if we look at these differences in combination – especially the focus on education and linkage to terminal values – we see that museum services are sufficiently different to require their own unique scale. Some would make a case for adapting a valid service quality measurement scale such as SERVQUAL. In fact, various service organizations have successfully adapted SERVQUAL to their specific situations. However, SERVQUAL was developed and tested in five service industries (appliance repair/maintenance, retail banking, long-distance telephone service, security brokerage and credit card banking) in which the services could be classified as frequently purchased, negatively motivated, predominantly utilitarian (non-hedonic) and involving relatively brief service encounters. Since the development of SERVQUAL was based on totally different kinds of service industries, an adaptation/expansion of SERVQUAL may not help to capture the construct of service quality as it relates to the museum experience. Development of a measurement scale capable of taking into account the unique nature of museum visiting is therefore indicated.

The purpose of this study is to explore the domain of service quality as it relates to the museum experience, determine what visitors consider important when evaluating the quality of a museum experience and develop a scale to measure the quality of a museum experience.

The article is divided into five sections. In the first section we review the literature on service quality in museums. In the second section we define the generation of items through personal interviews and the literature. In the third section we describe our data-collection and purification procedures. In the fourth section we cover validation issues. In the final section we discuss managerial applications and guidelines for further research.

Literature Review

Service Quality

Quality is believed to be a concept that is abstract in nature. Many authors have described how hard it is to define quality (Deming, 1986; Cronin and Taylor, 1992). There are fundamental differences in the way that different authors

ABSTRACT

The authors develop a scale for measuring museum service quality. Using a sequence of exploratory and confirmatory factor analyses, they develop and test scale structure and dimensionality. The scale displays high internal reliability and convergent and discriminant validity. The findings suggest that museum service quality is evaluated less at attribute levels and more at higher abstraction levels, such as ambiance, learning and self-actualization. Managers can use the scale's dimensions to strategically design a museum experience.

KEYWORDS

Museums, service quality, scale development, exploratory and confirmatory factor analyses

have tried to define quality. For some it is superiority or excellence (Zeithaml, 1988), while for others it is a perception (Gronroos, 1990), a set of attributes (Hill, 1986) or a relation such as a gap (Gronroos, 1982; Lehtinen and Lehtinen, 1982).

Service quality defined as the gap between what was expected and what is delivered (also known as P-E conceptualization) has received strong support from the operationalization of SERVQUAL in terms of difference scores. Although this operationalization has been criticized for its statistical inadequacies, it remains the most popular and frequently used definition. Parasuraman, Zeithaml and Berry (1985) formally define service quality as the degree and direction of discrepancy between a customer's perception of and expectation of a service.

What constitutes the domain of service quality has been the subject of intense debate. Research in this area revolves around the dimensionality of the service quality construct. Early service quality research focused on the differentiation of two distinct dimensions of technical and functional quality (Gronroos, 1982). Technical quality refers to what the customer receives and functional quality refers to how it is delivered. Which of the two is more important depends on the nature of the service – that is, whether it is a pure service or a product-service mix. In subsequent research, greater attention was given to functional quality, with Parasuraman, Zeithaml and Berry (1985) proposing a five-dimensional conceptualization that overwhelmingly focused on service delivery. Based on focus group interviews and industry surveys, five dimensions of service quality – tangibility, assurance, empathy, reliability and responsiveness – were operationalized through SERVQUAL. Although Parasuraman and Zeithaml (2002) recognize that customers consider both the outcome and the delivery process in making quality evaluations, they do not propose a framework to incorporate outcomes in service quality measurement. Rust and Oliver (1994) propose a third dimension, service environment, in addition to the focus on technical and functional quality. This conceptualization incorporates the seminal findings of Bitner (1990) from her study of SERVICESCAPES.

In summary, there is little consensus on how service quality should be defined and measured. Definitions and measurement scales reported in the literature represent unique perspectives. It is clear, however, that P-E remains the most popular conceptualization of service quality and SERVQUAL remains the most frequently used scale for measuring it.

Service Quality Evaluations and Terminal Value Attainment

The marketing communications literature has emphasized the role of terminal values in understanding consumer motivations (East, 1997). Researchers in services marketing, however, have not actively explored the link between service experience and attainment of terminal values (de Ruyter and Bloemer, 1999). One reason for the lack of emphasis may be that service research has focused more on discrete, functional aspects of services, such as hotel reservations or bank transactions (Price, Arnould and Tierney, 1995). The need to link service experiences with terminal values becomes even more important in the case of services that involve "more phenomenological (i.e., gestalt) evaluation with sustained sensory and expressive content and ritualistic processes" (de Ruyter and Bloemer, 1999) such as museum visits.

Use of consumer values makes more sense in extended service settings. Price, Arnould and Tierney (1995) list three characteristics of extended services that separate them from other encounters: the duration of the service encounter is much longer than the usual two- or three-minute transaction, the emotional component

RÉSUMÉ

Les auteurs mettent au point une échelle pour mesurer la qualité du service offert par les musées. En utilisant une séquence d'analyses factorielles exploratoires et confirmatoires, ils développent et testent la structure et la dimensionnalité de cette échelle. Celle-ci montre un haut degré de fiabilité interne et de validité concourante et discriminante. Les résultats donnent à penser que la qualité du service offert dans les musées est évaluée moins selon des éléments physiques et davantage selon des éléments abstraits comme l'ambiance, l'apprentissage et l'épanouissement personnel. Les managers peuvent utiliser les dimensions de l'échelle pour concevoir stratégiquement l'expérience muséale.

MOTS CLÉS

Musées, qualité du service, création d'une échelle, analyses factorielles exploratoires et confirmatoires

of the encounter is dominant in making quality evaluations and extended services are more intimate in terms of spatial proximity. This description fits museum visiting perfectly. It could be argued that terminal value attainment – beyond achieving functional benefits – is an important factor in overall service evaluation. Terminal value attainment can therefore be considered an important dimension of the "gestalt" evaluation of extended services.

