ALTERNATIVE MODES OF MEASURING STORE IMAGE: AN EMPIRICAL ASSESSMENT OF STRUCTURED VERSUS UNSTRUCTURED MEASURES

Jhinuk Chowdhury
University of North Texas

James Reardon
University of Northern Colorado

Rajesh Srivastava
University of Southwestern Louisiana

Store image has been measured frequently by means of structured scales. Some researchers exhort against the use of structured scales for the measurement of this construct and recommend the use of unstructured measures instead. They argue that structured scales are inadequate for capturing the "gestalt" associated with the perception of a store image. This research attempts, for the first time, to investigate empirically the relative efficacy of the structured scales and the unstructured measures of store image. The results reveal that the two types of measures have similar properties and that the structured scales are more correlated with a set of self-reported behavioral measures. Thus practitioners should feel more comfortable utilizing structured, semantic differential scales to assess their store image.

INTRODUCTION

Beginning with the pioneering work of Boulding (1956) and Martineau (1958), academic researchers have frequently examined store image as a potentially valuable theoretical construct. Trade publications and business media characterize it as a critical determinant of successful retailing strategy (e.g., Willmes 1990; Wilson 1993). Unlike many issues that are popular predominantly or even exclusively in academic circles, store image is believed to have concrete and consequential managerial relevance, especially with regard to its impact on profitability (Mitchell 1993).

Scholars have investigated many facets of the store image construct, including its conceptualization (e.g., Kasulis and Lusch 1981; Keaveney and Hunt 1992; Darden and Babin 1994) and operationalization (e.g., Kasulis and Lusch 1981; Golden, Albaum and Zimmer 1987; Ward, Bitner and Barnes 1992). While some researchers have examined store image as a criterion variable (e.g., Doyle and Fenwick 1975) or as a dependent measure (e.g., Ohanian and Tashchian 1992; Baker, Grewal and Parasuraman 1994), others have observed its contingent (i.e., interactive) effects (e.g., Thorelli, Lim and Ye 1989; Gupta and Cooper 1992). Interest in the potential value of this topic extends beyond the United States to business researchers in Europe (e.g., Coshall 1985) and

Japan (e.g., Wakabayashi 1988; Wakabayashi and Itoh 1994). The ongoing involvement with store image has not only contributed to an advancement of knowledge, but it has also generated several academic debates concerning the theoretical underpinnings of the store image construct. Partly as a result of these theoretical conflicts, a number of complex issues have remained shrouded in ambiguity and key questions have remained unanswered. Recurrently, scholars have described this area as one characterized by a high "noise" level (e.g., Peterson and Kerin 1983; Amirani and Gates 1993).

A fundamental area of conflict and confusion involves the relationship between the conceptual underpinning of the store image construct and its operationalization. Keaveney and Hunt (1992, p. 167) cite the work of Arons (1961), Dichter (1985), Mazursky and Jacoby (1986), Oxenfeldt (1974), and Zimmer and Golden (1988) to argue the point that "operationalizing retail store image along traditional attribute-based lines [cannot] account for the gestalt view of store image." Keaveney and Hunt's contentions include the view that individuals' images of stores are composite, synergistic, and gestalt in nature. The term "gestalt" is frequently used by researchers to convey the idea that the individual's perception of any object incorporates innumerable bits of separate information that are combined in a such a manner that the end result of the integration of the inputs amounts to more than the sum of its constituent parts. The information processing system transmutes the input signals (primarily visual cues) and combines them with internally stored knowledge to generate the resultant inference.

According to Keaveney and Hunt, traditional measures of store image are inadequate, perhaps even erroneous. They regard the use of paper and pencil tests focusing on attributes of stores, the application of "attribute-intensive" semantic differential, Likert-type, or Stapel scales, the use of attribute based-models of information processing, and the incorporation of multivariate statistical analyses including multidimensional scaling, discriminant analysis, and factor analysis as inappropriate, or at least deficient, for the purpose of measuring images of stores.

Among the research instruments, tools, and methods recommended by Keaveney and Hunt are those oriented toward category-based processing theory (cf. Sujan 1985; Fiske and Pavelchak 1986). These authors support Zimmer and Golden's (1988) use of an unstructured measurement technique and their attempt to capture the 'gestalt' of store image. Zimmer and Golden use content analysis of respondents' descriptions of the image of three national retail chains. In their study, a total of 894 respondents' answers to open-ended questions was first classified into 220 themes. Subsequently, through a sequence of sorting and reclassifying, the several image themes were reduced to ten dimensions.

RESEARCH QUESTION

The literature on store image is replete with exhortations on the value of adopting open-ended measures (e.g., Berry 1969; Jain and Etgar 1976; Zimmer and Golden 1988; Keaveney and Hunt 1992). In contrast, many empirical studies have adopted structured scales for appraising store image (e.g., Marcus 1972; Hirschman, Greenberg and Robertson 1978; Kasulis and Lusch 1981; Zelnio and Gagnon 1981; Sirgy and Samli 1985). The underlying rationale for adopting unstructured scales (that consumers form composite gestalt images of perceived objects) could presumably be extended to many other areas of marketing. Given the popularity and ubiquity of structured measurement scales in marketing, the following question assumes critical importance: Can structured scales adequately capture and measure the construct of store image? The purpose of this study is to empirically assess the comparative properties (and, subsequently, the relative usefulness) of structured scales and unstructured measures for evaluating store image. This comparative analysis appears especially well founded in view of the fact that such an attempt has never been made before.

The use of structured questionnaires is widespread in both marketing research and practice. In contrast, the use of unstructured measures is rare. This is true, in part, because unstructured measures are relatively more cumbersome. Typically, unstructured questionnaires require to be refined very carefully. More often than not, such questions are context sensitive and may yield a different set of measures in different samples. From the perspective of the average retail store manager, an imperative to use unstructured scales is likely to add to cost and confusion. Such real world considerations make it critical to explore the relative efficacy of the structured and unstructured scales in the context of the measurement of store image.

