Applied Scale Development: Measurement of Store Image

Dr. James Reardon, Business Administration, University of Tennessee, Knoxville
Dr. Chip E. Miller, Business Administration, Pacific Lutheran University
Dr. Barbara Coe, Marketing, University of North Texas

Abstract

The paper discusses the need for the development of measures that can be used in an applied domain. It suggests that applied scales differ from scales used for scientific research in several ways. The paper presents an applied scale developed for use by retail managers to assess store image. Unlike previous research, the priorities used in the development of the scale included the applicability of the scale by store managers.

Introduction

The validity of measures for research in the business disciplines has recently received a great deal of attention (Hemmasi et al. 1994; Martin and Eroglu 1993). Peter (1979, p.6) goes as far as to argue that "Valid measurement is the *sine qua non* of science." However, academics have largely ignored the very same issues as they pertain to the business practitioner. Academics can make a potential contribution to the applied practice of business by examining, among other issues, the conceptual differences between scale development for the purposes of science and application. In pursuit of this goal, an applied store image scale is presented in this paper to demonstrate the differences between applied and scientific scales.

Theoretical Framework

Questions relating to shortcomings associated with validity and unidimensionality of research measures have been the focus of academic researchers for the past twenty-five years (Churchill, 1979; Heeler and Ray, 1972; Jacoby, 1978; Kassarjian, 1971; Rogers, 1976). Peter (1979, 1981), Churchill (1979) and Gerbing and Anderson (1988) responded to these shortcomings by laying the groundwork for scale development in business disciplines.

Peter (1981) defines construct validity as "the degree of correspondence between constructs and their measures" and suggests that construct validity is a composite of both reliability and convergent, discriminant, face, and nomological validities. The reliability of a measure refers to the degree to which the measure is free from error and therefore yields consistent results (Peter 1979).

As suggested by Peter (1979), a scale should measure the magnitude and direction of all the characteristics of a construct and only those characteristics. In addition, Gerbing and Anderson (1988) discuss the dimensionality of a construct and its measure in much the same context. They suggest that a construct should be unidimensional in order to ensure that the measure is not polluted with partial measures of other constructs. While these authors have begun to address issues of scientific measurement, virtually no scholars discuss the need and development of applied scales.

Applied Scale Development

Research in the applied domain is different from scientific research (Calder et al. 1981, 1982, 1983). Applied research is more concerned with correspondence and not necessarily theoretical measures. In addition to meeting the traditional considerations of construct validity, applied scales must also be actionable, cost efficient, and generalizable.

Managers must be able to interpret measures, use them for diagnosis, and take corrective action as inadequacies are indicated by the measures. For the measure to be functional, the manager needs to be able to take action to improve the current situation. It does little good for a manufacturer to find that the customer is unsatisfied due to uncontrollable environmental factors such as the general state of the economy. Instead, the manager needs to focus on actionable and controllable variables.

To maximize efficiency, managers face several constraints in conducting research. First, and perhaps most important, are financial constraints. Managers must spend their research dollars wisely. The second constraint arises from the lack of interest of respondents. A manager must minimize both the cost of collecting information and the time and effort costs to the consumer of providing the information.

Measures should be generalizable across many situations. Generalizability is important because constructing a new measure for each specific situation is both wasteful and expensive (Cronbach 1971; Nunnally 1978). Ideally, a scale should meet conditions of external validity. Lynch (1982, 1983) argues that external validity is a necessary precursor to construct validity. As academic researchers develop and build scales for applied application, limited academic resources will demand that scales be useful across a broad range of situations.

Bridging the Gap

Many academics have lamented the gap between the academic community and business practitioners (e.g., Hirschman, 1986). The results of the 1986 AMA study found there were many instances of failure to communicate between academics and practitioners. For example, many managers think that the material published in academic journals is of no practical value. According to some practitioners, it is written by academicians for other academicians. The same comments can be made with regard to published research relating to scale development.

Scale development and validation often requires the use of the Multi-Trait Multi-Method (MTMM) approach (Campbell and Fiske 1959) or Structural Equation Modeling (Gerbing and Anderson 1988). Neither of these methods is popular with practitioners. The burden for constructing valid scales therefore falls upon the shoulders of those with both the resources and ability, namely academics. While the development of scales requires advanced techniques, the scales themselves are easily applied and analyzed with descriptive statistics.