Thyne (2001) uses a laddering technique within a means-end framework to understand the linkage between museum visit motivations and terminal values. The notion that customers consider some values more important than others has important marketing implications. The author reports wisdom, warm personal relationships (belongingness), sense of accomplishment (self-actualization) and self-fulfilment as important consumer values.

The importance of value attainment in quality evaluation has also been endorsed by Zeithaml (1988) and, in terms of services, by Lages and Fernandes (2005). In proposing a model for understanding consumer perceptions of quality, Zeithaml introduces a construct labelled "higher level abstractions." According to means-end theory, quality information is retained by the consumer at various levels of abstraction, ranging from very simple (attribute) to very complex (terminal values). In this view, product quality can easily be assessed at a lower level of abstraction – that is, the attribute level – whereas most services can be evaluated only after consumption and therefore at much higher levels of abstraction – that is, personal values. In order to fully explicate the domain of service quality, one must consider all levels of abstraction, including terminal values.

A review of the literature on museum marketing, service quality and personal value attainment leads to three observations. First, an instrument suitable for measuring the museum experience is indicated. Second, the museum experience differs significantly from other service experiences in terms of its nature and customer expectations; it may therefore not be wise to adapt existing measures such as SERVQUAL, as these are unlikely to capture the domain of service quality in a museum visit. Third, value attainment is an extremely important determinant of the overall quality of a museum visit and should be incorporated into any measurement scale; the development of service quality measurement scales has given little consideration to value attainment. In this study we intend to fill these three gaps.

Domain Specification

Correct domain specification is one of the most important steps in scale development (Churchill, 1979). Both a literature review and in-depth personal interviews were used for domain specification. Semi-structured interviews were conducted with 18 students who had recently visited a museum. The students' participation was voluntary.

Results of an extant literature review and subsequent in-depth interviews indicated that perception-level museum evaluations can be categorized into 10 dimensions.

Pleasure

Pleasure has been defined as the state of being gratified, resulting in joy and delight. The hedonic nature of the museum visit has been revealed in several studies, as visitors mention the pleasure and emotions derived from their visit. Pleasure in the museum context has been defined as the joy one feels when viewing beautiful or aesthetically pleasing objects that add value to the museum experience (Passebois, 2002).

RESUMEN

Los autores desarrollan una escala para medir la calidad de servicio de los museos. Usando una secuencia de factores exploratorios y confirmatorios de análisis, desarrollan y ponen a prueba la estructura y dimensionalidad de la escala. La escala exhibe una alta fiabilidad interna, convergencia y validez discriminatoria. Los resultados sugieren que la calidad de servicio de los museos se evalúa menos a nivel de atributos y más a niveles abstractos superiores, como ambiente, aprendizaje y autoactualización. Los gestores pueden usar dichas dimensiones para diseñar estratégicamente la experiencia para los visitantes del museo.

PALABRAS CLAVE

Museos, calidad de servicio, desarrollo de escalas, análisis por factor exploratorio y confirmatorio

Relaxation

Relaxation is defined as refreshment of body or mind. In their study of a Texas museum, Tian, Crompton and Witt (1996) describe the hedonic dimension of the museum experience as one that relieves stress, is "relaxing, allows the visitor to get away from the usual demands of life" and "adds variety [to] a person's daily routine" (p. 38).

Learning (Education)

One of the core purposes of a museum is to educate. The visitor gains knowledge via historical recreations, art exhibits, guided tours and audio guides describing the works. Various museum studies describe different aspects of learning, such as allowing visitors to challenge themselves through new experiences and active participation (Hood, 1983) and satisfying curiosity/sense of discovery (Tian, Crompton and Witt, 1996).

Entertainment

Entertainment is something that amuses, pleases or diverts one's attention. People visit museums not only for educational purposes but also to enjoy a social outing. This point is well illustrated in Thyne's (2001) work, which discusses the significance and rising popularity of "edutainment," a term coined to describe a mixture of education and entertainment. Visitors' desire for and expectations with regard to edutainment is a key factor in their perception of the visit as satisfying.

Solitude

Solitude is defined as the state or quality of being alone or remote from others. It implies the absence of all others. In contrast to the need for social interaction, many individuals visit museums for the solitude and solace it brings them. According to Merriman (1995), the quest for contemplation and solitude is a compelling factor in the museum visit. The individual experience allows the visitor to internalize and meditate on the visit. The internalization and introspection of the visit seem to be intimately linked to what Passebois (2002) calls the spiritual experience. Social interaction is the other extreme of the construct of solitude. The role of the museum as a place for social interaction is recognized by several authors (Tian, Crompton and Witt, 1996; Caldwell, 2002; Goulding, 2000). The museum visit may be an outing that reinforces the bond between family and friends. It has been argued that a large number of museum visitors have socially oriented values and wish to "do something with their family and enjoy the company of those who came along" (Tian, Crompton and Witt, 1996, p. 38).

Self-actualization

The museum experience may be a way for individuals to seek solace and to view themselves as belonging to a higher order. The literature points to the deeper needs that are satisfied by a museum visit. Some museum visitors value "doing something worthwhile and having a challenge of new experiences" (Hendon, Costa and Rosenberg, 1989, p. 234). It has been reported that museum visiting makes one "feel like a better person" and increases one's sense of self-worth and "sense of accomplishment" (Tian, Crompton and Witt, 1996, p. 38). Furthermore, an individual may use a museum visit to differentiate himself as belonging to an elite group or to secure images of the self and enhance acceptance in significant groups (Goulding, 2000).

Social Recognition

Tian, Crompton and Witt (1996) report that the museum experience also has social recognition value. Visitors sense that others will think more highly of them and that they will impress others by demonstrating their intelligence or knowledge.