In the following sections, this study describes the development and refinement of a structured scale for measuring store image. An unstructured measure for assessing consumers' store image is also reported. Subsequently, the information obtained from the two measures are compared with each other and then contrasted in the context of several self-reported behaviors.

THE STRUCTURED SCALE

The development of the structured scale was conducted in systematic conformance with the standard principles of scale construction (cf. Churchill 1979; Gerbing and Anderson 1988). The generation of the sample items was followed by purification of the measure involving tests of reliability,

1988). The generation of the sample items was followed by purification of the measure involving tests of reliability, unidimensionality, convergent validity, discriminant validity, and nomological validity. (Please refer to Appendix A for definitions of these properties of measurement scales.)

Generation of Sample Items

An exhaustive review of the current literature on store image was undertaken in order to generate a superset of store image dimensions. (See Kelly and Stephenson 1967; Kunkel and Berry 1968; Berry 1969; Lindquist 1974; McDougall and Fry 1974; Pathak, Crissy and Sweitzer 1974; Hawkins, Albaum and Best 1976; James, Durand and Dreves 1976; Hansen and Deutscher 1977; Hirschman, Greenberg and Robertson 1978; Pessemier 1980; Kasulis and Lusch 1981; Zelnio and Gagnon 1981; Malhotra 1983; Sirgy and Samli 1985; Mazursky and Jacoby 1986; Hildebrandt 1988; Dickson and MacLachlan 1990; Ohanian and Tashchian 1992; Baker, Grewal and Parasuraman 1994; Darden and Babin 1994). These store image dimensions and the scale items were initially checked for ambiguity and redundancy (i.e., recurrence) by three researchers. Subsequently, each judge individually labeled the dimensions that each scale item was purporting to measure. Next, the judges met as a group to refine the labeling further. They conferred until there was agreement on the assignment of each scale item to a dimension of store image. Those scale items that could not be unanimously assigned to a specific dimension were considered inappropriate and were discarded. The result of this exercise yielded six dimensions of store image encompassing thirty-eight scale items. These scales items, that were retained for further purification of the measures, are listed in Table 1.

TABLE 1
Structured Questionnaire - Store Image

The following items were used in the original instrument. A 7-point scale, anchored by (1) Strongly Agree and (7) Strongly Disagree, was used for the items. Items marked with an asterisk (*) were included in the final (purified) scale. Items marked with a (R) were reflexed.

Employee Service

- <Store name> employees are very friendly.
  <Store name> employees are honest with customers.
  <Store name> employees are rude. (R)
- * The service at <Store name> is excellent.
- * I am pleased with the service I receive at <Store name>.

Product Quality

- * <Store name> sells only high quality products.
  <Store name> produce is never fresh. (R)
- * I like <Store name> brand products.
- * I can count on the products I buy at <Store name> being excellent.
  The products at <Store name> are unsatisfactory. (R)*Product Selection*

  <Store name> never has what I want in stock. (R)
- * <Store name> has a large variety of products.
- * Everything I need is at <Store name>.
- * <Store name> carries many national brands.
  <Store name> always seems to lack the size of the package that I buy. (R)

TABLE 1 (CONTINUED)
Structured Questionnaire - Store Image

Atmosphere

	I feel comfortable shopping in <Store name>.
*	The appearance of <Store name> is appealing.
*	<Store name> is always dirty. (R)
*	<Store name> is old-fashioned. (R)
	<Store name> keeps the interior temperature much too hot. (R)
	<Store name> has the ugliest buildings. (R)
	The area around <Store name> is clean.
	<Store name> is located in a nice area.
	The <Store name> store is appealing.
	<Store name> is a pleasant place to shop.
	<Store name> is a nice place.

Convenience

*	<Store name> is easily accessible.
*	<Store name> is convenient.
	<Store name> is well organized.
	I can never find what I'm looking for. (R)
*	It is easy to get into the store.
	It is difficult to reach the products. (R)
	The price tags are easy to find in <Store name>.

Prices/Value

	<Store name> charges the highest prices.
*	The prices at <Store name> are fair.
*	I obtain value for my money at <Store name>.
*	I can buy products for less at <Store name>.
	<Store name> is too expensive. (R)

Self-reported Behavioral Measures (Dependent Variables)

I often shop at <Store name>.	(DV1)
I would recommend <Store name> to my friends.	(DV2)
I am satisfied with <Store name>.	(DV3)
<Store name> is a good place to shop.	(DV4)
Members of my family often shop at <Store name>.	(DV5)

Purification of Measure

A sample of 189 individuals in a major metropolitan area in the southwest United States responded to the survey in which they rated a large regional grocery store chain. The survey was given only if the respondent indicated familiarity with the store. Besides the thirty-eight structured items (measured on a seven-point scale anchored by "strongly agree" and "strongly disagree"), the survey also included the unstructured questionnaire. (The unstructured questionnaire is reported in a following section.) For approximately half the sample, the unstructured questionnaire was preceded by the presentation of the structured items. For the remaining half, the order was reversed. The respondents had an average income of $38,038, age of 33 years, 13.57 years of formal education, and 44 percent were female.

The structured measure of store image was purified by means of exploratory factor analysis, item-to-total reliability analysis, and covariance structure analysis. Reliability analyses, using coefficient alpha (Cronbach 1951), and item-to-total correlations suggested that the scale reliability of a

number of dimensions could be improved by selective elimination of items. An exploratory factor analysis revealed that several items loaded high (above .5) on more than one dimension. Items indicated for deletion by the reliability analysis and the factor analysis were discarded. A measurement model using covariance analysis with LISREL 8 (cf. Jöreskog and Sörbom 1993) was examined to purify the scale further. The modification indices and standardized residuals were used as gauges for that purpose. In each of the six dimensions of the store image scale, three items were finally retained. The surviving items are identified by asterisks in Table 1 (in previous section).