Courtesy

The impact of employee behaviour, including courtesy, is well documented in the service marketing literature. Although no museum study has explicitly examined the impact of employee behaviour on service quality evaluations, there is some evidence of the role played by employee behaviour in service quality evaluation in the leisure sector (Haahti and Yavas, 2005). Courtesy, defined as politeness, willingness or generosity in providing something needed, was cited as an important factor in most of the in-depth interviews.

Design

The importance of good design in delivering service quality is well established in the service literature (Baker, 1987; Bitner, 1990). Goulding (2000) discusses layout in terms of routing and mapping that facilitate visitors' movement through a museum. Easy access to areas of interest can enhance the perception of superior quality. Length of time spent in the museum and ability to gather and analyze information are partially dependent on the quality of design and layout. A good spatial design will allow for large numbers of people in an area without the perception of crowding. Design and layout "can serve to either enhance or detract from the visit" (Goulding, 2000, p. 273).

Aesthetics

Aesthetics has been defined as appreciation of beauty or good taste. Conforming to accepted notions of good taste, then, results in something that is aesthetically appealing.

Item Pool Generation

Three sources were used to generate an initial pool of items. First, items were drawn from the extant literature, including items representing measurement scales that tapped into the domain of museum service quality. The use of items with existing reliable and valid scales helped us to ensure that minimum errors were made in terms of wording, clarity and redundancy. Second, we tapped into another source of item generation: the population of interest or the population from which the survey respondents were drawn. A group of 35 students were asked to write down in their own words what they considered important in a museum visit. The most frequent responses, such as "museum visit provides me with sense of discovery," "I am able to express my knowledge authoritatively," "I feel like a better person" and "a museum visit is stimulating," can be easily accommodated under the themes identified in the literature review, such as education, social benefits, self-actualization and ambiance. Third, three local museum managers were also interviewed informally and their views were accommodated in the initial item pool generation.

In line with DeVellis's (2003) recommendation, the items were examined by four marketing academics for the purpose of evaluating face and content validity. These experts were asked to rate each of the items as either "not at all representative," "somewhat representative" or "clearly representative" of the construct definition. For an item to be retained, it had to be rated by the majority of the experts as "clearly representative." This step resulted in the deletion of some items. The remaining items were considered to have sufficient face and content validity and thus ready to be administered to a development sample for further scale purification.

Scale Development: Student Sample

For the purposes of data collection, we used a convenience sample of students (those enrolled in various sections of introductory marketing courses at the undergraduate level). The students were directed to a Web site where they could self-administer the questionnaire. To encourage student participation, respondents were awarded two bonus points towards their final grade.

We screened the respondents for being recent museum visitors. Anyone who had visited a museum in the preceding three months was classified as a recent visitor. We received a total of 196 responses, 188 of which were usable. The average age of respondents was 20 years and the male-to-female ratio was 61:39.

The survey data were analyzed using standard scale-development procedures for item reduction and refinement. We first conducted a reliability (coefficient alpha) analysis for 10 a priori conceptual dimensions, as listed in the section on domain specification. Item-to-total correlation values within each dimension were used as the basis for eliminating items that would improve the coefficient alpha. We continued eliminating items until alpha values for each dimension went above 7.0. Since alpha values are strongly influenced by the number of items, correlations between items and item redundancy, we took care to minimize the negative impact of these factors by eliminating item redundancy and parsimoniously selecting the items that represented a dimension. This process resulted in the elimination of 16 items, leaving us with 65 items and 10 dimensions.

In conducting factor analysis we chose principal component as the extraction method. A

factor analysis usually has the dual purpose of structure exploration and data reduction. In general, principal axis factoring is used when the primary focus is explaining factor structure and principal component is used when the primary focus is data reduction. In this case, we had a good theoretical basis for predicting domain structure, and thus our main concern was to represent that structure parsimoniously. We therefore selected principal component analysis.

These two extraction techniques, however, provide similar results if the data set displays high variable reliabilities and/or if it has a large number of items. The main difference between the two techniques is that principal component analysis uses the value of 1 in diagonal of the correlation matrix while principal axis factoring replaces diagonal values of 1 with reliability values. As reliability scores approach a value of 1, the results for these two techniques converge (Thompson and Vidal-Brown, 2001). In our study, all of the variables had alpha values higher than 0.7.

An initial unrotated factor analysis asking for ten dimensions yielded a solution in which only nine factors had an Eigen value greater than 1, explaining 67% of the variation in the data set. Of the nine factors, two were meaningless, as all variable loadings on them were below 0.5. This meant that the data only partially supported the ten-factor exhaustive model and we were left with a seven-factor solution.

At this stage, all items with low loadings on all factors or high cross-loading on two or more factors were eliminated and the remaining items were subjected to an a priori seven-factor solution. A varimax rotation provided a much clearer solution. Again, low-loading or cross-loading items were deleted. This iterative process resulted in a 27-item, 7-factor initial scale, as shown in Table 1. The factors were labelled as follows:

1. *Employee behaviour* (α = 0.926): civil, courteous and friendly employee behaviour

2. *Self-actualization* (α = 0.868): the extent to which the museum visitor has a sense of accomplishment and self-worth following the visit

3. *Customer compatibility* (α = 0.929): the extent to which the presence of other visitors enhances or detracts from the museum experience