Scale Reliability and Validity

Face validity of the scale was established as the items were generated through the matching technique. The reliabilities of the six dimensions (measured by Cronbach's alpha) range from .76 to .92. Additional validity and unidimensionality of each dimension of the purified scale were tested with confirmatory factor analysis using LISREL 8. Convergent validity was tested by examining the t-values of the Lambda-X matrix (Bagozzi 1981). These values range from 8.60 to 18.60 and are well above the 2.00 cutoff level adopted by Kumar et al. (1992). Discriminant validity was tested by setting the individual paths of the Phi matrix to one and testing the resultant model against the original (Gerbing and Anderson 1988; Kumar, Stern and Achrol 1992) using the D-square statistic (Jöreskog and Sörbom 1993). (Please see Appendix B for a listing of the values of Cronbach's alpha, the t-values of the Lambda-X matrix for each of the three items within a given dimension, and the values of the D-square statistic.)

The fit of the overall measurement model indicated an acceptable level of unidimensionality of each measure (i.e., each dimension of store image). Although the chi-square goodness-of-fit value is significant, the *p*-value associated with RMSEA test of a close fit (Root Mean Square Error of Approximation < .05) is .31. This suggests that the hypothesis of a good fit cannot be rejected. An additional indication that the model fits well (and thus assures unidimensionality of the measures) is that the value of Expected Cross Validation Index (ECVI) for the measurement model (1.47) is less than the same for the saturated model (1.71) (Jöreskog and Sörbom 1993). Finally, the GFI is .91 which is above the generally recommended limit of .9 (Marsh, Balla and Roderick 1988; Bentler 1990; Lichtenstein, Ridgway and Netemeyer 1993). Overall, it may be concluded that each dimension of the store image scale is unidimensional.

Nomological Validity

To test for the nomological validity (i.e., the degree of association with related behaviors) of the store image scale, five self-reported behavioral measures were collected. These are listed in the Table 1 as DV1 through DV5. These five dependent measures were separately regressed on the all six store image dimensions to test for the predictive ability of the store image scale. The F values of the regression models are all significant at the .001 level, indicating that the store image correlates highly with these self-reported behaviors. The amount of variance explained (measured by values of R^2) ranges from .41 to .70 (see Appendix C).

UNSTRUCTURED QUESTIONNAIRE

One major objective of unstructured measurement entails preserving the respondents' own schema of the measured object and taking care not to sensitize them to specific aspects of it by offering them a set of image dimensions (Zimmer and Golden 1988). Consequently, in conformance with the research literature on store image, the unstructured questions were designed to capture the respondents' global (or "gestalt") impressions of the chain store. (Please see Table 2 for the entire questionnaire.) In the questionnaire, the first question reflects an attempt to elicit a completely unprompted response from the respondent. It was designed to preclude the possibility of directing the respondent toward specific attributes of stores or any particular affective or cognitive dimensions. The remaining questions were included to capture other unprompted reports of affective, cognitive, and conative reactions from the respondents.

The coding procedure involved the assignment of each response to one of the six dimensions used in the structured scale development. The purpose of choosing these six dimensions is twofold. First, these six dimensions have been culled from a large body of authoritative research and are make most sense in theoretical terms. Second, since the primary purpose of the paper is to compare the structured and the unstructured measures, the two types of responses must necessarily be reduced to some common denominator so that the comparative analysis may be accomplished.

For those responses that could not be meaningfully categorized in any of the six categories, two additional groups were created. These categories — "Other - cognitive" and "Other - affective" — were created to hold miscellaneous thought-related and feeling-related statements respectively. Within each of these eight groupings, the response was assigned a positive, negative, or neutral subcategory. Initially, the responses were independently coded by two trained business students and each was subsequently verified by one researcher. These category assignments were reviewed by the researchers and discrepancies were resolved through deliberations. Responses that were not unanimously categorized after discussions were rejected.

TABLE 2
UNSTRUCTURED QUESTIONNAIRE

1. What is the first thing that comes to your mind when you think of <Store name>?
2. Please list any *feelings* that come to mind when you think of <Store name> grocery stores. (Be as descriptive as possible.)
3. Please list any *thoughts* that come to mind when you think of <Store name> grocery stores. (Be as descriptive as possible.)
4. How would you describe <Store name> to a long time friend? (Please be specific.)
5. What do you like MOST about <Store name>?
6. What do you like LEAST about <Store name>?

Two types of store image measures were generated from these classifications. The first, Unstructured/Net-Positive, was created by subtracting the number of negative responses from the number of positive responses. This measure incorporates the assumption that positive and negative responses are equal in valence and opposite in direction. The second measure, Unstructured/Total-Positive (or Unstructured/Aggregate-Positive), simply reflects the total number of positive responses only.

In sum, responses to the unstructured questionnaire were coded to yield frequencies of net-positive and aggregate-positive responses corresponding to eight categories (i.e., six dimensions associated with the structured scale and two additional categories labeled "other - cognitive" and "other - affective").

SCALE COMPARISON

The scores on the structured and unstructured scales were standardized before tests were done to compare the two scales. Given the nature of the unstructured scale, traditional assumptions regarding normality of distributions are likely to be tenuous. Consequently, nonparametric methods (which involve tests without making distributional assumptions) are indicated (Iman and Conover 1983).

Test for Differences in Distribution

Within each of the six relevant dimensions, a Wilcoxon nonparametric test (see Siegel and Castellan 1988) was conducted to test the hypothesis that there are no differences between the two paired populations of ordered metric scores. The Z statistic and the corresponding p-values for the 2-tailed tests are reported in Appendix D. For each of the six dimensions the tests indicate that there are no differences between the distributions of the unstructured and the structured scales. This supports the proposition that the two measurement processes result in the same type of distribution of responses. In other words, if some customers associate a store with great service while others regard its service to merely adequate, this response pattern is likely to be independent of the measurement technique used. Furthermore, this is true regardless of whether the unstructured scales are formulated with net-positive or aggregate-positive (i.e., positive only) scores.

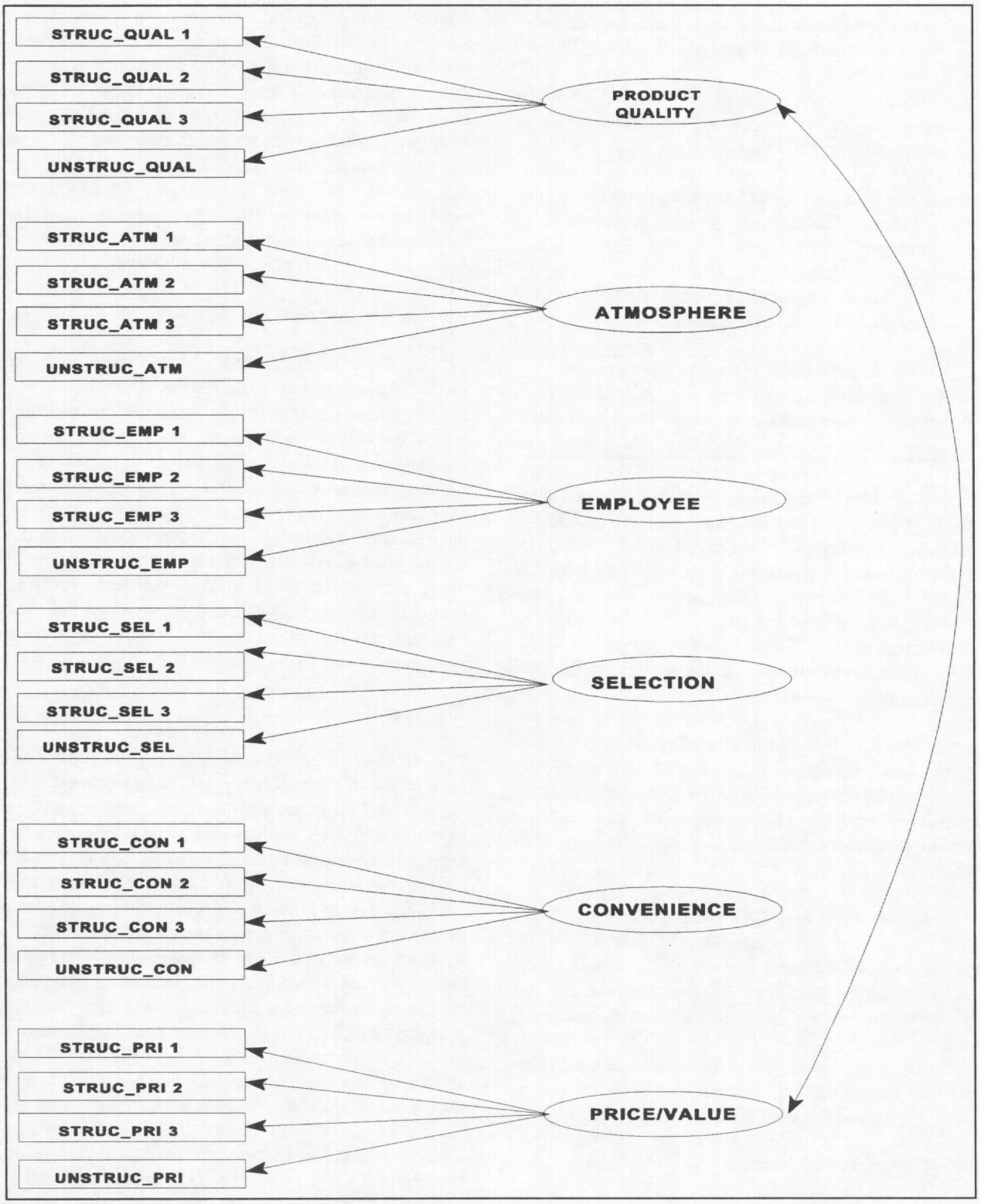
Correlations Between the Measures

The next step in the process of comparison of the scales involves an investigation of the degree of correlation between the respondents' scores on the structured and unstructured scales by within each dimension. The entire sets of correlation coefficients are listed in Appendix E. The negative signs associated with the correlation coefficients are a consequence of the reverse polarities of the structured and the unstructured scales. These numbers suggest that there is a very good association between measures of store image using a structured versus unstructured scale. Again, this lends further evidence to support the proposition that retailers can use either method without significant change in aggregated results.

Convergent Validity

The principle of reflective scaling (Bagozzi 1981) was used to test for the possibility that the structured and the unstructured scales are measuring the same construct. This principle embodies the idea that if two scales measure the same underlying construct, the scores may be expected to reflect that phenomenon. A confirmatory factor analysis model (see Figure 1) that includes both the structured and the unstructured measures was estimated to examine the degree to which the scales measure the same dimensions (Gerbing and Anderson 1988). The t-values associated with the maximum likelihood estimates of the KSI coefficients provided by LISREL are greater than 2.00 and, consequently, indicate an acceptable level of convergence of both the structured and the unstructured measures with the same underlying dimensions. The p-value associated with the test for a close fit (RMSEA < .05) is .02, indicating that the hypothesis of a close fit cannot be rejected at the .01 level of confidence. Additionally, the ECVI for the saturated model (3.57) exceeds the ECVI for the model tested (3.13) and provides further indication of a close fit. However, the chi-

FIGURE 1
STRUCTURAL EQUATIONS MODEL



square goodness-of-fit statistic is significant at the .001 level and, consequently, it does not corroborate the preceding conclusions. It is likely, though, that the differences in scaling and measurement may have contributed to poor chi-square results. Overall, this analysis suggests both the structured and the unstructured measures reflect the same underlying dimensions. Since both measurement methods are measuring the same construct, this test also lends support that retailers should feel comfortable choosing either method to measure store image.