TABLE 1

EXPLORATORY FACTOR ANALYSIS

	Cronbach alpha	Factor loadings
Employee behaviour	0.926	
1. Employees very courteous		.735
2. Employees civil		.745
3. Employees friendly		.743
Self-actualization	0.868	
4. Feelings of self-worth		.757
5. Feel like a better person		.857
6. Sense of accomplishment		.808
Customer compatibility	0.929	
7. Identify with the typical museum visitor		.880
8. Behaviour of other visitors irritated me		.874
9. Visitor is very much like me		.857
Crowding	0.954	
10. Museum is not crowded		.942
11. Does not have too many people		.943
Design/layout	0.910	
12. Designed aesthetically		.681
13. Simple layout		.453
14. Well designed		.744
15. Easy to follow signage		.803
16. Provide detailed maps		.871
17. Exhibits are clearly identified		.854
Learning/education	0.862	
18. Satisfies my curiosity		.501
19. Sense of exploration		.655
20. Sense of discovery		.578
Ambiance	0.902	
21. Museum is cheerful		.715
22. Museum is stimulating		.659
23. Museum is lively		.682
24. Very pleasing activity		.706
25. Sense of excitement		.796
26. Sense of adventure		.633
27. Novel experience		.640

4. *Crowding* (α = 0.954): the negative impact of spatial crowding on the museum experience

5. *Design/layout* (α = 0.910): the impact of design/layout on the museum experience

6. *Education* (α = 0.862): the extent to which a museum visit adds to the visitor's knowledge

7. *Ambiance* (α = 0.902): the overall feeling about the visit

Of the original domain specification, the dimensions of solitude, relaxation, social recognition and social interaction were not supported by the data. Pleasure and entertainment were partially captured in the ambiance dimension. Learning/education, self-actualization, courtesy and design, as specified in the domain, found support. On the other hand, the data uncovered the new dimensions of customer compatibility and crowding. The collapse of a large number of theoretically defined dimensions into a smaller set of dimensions during scale development is normal and at times can be expected. During the development of SERVQUAL, the authors began with ten dimensions and finished with five.

Scale Confirmation: Museum Visitor Sample

Having established the 7-factor, 27-item scale through exploratory factor analysis, we proceeded to collect data to confirm the factor structure and assess its validity. The data for confirmation were collected at a museum of natural history in a New England city (in the United States). Visitors were approached as they exited the museum and were invited to participate in a self-administered survey. Participation was voluntary. Once they agreed, respondents were led to a makeshift booth where they could complete the survey in private. The average time taken to complete the survey was 25 minutes. A male-to-female ratio of 46:54 ensured balanced gender representation. Almost all of the respondents (98%) had one or more children, which suggests that the museum visit was a group (family) affair.

We first ran confirmatory factor analysis on the newly collected data at the museum using AMOS version 6. We did not expect the exploratory factor analysis structure to be fully confirmed, as student data are typically less reliable than data collected in the field. As expected, the 7-factor specification failed to converge. When we analyzed the modification indices, we thought that removal of the construct "crowding" might cause the data to converge. Removing a construct solely on the basis of modification indices is not a sound strategy. It is, however, possible to justify removal if one considers the fact that part of the "crowding" construct was captured by the "customer compatibility" construct. The 6-factor solution – that is, after removal of the "crowding" construct – produced a valid structure. Further, weak-loading items had to be removed from the model to improve fit statistics. The confirmed model, as shown in Table 2, displayed reasonable fit statistics (χ^2 / *df* = 2.12, CFI = 0.935, NFI = 0.921, TLI = 0.921, RMSEA = 0.068 ... Lo .055 Hi 0.080). The final structure, after removal of the weak items, is displayed in Figure 1.

Individual item dimensionality is assumed when it measures only one construct. In other words, the unique variance of each item should be the only measurement error and should not be significantly correlated with the unique variance of another item. Within the structural equation model we looked at the standardized residuals and none of these had a value above 2.58 (a value at which p = < .01), thus confirming item unidimensionality. At the scale level, the average interfactor correlation was 0.32, further supporting the unique factor structure.

In line with the recommendation of Fornell and Larcker (1981), discriminant validity was assessed by comparing the squared correlation of any two dimensions with the variance extracted estimates. If both variance extracted estimates are greater than the squared correlation, discriminant validity is demonstrated. Table 3 shows the results. For example, the correlation between self-actualization and customer compatibility is 0.344 and the squared correlation 0.119. The variance extracted estimate is 0.629 for customer compatibility and 0.618 for self-actualization. Because the variance extracted estimates are greater than the square of the interfactor correlation, the test supports the discriminant validity of these two dimensions. Examination of the other variance extracted estimates and squared correlation coefficients supported discriminant validity within the model.

High factor loadings (0.61–0.93) and alpha values (> 0.89) indicate a strong association between constructs and items measuring them.