Overall Scales

Comparing the overall (or summated) scales leads to results that are comparable with those obtained from the analyses by dimensions. The Pearson product-moment correlation between the structured and the unstructured/net-positive measures is -.69 ($p < .0001$). The correlation between the structured and the unstructured/total-positive measures is -.49 ($p < .0001$). The Wilcoxon nonparametric test for differences in the distributions between the structured and the unstructured/net-positive scores yields a Z statistic of -.44 associated with a p-value of .66, indicating that the hypothesis of the same underlying distribution cannot be rejected. The same test with the structured and the unstructured/total-positive scores yields a Z statistic of .07 associated with a p-value of .94. These tests indicate that not only do the scales perform similarly when tested on each dimension, but also when tested at the aggregate level. Again, these tests lend additional support in favor of the proposition that the two measurement techniques are equivalent for the purpose of both retailers and academics.

A canonical correlation analysis with the variables of the structured scale and those of the unstructured/net-positive scale shows that the canonical functions of the structured and the unstructured/net-positive scales are significantly correlated. Similarly, the canonical functions of the structured and the unstructured/total-positive scales were significantly correlated as well. The relevant statistics are summarized in Appendix F. These results also indicate the relative ability of each scale to measure store image.

Separate regression analyses with the five self-reported behavioral measures as dependent variables and all the scores on the eight net-positive unstructured measures are shown in Appendix G. Analogous statistics for the *structured* scale are listed in Appendix C and repeated in Appendix G for comparison. Simply stated, the structured scales outperforms the unstructured scale when predicting self reported shopping behavior. In other words, the structured store image scale was highly correlated with a propensity to shop at a specific store, recommend the store to friends, be satisfied with the ore, agree that it is a good place to shop, and have members of the family shop at the store. Thus while all other tests indicate equality between the use of structured versus unstructured scales to measure store image, this last test indicates a potential superiority of the structured scale.

CONCLUSIONS AND MANAGERIAL IMPLICATIONS

The managerial significance of the store image construct can hardly be overemphasized. While the confusion regarding the appropriate measure of store image cannot be cleared in a short time, a few tentative conclusions may be drawn. The results indicate that there exists a very high degree of correspondence between the structured store image scale and the variables derived from the coding of the unstructured measures. While both of the two types of scales predict these self-reported behaviors significantly, the structured scale appears to explain a greater amount of variance. The results of this study strengthen the case for a properly constructed structured scale to measure store image (and also a scale derived from unprompted measures).

There are many disadvantages associated with the unstructured scale, including the necessity of painstaking content analyses and coding procedures, potential verbosity or articulation bias, and unmotivated respondents (Zimmer and Golden 1988). Additional drawbacks include bias introduced through the coding procedure, the added time and costs associated with processing such an instrument, and the lack of reliability and validity indices for such measures. In the face of such a disconcerting choice, this study should contribute toward reinforcing both managerial and academic confidence in the traditional structured scales that are ubiquitously used in marketing for measuring images of stores, brands, products, salespeople, and companies among other things.

Admittedly, the dependent variables used were not measures of actual behaviors. Instead, self-reports of behaviors were collected. It can be argued that the structured scale predicts the self-reported behaviors better because variables from one structured scale are regressed upon variables from another structured scale (the dependent variables). For the unstructured measures, variables *derived* from coding and transformation of the unstructured measures are being regressed upon variables based on structured scales. As such, they are "a few steps removed" from the dependent variables. Nevertheless, in order to go beyond mere investigative explorations of consumers' perceptions of image and to put the construct to work in quantitative analyses, coding the unstructured measures of store image and transforming them into variable is imperative. Following the logic of studies espousing unstructured scales, it is reasonable to expect such derived measures to contain more "raw" information since they are purportedly less constraining than structured scales. Consequently, these measures should be expected to predict the relevant dependant variables better than measures

Consequently, these measures should be expected to predict the relevant dependant variables better than measures incorporating traditional scales. In this study, however, the unstructured measure does not appear to do so.

Epilogue

Empirical validation of conceptual and logical extensions of theory is widely considered a critical element of the scientific process. In spite of the many theoretical discourses on the relative merits of structured and unstructured formats of the store image scale, there exists no prior attempt to empirically substantiate the comparative efficacy of the two. This study is an initial effort toward that end. As a neonate in such an endeavor, this research is encumbered with a number of limitations. Had the self-reported behaviors been replaced by actual behavioral measures, this study clearly would have afforded a higher level of confidence in the conclusions. Additionally, there are several possible ways in which unstructured measures may be collected and subsequently coded. The method chosen in this study was intended to represent the salient perspectives in current literature. This research cannot and does not purport to address all variants of normative issues in the construction of unstructured measures. However, in spite of these limitations, this study addresses an important issue and provides answers to some significant questions about the store image construct.

REFERENCES

Amirani, Shahrzad and Roger Gates (1993), "An Attribute-Anchored Conjoint Approach to Measuring Store Image," *International Journal of Retail & Distribution Management*, 21 (5), 30-39.

Arons, Leon (1961), "Does Television Viewing Influence Store Image and Shopping Frequency?", *Journal of Retailing*, 37, 1-13.

Bagozzi, Richard P. (1981), "Evaluating Structural Equation Models With Unobservable Variables and Measurement Error: A Comment," *Journal of Marketing Research*, 18 (August), 375-381.

Baker, Julie, Dhruv Grewal, and A. Parasuraman (1994), "The Influence of Store Environment on Quality Inferences and Store Image," *Journal of the Academy of Marketing Science*, 22 (4), 328-339.

Bentler, Paul M. (1990), "Comparative Fit Indices in Structural Equation Models," *Psychological Bulletin*, 107, 238-246.

Berry, Leonard L. (1969), "The Components of Department Store Image: A Theoretical and Empirical Analysis," *Journal of Retailing*, 45 (1), 3-20.

Boulding, Kenneth E. (1956), *The Image*. Ann Arbor, MI: University of Michigan Press.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (February), 64-73.

Coshall, J. (1985), "The Form of Micro-Spatial Consumer Cognition and Its Congruence With Search Behavior," *Tijdschrift voor economische en sociale geographie*, 70, 345-355.

Cronbach, Lee J. (1951), "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, 16, 297-334.

Darden, William R. and Barry J. Babin (1994), "Exploring the Concept of Affective Quality: Expanding the Concept of Retail Personality," *Journal of Business Research*, 29 (2), 101-109.

Dichter, Ernest (1985), "Commentary: What's an Image," *Journal of Consumer Marketing*, 2 (1), 75-81.

Dickson, John P. and Douglas L. MacLachlan (1990), "Social Distance and Shopping Behavior," *Journal of the Academy of Marketing Science*, 18 (2), 153-161.

Doyle, Peter and Ian Fenwick (1975), "How Store Image Affects Shopping Habits in Grocery Chains," *Journal of Retailing*, 50 (4), 39-52.

Fiske, Susan T. and Mark A. Pavelchak (1986), "Category-Based Versus Piecemeal-Based Affective Responses: Developments in Schema-Triggered Affect," in *The Handbook of Motivation and Cognition: Foundations of Social Behavior*, Richard M. Sorrentino and E. Tory Higgins, eds. New York: Guilford Press, 167-203.

Gerbing, David W. and James C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (May), 186-192.

Golden, Linda L., Gerald Albaum, and Mary Zimmer (1987), "The Numerical Comparative Scale: An Economical Format for Retail Image Measurement," *Journal of Retailing*, 63 (4), 393-410.

Gupta, Sunil and Lee G. Cooper (1992), "The Discounting of Discounts and Promotion Thresholds," *Journal of Consumer Research*, 19 (3), 401-411.

Hansen, Robert A. and Terry Deutscher (1977-1978), "An Empirical Investigation of Attribute Importance in Retail Store Selection," *Journal of Retailing*, 53 (4), 59-72,95.

Hawkins, Del I., Gerald Albaum, and Roger Best (1976), "Reliability of Retail Store Images As Measured By the Stapel Scale," *Journal of Retailing*, 52 (4), 31-38.

Hildebrandt, Lutz (1988), "Store Image and the Prediction of Performance in Retailing," *Journal of Business Research*, 17 (1), 91-100.

Hirschman, Elizabeth C., Barnett Greenberg, and Dan H. Robertson (1978), "The Intermarket Reliability of Retail Image Research: An Empirical Examination," *Journal of Retailing*, 54 (1), 3-12.

Iman, Ronald L. and W. J. Conover (1983), *A Modern Approach to Statistics*, Wiley series in probability and mathematical statistics. New York: Wiley.

Jain, Arun K. and Michael Etgar (1976-1977), "Measuring Store Image Through Multidimensional Scaling of Free Response Data," *Journal of Retailing*, 52 (4), 61-70,95-96.

James, Don L., Richard M. Durand, and Robert A. Dreves (1976), "The Use of a Multi-Attribute Attitude Model in a Store Image Study," *Journal of Retailing*, 52 (2), 23-32.

Joreskog, Karl G. and Dag S`rbom (1993), *LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language*. Chicago, IL: Scientific Software International, Inc.

Judd, Charles, Eliot Smiiith, and Louise Kidder (1991), "Research Methods in Social Relations,"6th edition, Holt, Rinehart and Winston, Inc. Chicago.

Kasulis, Jack J. and Robert F. Lusch (1981), "Validating the Retail Store Image Concept," *Journal of the Academy of Marketing Science*, 9 (4), 419-435.

Keaveney, Susan M. and Kenneth A. Hunt (1992), "Conceptualization and Operationalization of Retail Store Image: A Case of Rival Middle-Level Theories," *Journal of the Academy of Marketing Science*, 20 (2), 165-175.

Kelly, Robert F. and Ronald Stephenson (1967), "The Semantic Differential: An Information Source for Designing Retail Patronage Appeals," *Journal of Marketing*, 31, 43-47.

Kerlinger, Fred N. (1986), "Foundations of Behavioral Research," 3rd edition, Harcourt Brace Jovanocvich College Publishers, 1986, New York.

Kumar, Nirmalya, Louis Stern, and Ravi Achrol (1992), "Assessing Reseller Performance from the Perspective of the Supplier," *Journal of Marketing Research*, 29 (May), 238-253.

Kunkel, John H. and Leonard L. Berry (1968), "A Behavioral Conception of Retail Image," *Journal of Marketing*, 32, 21-27.

Lichtenstein, Donald R., Nancy M. Ridgway, and Richard G. Netemeyer (1993), "Price Perceptions and Consumer Shopping Behavior: A Field Study," *Journal of Marketing Research*, 30 (2, May), 234-245.