TABLE 2

RESULTS OF EXPLORATORY AND CONFIRMATORY FACTOR ANALYSES

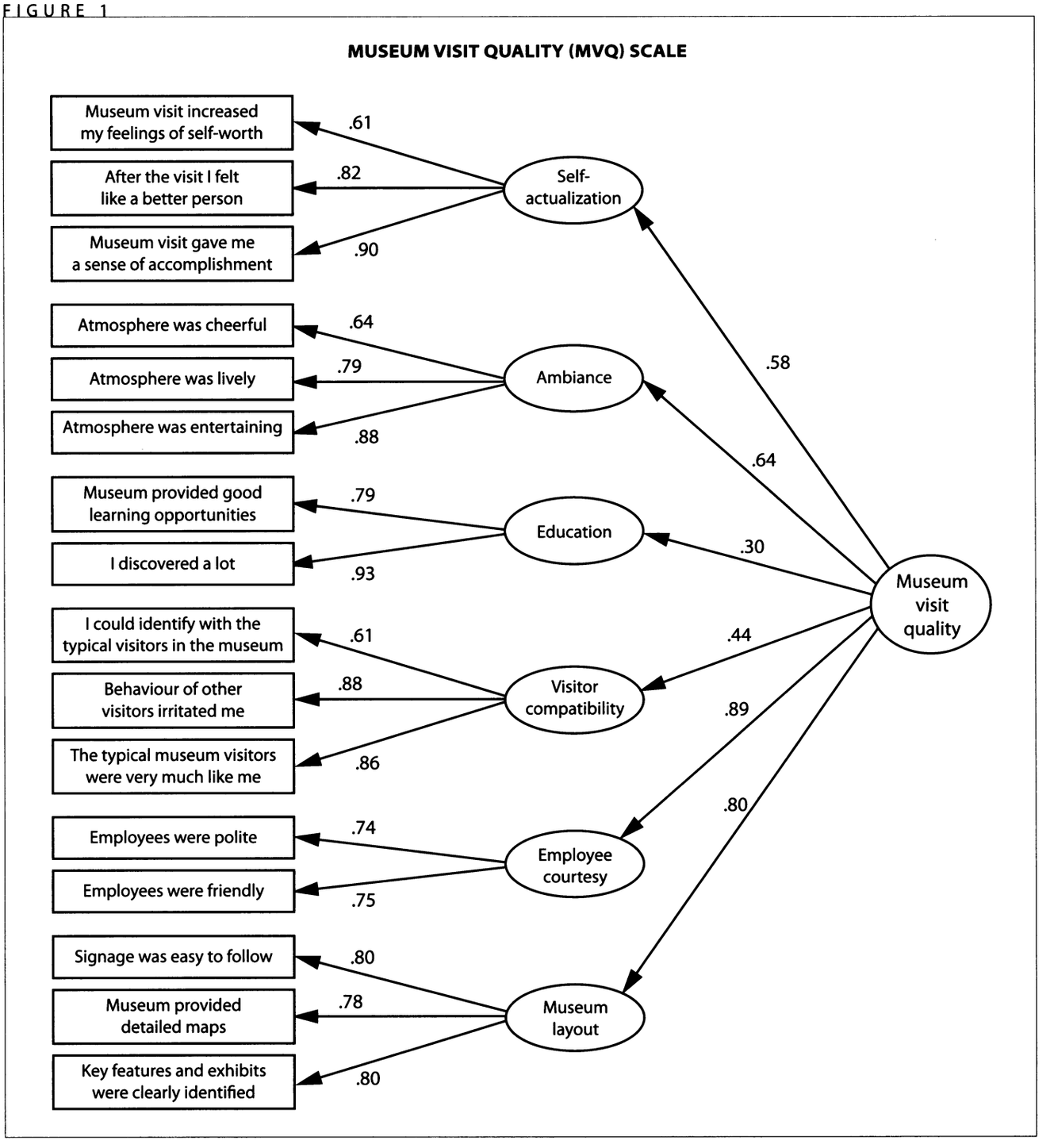
Factors	Cronbach alpha (dimensions)	EFA rotated loadings	CFA standardized loadings
Employee behaviour	0.926		
Employees very courteous		.735	.72
Employees civil		.745	.28
Employees friendly		.743	.78
Self-actualization	0.868		
Feelings of self-worth		.757	.61
Feel like a better person		.857	.82
Sense of accomplishment		.808	.90
Customer compatibility	0.929		
Identify with the typical museum visitor		.880	.61
Behaviour of other visitors irritated me		.874	.88
Visitor is very much like me		.857	.86
Crowding	0.954		
Museum is not crowded		.942	Dropped
Does not have too many people		.943	Dropped
Design/layout	0.910		
Designed aesthetically		.681	Dropped
Simple layout		.453	Dropped
Well designed		.744	Dropped
Easy to follow signage		.803	.80
Provide detailed maps		.871	.78
Exhibits are clearly identified		.854	.80
Learning/education	0.862		
Satisfies my curiosity		.501	Dropped
Sense of exploration		.655	Dropped
Sense of discovery		.578	.95
Good learning		.634	.78
Ambiance	0.902		
Museum is cheerful		.715	.64
Museum is stimulating		.659	Dropped
Museum is lively		.682	.80
Very pleasing activity		.706	Dropped
Very entertaining		.796	.88
Sense of adventure		.633	Dropped
Novel experience		.640	Dropped

Fit Statistics: $\chi^2 = 233.4$; Df = 110; χ^2 / df = 2.12; CFI = 0.935; NFI = 0.936; TLI = 0.921; RMSEA = 0.068

Legend: CFA = Confirmatory Factor Analysis; EFA = Exploratory Factor Analysis; CFI = Comparative Fit Index; NFI = Normed Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Estimation.

Total variance extracted = 71% with Varimax Kaiser Normalization Rotation

FIGURE 1



Reasonable loading (0.30–0.8 with p = .000) between first- and second-order constructs also attests to convergent validity. Convergent validity was further assessed by looking at the association between scale scores and the scores on questions asking respondents to provide an overall measure of visit quality by checking one of three categories – excellent, fair or poor. One-way ANOVA was used to examine the correspondence between scale score and overall quality responses. The results are displayed in Table 4. The scale scores for those in the "excellent" category were significantly higher than those in the "fair" category. Similarly, scores for those in the "fair" category were significantly higher than those in the "poor" category.

TABLE 3

TEST OF DISCRIMINANT VALIDITY

	Employee behaviour	Self-actualization	Customer compatibility	Design/ layout	Learning/ education	Ambiance
Employee behaviour	***0.402****					
Self-actualization	0.089	***0.618****				
Customer compatibility	0.046	0.119	***0.629****			
Design/layout	0.111	0.052	0.064	***0.629****		
Learning/education	0.084	0.178	0.151	0.139	***0.486****	
Ambiance	0.104	0.035	0.077	0.141	0.070	***0.608****

*Average Variance Extracted (AVE) for the column dimension

Nomological validity is assessed by investigating the relationship between the construct of interest (i.e., service quality) and other constructs that are theoretically related to it. A number of studies have found that consumers receiving a good-quality service are likely to be satisfied and thus to repurchase (see, e.g., Dabholkar, 1995; Caruana, 2002). To test the nomological validity of the 6-dimension, 16-item scale, we calculated the correlations between individual dimensions and the constructs of satisfaction and repurchase intention. The results are presented in Table 5. As can be seen, all of the correlations are in the expected direction and are significant, except the correlation between self-actualization and repurchase intention. This lack of significance may be attributable to the fact that achieving a sense of accomplishment or of being a better person would not require repeat visits. A few other results are worth noting. The strong correlation (0.62) between visitor compatibility and repurchase intention highlights the need to manage visitor segments. The probability of revisiting is high when people visit museums in groups of like-minded individuals. Similarly, the strong correlation between ambiance and satisfaction highlights the role of ambiance in delivering satisfaction.

Discussion

Of the six final dimensions, only one – design and layout – proved to be directly related to physical attributes. This finding is not surprising considering the overwhelming hedonic nature of museum visiting. Hedonic refers to the multisensory nature of the consumption experience, which includes taste, sound, tactile impressions and visual images (Hirschman and Holbrook, 1982). In hedonic situations the focus is on fun, feelings and fantasy rather than on the physical attributes of the service. The issues of customer compatibility, ambiance and employee courtesy are directly related to the hedonic nature of museum visiting, while the constructs of self-actualization and education are partially hedonic in nature. This focus on hedonism does not diminish the importance

TABLE 4

TEST OF CONVERGENT VALIDITY

	Overall quality					
	Excellent		Fair		Poor	
Self-actualization	6.3	*	4.2	*	3.6	*
Ambiance	6.1	*	4.3	*	2.9	*
Education	6.4	*	5.1	*	3.1	*
Museum layout	5.9	*	3.8	*	2.9	*
Visitor compatibility	6.2	*	4.3	*	3.1	*
Employee courtesy	6.5	*	4.2	*	2.8	*
		*		*		*
Combined scale score	6.2	*	4.3	*	3.1	*

*All scores are significantly different within each dimension and for combined scale scores.

TABLE 5

SUMMARY OF NOMOLOGICAL VALIDITY ASSESSMENTS

Employee courtesy	
Satisfaction	0.31**
Repurchase intention	0.28*
Museum layout	
Satisfaction	0.35**
Repurchase intention	0.21*
Ambiance	
Satisfaction	0.60**
Repurchase intention	0.51**
Education	
Satisfaction	0.44**
Repurchase intention	0.32**
Self-actualization	
Satisfaction	0.23*
Repurchase intention	0.18
Visitor compatibility	
Satisfaction	0.41**
Repurchase intention	0.62**

*Significant at .05 (2-tailed)
**Significant at .01 (2-tailed)

of the physical aspect of the visit. It is worth remembering that the hedonic experience is delivered through a combination of exhibits, design and people. It appears that, when evaluating museums, visitors take into account all physical aspects, such as exhibits and the building itself, but tend to evaluate them in terms of a holistic outcome such as ambiance.

Confirmation of the customer compatibility dimension is a significant result. Although the role of the museum as a place of social interaction is well documented in the literature, the focus has been on interaction within groups of friends and family members. The impact of other visitors in the museum has not been investigated. It seems that the presence of visitors similar to oneself in terms of appearance and behaviour can greatly affect the overall experience. In a service such as a museum, where visitors share facilities, space and time, the role of others in the final evaluation should not be underestimated. The theory of social distance suggests that people prefer to interact with members of their own social class. They will interact (albeit to a lesser degree) with those in a higher social class but will avoid those in social classes lower than their own. Social identity regulates visitor behaviour such that visitors with certain beliefs tend to associate with those who hold similar beliefs and to disassociate from those who do not (Smith, Bruner and White, 1956). Social identity is known to influence brand loyalty and organizational patronage.

Although self-actualization and ambiance were not a surprise, they are important enough to merit discussion. In describing purchase motivations for products and services, Rossiter, Percy and Donavan (1991) list two basic sets of these motivations. People are known to purchase products and services either for the purpose of problem removal/avoidance or for the purpose of improving their existing sensory, mental or social state. Sensory gratification and intellectual stimulation can be considered two important motivations for visiting a museum. Sensory gratification can be a function of exhibit quality and ambiance whereas intellectual stimulation can be delivered through the authenticity and creativity of the museum's collection. Intellectual stimulation is also closely related to learning (education). It was, therefore, easy to speculate that visitors would be evaluating the visit for the motivational objectives of self-actualization and ambiance.

The absence of exhibit quality among the final dimensions is rather surprising, particularly since respondents in the in-depth interviews consistently mentioned it as an important factor. One possible explanation for this is the fact that the impact of exhibit quality was captured via the higher-order value constructs of education, ambiance and self-actualization. Authenticity, an integral part of exhibit quality, can be a part of education, as little knowledge is imparted unless the exhibits are authentic. A quality exhibit is expected to be entertaining and lively. The ambiance dimension explicitly captures both of these effects.

Managerial Implications

Two kinds of implications for museum managers can be drawn from the results. The first kind relates to service businesses in general, while the second relates to hedonic-specific services such as museums. At a general level, the importance of design and employee

behaviour in delivering good service experiences is recognized and well documented. To improve the quality of human interactions, museum managers would do well to focus on employee training, empowerment and fair remuneration and on visitor complaint management. Museum managers should also focus on choosing designs that are compatible with the nature of the exhibits and with visitors' movements, making it easy for visitors to move around and locate objects of interest to them.

In designing a spatial experience, museums should allow visitors to take their time and to learn without a fixed format. Surprise elements such as variation in light and space, places for visitors to reflect and rest, and reduction in noise and congestion are all elements of good design. Visitors should be able to find their way through the museum in a logical sequence. It is easy to lose one's way in a large museum! While a good map design is key to way-finding, technological features such as location-based audio guides can also be useful. Museum managers could think about podcasting as a way to customize way-finding – allowing visitors to select their preferred exhibit sections.