Lindquist, Jay D. (1974-1975), "Meaning of Image," *Journal of Retailing*, 50 (4), 29-38.

Malhotra, Naresh K. (1983), "A Threshold Model of Store Choice," *Journal of Retailing*, 59 (2), 3-21.

Marcus, Burton H. (1972), "Image Variation and the Multi-Unit Retail Establishment," *Journal of Retailing*, 48 (2), 29-43.

Marsh, Herbert W., John R. Balla, and P. McDonald Roderick (1988), "Goodness-Of-Fit Indices in Confirmatory Analysis: The Effect of Sample Size," *Psychological Bulletin*, 103, 391-410.

Martineau, Pierre D. (1958), "The Personality of the Retail Store," *Harvard Business Review*, 36 (January-February), 47-55.

Mazursky, David and Jacob Jacoby (1986), "Exploring the Development of Store Images," *Journal of Retailing*, 62 (2), 145-165.

McDougall, G. H. G. and J. N. Fry (1974-1975), "Combining Two Methods of Image Measurement," *Journal of Retailing*, 50 (4), 53-61.

Mitchell, Jim (1993), "Perception Is Everything," *NPN: National Petroleum News*, 85 (5), 47.

Ohanian, Roobina and Armen Tashchian (1992), "Consumers' Shopping Effort and Evaluation of Store Image Attributes: The Roles of Purchasing Involvement and Recreational Shopping Interest," *Journal of Applied Business Research*, 8 (4), 40-49.

Oxenfeldt, Alfred R. (1974-1975), "Developing a Favorable Price-Quality Image," *Journal of Retailing*, 50 (4), 8-14,115.

Pathak, Dev S., William J. E. Crissy, and Robert W. Sweitzer (1974-1975), "Customer Image Versus the Retailer's Anticipated Image," *Journal of Retailing*, 50 (4), 21-28,116.

Pessemier, Edgar A. (1980), "Store Image and Positioning," *Journal of Retailing*, 56 (1), 94-106.

Peterson, Robert A. and Roger A. Kerin (1983), "Store Image Measurement in Patronage Research: Fact and Artifact," in *Patronage Behavior and Retail Management*, William R. Darden and Robert F. Lusch, eds. New York: North-Holland, 293-306.

Siegel, Sidney and N. John Castellan Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. New York: McGraw-Hill.

Sirgy, M. Joseph and A. Coskun Samli (1985), "A Path Analytic Model of Store Loyalty Involving Self-Concept, Store Image, Geographic Loyalty, and Socioeconomic Status," *Journal of the Academy of Marketing Science*, 13 (3), 265-291.

Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," *Journal of Consumer Research*, 12, 31-46.

Thorelli, Hans B., Jeen-Su Lim, and Jongsuk Ye (1989), "Relative Importance of Country of Origin, Warranty and Retail Store Image on Product Evaluations," *International Marketing Review*, 6 (1), 35-46.

Wakabayashi, Yoshiki (1988), "Multidimensional Analysis of Consumers' Store Image," *Geographical reports of Tokyo Metropolitan University*, 23, 1-11.

Wakabayashi, Yoshiki and Satoru Itoh (1994), "Reexamination of the Components of Distortions in Cognitive Maps - A Case Study in Kanazawa City," *Chigaku zasshi*, 103 (3), 221.

Ward, James C., Mary J. Bitner, and John Barnes (1992), "Measuring the Prototypicality and Meaning of Retail Environments," *Journal of Retailing*, 68 (2), 194-220.

Willmes, Francis E. (1990), "Continuity Programs Can Boost Store Image," *Supermarket Business*, 45 (3), 89,92.

Wilson, Marianne (1993), "J.C. Penney's New Look," *Chain Store Age Executive*, 69 (6), 70-71.

Zelnio, Robert N. and Jean P. Gagnon (1981), "The Construction and Testing of an Image Questionnaire," *Journal of the Academy of Marketing Science*, 9 (3), 288-299.

Zimmer, Mary R. and Linda L. Golden (1988), "Impressions of Retail Stores: A Content Analysis of Consumer Images," *Journal of Retailing*, 64 (3), 265-293.

APPENDIX A
DEFINITION OF KEY TERMS

Validity	The extent to which a measure reflects only the desired construct without contamination from other systematically varying constructs. The validity measure provides the best available approximation to the truth of propositions, including propositions about cause.

Convenience	.84	13.88, 18.44, 11.23	66.37
Prices/Value	.88	11.27, 13.21, 13.40	43.64

[a] For all values of D-square, $p < .0001$

APPENDIX C
REGRESSION ANALYSES WITH BEHAVIORAL MEASURES AS DEPENDENT VARIABLES AND THE STORE IMAGE DIMENSIONS (STRUCTURED SCALE) AS INDEPENDENT VARIABLES

Dependent Variable in Model	Value of R-square	F-Value[a]
DV1 (I often shop at <Store name>)	.53	30.42
DV2 (I would recommend <Store name> to my friends)	.65	50.46
DV3 (I am satisfied with <Store name>)	.70	62.56
DV4 (<Store name> is a good place to shop)	.67	53.55
DV5 (Members of my family often shop at <Store name>)	.41	18.61

[a] All F-values are associated with d.f. (6, 121) and are significant at .0001.