The role of museums in knowledge enhancement is well documented. Our findings support those reported in the literature. Managers would do well to follow the extensive guidelines available on this subject. It is not enough to provide supporting materials such as background information on exhibits. Visitors learn by questioning, touching, hearing, seeing and playing. Museum managers need to consider using multiple media to explain and educate. Providing opportunities for visitors to touch can serve to accelerate their learning. Walking through a hall full of specimens lined up on shelves and encased in glass may not be best way to learn. It is important for museum managers to make use of technology when presenting learning opportunities.

The issue of customer compatibility is an important one. Traditionally, a museum visitor has been profiled as an active community member belonging to upper educational, occupational and income groups. This profile suggests that compatibility will not be an issue. It appears, however, that museums, especially large ones, attract visitors of significant variation in demographic and psychographic profile. What can be done to improve compatibility? Typically, the exhibit type will determine the kind of visitor it attracts. Therefore, narrowly defining museum exhibits will enhance compatibility. Another factor in customer compatibility is the museum's location. Museums in upscale areas attract upscale visitors and museums in suburban areas attract suburban visitors.

Museums can also try to group visitors in such a way as to increase within-group homogeneity. The purpose of such grouping is to minimize irritation and conflict within the group (Pranter and Martin, 1991). For example, it may be possible to separate experts from laypersons, regular visitors from first-time visitors, and people with children from those without children. Similarly, age-based grouping can enhance compatibility. Designated smoking and non-smoking areas in hotels is an example of grouping similar customers.

Limitations and Future Research

As with any study, the present one has several limitations. Collection of validation data at only one museum, the Peabody Museum of Natural History at Yale University, is the most obvious limitation. Although one can expect the visitors of natural history museums to be similar to those of other museums, it would have been useful to include visitors of other museum types, such as art or science and technology, in the validation sample. Extreme care should therefore be taken in assuming generalizability of the scale. It should be viewed as a preliminary scale, because it is based on two samples of modest size that do not permit comprehensive psychometric assessment.

At the exploratory stage, we used student data. Although a student sample represents a much richer opinion pool than a single museum data set, it was a convenience sample nonetheless. Establishing scale validity is a process that takes time. Future research could focus on validating this scale across a variety of situations and with a variety of samples. Replication in other museum types and in different cultures would increase confidence in the scale.

Future research on museum quality could make several extensions of this study. For example, the scale does not expand on the relative importance that visitors attach to each of the dimensions. Different dimensions are expected to make varying contributions towards the overall holistic evaluation. This knowledge of relative importance is extremely useful in developing a

quality-improvement strategy. Given their limited financial resources, museums need to spend wisely. Knowledge of which of the six dimensions visitors consider most important provides a good basis for an improvement strategy.

The scale developed in this study makes it possible to investigate the impact of the consumers' perception of quality on some important marketing constructs, such as satisfaction, loyalty, value, repurchase intention and profitability. Future research could focus on confirming some commonly held notions about the positive impact of quality on satisfaction and so forth. Similarly, the scale could be used to examine the antecedents of customer-perceived quality. Linking the scale dimensions to specific service attributes, such as authenticity of artifacts, new arrivals and ease of access, will provide a basis for developing an effective marketing strategy.

References

Arnould, E.J., and L.L. Price. 1993. "River Magic: Extraordinary Experience and Extended Service Encounter." *Journal of Consumer Research,* Vol. 20, June, p. 24–45.

Baker, J.A. 1987. "The Role of Environment in Marketing Services: The Consumer Perspective." In *The Services Challenge: Integrating for Competitive Advantage,* J.A. Czepiel, C.A. Congram and J. Shanahan, eds. (p. 79–84). Chicago: American Marketing Association.

Bitner, M.J. 1990. "Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses." *Journal of Marketing,* Vol. 54, nº 2, p. 69–82.

Caldwell, N. 2002. "(Rethinking) the Measurement of Service Quality in Museums and Galleries." *International Journal of Nonprofit and Voluntary Sector Marketing,* Vol. 7, nº 2, p. 161–171.

Caruana, A. 2002. "Service Loyalty: The Effects of Service Quality and the Mediating Role of Customer Satisfaction." *European Journal of Marketing,* Vol. 36, nº 7/8, p. 811–828.

Churchill, G.A. 1979. "A Paradigm for Developing Better Measures of Marketing Constructs." *Journal of Marketing Research,* Vol. 16, nº 1, p. 64–73.

Cronin, J.J., and S.A. Taylor. 1992. "Measuring Service Quality: A Reexamination and Extension." *Journal of Marketing,* Vol. 56, nº 3, p. 55–68.

Dabholkar, P.A. 1995. "The Convergence of Customer Satisfaction and Service Quality Evaluations With Increasing Customer Patronage." *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior,* Vol. 8, p. 32–43.

Deming, W.E. 1986. *Out of the Crisis.* Cambridge, MA: Centre for Advanced Educational Services, Massachusetts Institute of Technology.

de Ruyter, K., and J. Bloemer. 1999. "Customer Loyalty in Extended Service Settings: The Interaction Between Satisfaction, Value Attainment and Positive Mood. *International Journal of Service Industry Management,* Vol. 10, nº 3, p. 320–336.

DeVellis, R.F. 2003. *Scale Development: Theory and Applications,* 2nd ed. Thousand Oaks, CA: Sage.

East, R. 1997. *Consumer Behavior: Advances and Applications in Marketing.* Harlow, UK: Prentice Hall.

Fornell, C., and D. Larcker. 1981. "Structural Equation Models With Unobservable Variables and Measurement Error." *Journal of Marketing Research,* Vol. 18, nº 1, p. 39–50.

Goulding, C. 2000. "The Museum Environment and the Visitor Experience." *European Journal of Marketing,* Vol. 34, nº 3/4, p. 261–278.

Gronroos, C. 1982. "An Applied Service Marketing Theory." *European Journal of Marketing,* Vol. 16, nº 7, p. 30–41.

Gronroos, C. 1990. *Service Management and Marketing: Managing the Moments of Truth in Service Competition.* Lexington, MA: Lexington Books.

Haahti, A., and U. Yavas. 2005. "The Effect of Customer Satisfaction With Leisure Services on Behavioral Intentions: A Study of Visitors to SantaPark in Lapland." *Services Marketing Quarterly,* Vol. 26, nº 3, p. 1–12.

Hendon, W.S., F. Costa and R.A. Rosenberg. 1989. "The General Public and the Art Museum: Case Studies of Visitors to Several Institutions Identify Characteristics of Their Publics." *American Journal of Economics and Sociology,* Vol. 48, nº 2, p. 230–243.

Hill, D.J. 1986. "Satisfaction and Consumer Services." *Advances in Consumer Research,* Vol. 13, p. 311–314.

Hirschman, E.C., and M.B. Holbrook. 1982. "Hedonic Consumption: Emerging Concepts, Methods and Propositions." *Journal of Marketing,* Vol. 46, nº 3, p. 92–101.

Hood, M.G. 1983. "Staying Away: Why People Choose Not to Visit Museums." *Museum News,* Vol. 61, nº 4, p. 50–57.

Kelly, L. 2004. "Evaluation, Research and Communities of Practice: Program Evaluation in Museums." *Archival Science,* Vol. 4, nº 1/2, p. 45–69.

Lages, L.F., and J.C. Fernandes. 2005. "The SERPVAL Scale: A Multi-item Instrument for Measuring Service Personal Values." *Journal of Business Research,* Vol. 58, nº 11, p. 1562–1572.

Lehtinen, J., and U. Lehtinen. 1982. *Service Quality: A Study of Quality Dimensions.* Unpublished paper. Helsinki: Service Management Institute.

Merriman, N. 1995. "Hidden History: The Peopling of London Project." *Museum International,* Vol. 47, nº 3, p. 12–16.

Parasuraman, A., and V.A. Zeithaml, 2002. "Understanding and Improving Service Quality: A literature Review and Research Agenda." In *Handbook of Marketing,* R. Wensley and B. Weitz, eds. (p. 339–366). Thousand Oaks, CA: Sage.

Parasuraman, A., V.A. Zeithaml and L.L. Berry. 1985. "A Conceptual Model of Service Quality and Its Implications for Future Research." *Journal of Marketing,* Vol. 49, p. 41–50.

Passebois, J. 2002. *Building Consumer/Arts Institution Relationships: The Case of Contemporary Art Museums.* Montpellier: CREGO, Université Montpellier 11.

Pranter, C., and C. Martin. 1991. "Compatibility Management: Roles in Service Performance." *Journal of Services Marketing,* Vol. 5, nº 2, p. 43–53.

Price, L.L., E.J. Arnould and P. Tierney. 1995. "Going to Extremes: Managing Service Encounters and Assessing Provider Performance." *Journal of Marketing,* Vol. 59, nº 2, p. 83–97.

Rossiter, J.R., and L. Percy. 1997. *Advertising Communication and Promotion Management.* New York: McGraw-Hill.

Rossiter, J.R., L. Percy and R.J. Donovan. 1991. "A Better Advertising Planning Grid." *Journal of Advertising Research,* Vol. 5, p. 11–22.

Rust, R.T., and R.L. Oliver, eds. 1994. "Service Quality: Insights and Managerial Implications From the Frontier." In *Service Quality: New Directions in Theory and Practice,* R.T Rust and R.L. Oliver, eds. (p. 1–20). London: Sage.

Smith, M., J. Bruner and R. White. 1956. *Opinions and Personality.* New York: John Wiley.

Szymanski, D.M., and R.T. Hise. 2000. "e-Satisfaction: An Initial Examination." *Journal of Retailing,* Vol. 76, nº 3, p. 309–322.

Thompson, B., and S.A. Vidal-Brown. 2001. *Principal Components Versus Principal Axis Factoring: When Will We Ever Learn?* Paper presented at annual meeting of Southwest Educational Research Association, February 1–3, New Orleans.

Thyne, M. 2001. "The Importance of Values Research for Nonprofit Organizations: The Motivation-Based Values of Museum Visitors." *International Journal of Nonprofit and Voluntary Sector Marketing,* Vol. 6, nº 2, p. 116–130.

Tian, S., J.L. Crompton and P.A. Witt. 1996. "Integrating Constraints and Benefits to Identify Responsive Target Markets for Museum Attractions." *Journal of Travel Research,* Vol. 35, nº 2, p. 34–45.

Verhoef, P.C., G. Antonides and A.N. de Hoog. 2004. "Service Encounters as a Sequence of Events: The Importance of Peak Experiences." *Journal of Service Research,* Vol. 7, nº 1, p. 53–64.

Weil, S. 1999. "From Being About Something to Being for Somebody: The Ongoing Transformation of the American Museum." *Daedalus,* Vol. 128, p. 229–258.

Zeithaml, V.A. 1988. "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence." *Journal of Marketing,* Vol. 52, nº 3, p. 2–22.

APPENDIX 1

SCALE QUESTIONNAIRE
Employees were polite.
Employees were friendly.
Signage was easy to follow.
Museum provided detailed maps.
Key features and exhibits were clearly identified.
Museum provided learning opportunities.
I discovered a lot.
Atmosphere was cheerful.
Atmosphere was entertaining.
Atmosphere was lively.
Museum visit increased my feelings of self-worth.
After the visit I felt like a better person.
Museum visit gave me a sense of accomplishment.
I could identify with the typical visitor in the museum.
The behaviour of other visitors irritated me.
The typical museum visitor was very much like me.