APPENDIX D
RESULTS OF WILCOXON NONPARAMETRIC TESTS FOR DIFFERENCES BETWEEN STRUCTURED AND UNSTRUCTURED SCALES - BY DIMENSION

	Structured vs. Unstructured/Net-positive		Structured vs. Unstructured/Total-Positive	
Dimension	*Z* Statistic	*p*-value	*Z* Statistic	*p*-value
Employee Service	-.23	.82	-1.00	.32
Product Quality	-.24	.81	-1.08	.28
Product Selection	-.36	.72	-.18	.86

APPENDIX G

REGRESSION ANALYSES WITH THE FIVE SELF-REPORTED BEHAVIORAL MEASURES AS SEPARATE DEPENDENT VARIABLES

		Independent Variables					
		Unstructured/ Net-positive		Unstructured/ Total-positive		Structured[a]	
Dependent Variables		Overall F[b]	R-square	Overall F[b]	R-square	Overall F[b]	R-square
DV1	(I often shop at <Store name>)	11.68	.34	9.53	.28	30.42	.53
DV2	(I would recommend <Store name> to my friends)	16.65	.42	11.72	.33	50.46	.65
DV3	(I am satisfied with <Store name>)	18.65	.45	14.10	.37	62.56	.70
DV4	(<Store name> is a good place to shop)	16.84	.42	11.61	.30	53.55	.67
DV5	(Members of my family often shop at <Store name>)	7.89	.26	5.87	.20	18.61	.41

[a] The statistics associated with the structured scale are repeated from Appendix C and are included for comparison.

[b] All F values are significant at the .0001 level.

AUTHOR BIOGRAPHY

Jhinuk Chowdhury (Ph.D., University of Florida) is an associate professor of marketing at the University of North Texas. Dr. Chowdhury's research interests include strategic marketing and behavioral issues in channels of distribution. His articles have appeared in the *Journal of Marketing Research*, *Journal of Marketing Management*, *Journal of Managerial Issues*, *Journal of Current Issues & Research in Advertising*, *Journal of Services Marketing*, *Journal of International Consumer Marketing*, *Health Marketing Quarterly*, and the *Proceedings of AMA Educators' Conferences.*

AUTHOR BIOGRAPHY

James Reardon (Ph.D., University of North Texas) is an assistant professor at the University of Northern Colorado. His research interests include retailing, measurement issues, and international channel relationships. His research has appeared in the *Journal of Retailing*, *Journal of Marketing Theory and Practice*, *Journal of Marketing Management*, *Journal of International Advertising*, *Journal of Applied Business Research*, *International Journal of Management*, and the *Journal of Business and Economic Perspectives*. He has also recently coauthored a textbook with Ronald Hasty, titled *Retail Management*.

AUTHOR BIOGRAPHY

Rajesh Srivastava (Ph.D., University of North Texas) is an assistant professor at the University of Southwestern Louisiana. His research interests include behavioral issues in salesperson work environments and retailing. His research has been published in *International Journal of Management* and the proceedings of *AMA Educators' Conference*, *Academy of Marketing Science Conference*, and *National Conference for Sales Management*.

Convergent Validity	Refers to the overlap between alternative measures that are intended to tap the same construct but that have different sources of irrelevant, undesired variation. Convergent validity provides support for the notion that evidence from different sources gathered in different ways all indicates the same or similar meaning of the construct
Discriminant Validity	Refers to the degree to which a measure uniquely captures the construct it is designed to determine. It is the extent to which a measure fails to correlate with measures that are supposed to tap different constructs. Discriminant validity helps to empirically differentiate the construct from other constructs that may be similar, and identify that which is unrelated to the construct.
Nomological Validity	Refers to the theoretically derived set of relationships with other constructs that serves to define the target construct.
Reliability	Reliability is defined as the extent to which a measure is free from random error components. It is the accuracy or precision of a measuring instrument. Synonyms for reliability are: dependability, stability, consistency, predictability, and accuracy.
Unidimensionality	If a series of variables all measure a single general characteristic of a construct, the variables should all be highly interrelated. The construct is then said to be unidimensional. In contrast, multidimensionality refers to the fact that the variables measure more than an unique general characteristic.

SOURCES:

Kerlinger, Fred N. (1986), "Foundations of Behavioral Research," 3rd edition, Harcourt Brace Jovanovich College Publishers, 1986, New York.

Judd, Charles, Eliot Smith, and Louise Kidder (1991), "Research Methods in Social Relations," 6th edition, Holt, Rinehart and Winston, Inc. Chicago.

APPENDIX B
RELIABILITY AND VALIDITY INDICES FOR THE DIMENSIONS OF THE STORE IMAGE SCALE

Dimension	Reliability (Cronbach's alpha)	Convergent Validity (Lambda-X t-values)	Discriminant Validity (D-square[a])
Employee service	.92	12.49, 17.04, 18.60	121.82
Product quality	.76	13.39, 8.60, 12.48	128.79
Product selection	.84	12.48, 13.86, 12.21	173.30
Atmosphere	.90	15.90, 14.07, 11.67	119.04

Atmosphere	-.68	.50	-.42	.68
Convenience	-.16	.87	-.28	.78
Prices/Value	-.41	.68	-.69	.49

APPENDIX E
CORRELATIONS BETWEEN SCORES ON THE STRUCTURED AND UNSTRUCTURED SCALES - BY DIMENSION

Dimensions	Structured vs. Unstructured/Net-Positive	Structured vs. Unstructured/Total-Positive
Employee Service	-.73	-.48
Product Quality	-.40	-.24
Product Selection	-.62	-.39
Atmosphere	-.65	-.37
Convenience	-.40	-.32
Prices/Value	-.65	-..41

Note: All correlation coefficients are significant at the .001 level.

APPENDIX F
STATISTICS FROM THE CANONICAL CORRELATION ANALYSES BETWEEN THE STRUCTURED AND THE UNSTRUCTURED SCALES

Statistics	Structured vs. Unstructured/Net-positive	Structured vs. Unstructured/Total-positive
Wilks' lambda	.13[a]	.34[a]
Pillai's trace	1.55[a]	.91[a]
Hotelling-Lawley trace	2.79[a]	1.25[a]
Shared variance	.77	.76

[a] Significant at the .0001 level.