Development of an Applied Image Scale

A review of the current literature indicates that one area which offers high potential for the development of an applied scale is store image. Store image is an important area for emphasizing expanded application given the increasing relevance of store image as a competitive tool for retailers and the current controversy surrounding the development of store image scales. From a theoretical standpoint, store image has received a great deal of attention including being the focus of a special issue of the

Journal of Retailing (Winter 1974-75). Despite the extensive research in this area, neither the construct or a universal scale has yet been developed. Nor has anyone attempted to create a scale solely for use by retailers. The development of an applied store image scale is the focus of this article.

Store Image as a Strategic Tool

Store image has emerged as a major strategic tool in the highly competitive retailing environment of the 1990s for a number of reasons. Steenkamp and Wedel (1991) argue that it is one of the retailers' most valuable assets. Image is one basis used by consumers to determine how suitable they are as customers for a particular store (Dickson and MacLachlan 1990). Image affects shopping behavior and the ultimate choice of retail stores to be patronized (Donovan and Rossiter, 1982). Increased competitive pressures are challenging retailers to determine current image, make necessary changes, and tailor a marketing strategy to attract and develop loyalty among the targeted customer group.

Despite the critical role of store image, store managers are often frustrated by their inability to alter individual store image to a significant degree. Yet this ability to identify and, if necessary, modify store image is a critical managerial skill in the retail environment of the 1990s, where targeted store image has become a key competitive tool.

Store Image Measurement

While store image has traditionally been measured using bi-polar scales (Doyle and Fenwick, 1975), it has been suggested in the literature that content analysis and other open-ended or unstructured elicitation techniques should be used to measure store image (Zimmer and Golden 1988; Keaveney and Hunt, 1992). These methods offer advantages in the measurement of store image including: 1) the researcher does not impose structure on the respondent through language or dimensions; 2) each respondent is able to describe dimensions of store image in a unique fashion, thereby reducing errors of omission on the part of the researcher; and 3) the researcher captures a more robust picture of the image of a store.

Open-ended elicitation techniques for measuring store image exhibit several disadvantages including: 1) results may be difficult to interpret because subjective assessments by the respondent depend heavily on their frame of reference, making generalization very difficult; 2) coding by the researcher often introduces bias into the results; 3) the nature of the data makes powerful statistical analysis difficult, requiring use of special analytical techniques or

nonparametric statistics; 4) variations in the capacity of respondents to articulate their thoughts and feelings affect the quality of the data; 5) the techniques allow a more gestalt measurement but often give no measurement of the degree to which an attribute or feeling is present; and 6) the results obtained from this type of elicitation method may not be actionable by the retailer, perhaps the greatest fault of open-ended elicitation techniques. While it is arguable that the construct of store image should be measured using open-ended elicitation techniques for the purposes of scientific research, this is hardly practical for applied use. The development of this store image scale focuses on application.

Methodology

Scale Criteria

Four criteria were used in the development of the scale: validity, actionability, efficiency, and generalizability. Each of these was considered in relation to the practical utility of the scale to retail managers.

Validity

In development of the current scale, construct validity was a primary concern and included the assessment of convergent, discriminant and face validity. Applied scales should meet strict validity requirements. Hirschman et al. (1978) indicated that "little has been done to evaluate the reliability and validity of store image components..." They continue, "it would appear that a minimum of scientific rigor is to be found in many studies of store image." One purpose of the research was to address these issues.

Actionability

A positive image is very important to a retailer's survival. However, measuring store image without being able to identify specifically what is being done well and what poorly is inefficient. The retailer must be able to interpret the survey results and take corrective action to improve the store's image if deficiencies are found. The purposed scale concentrates on specific aspects of store image that are under the direct control of the store manager. These aspects include merchandise, atmosphere, value, service levels, the target market (clientele) of the store, and customer convenience. These elements were chosen after extensive review of works such as May (1981). She noted that "stores possess less potential for change (than products) because neither their physical properties nor surrounding trade environment are easily changed." Hence, inclusion of unalterable items that measure the holistic view of store

image will only add to the theoretical construct, not bring meaningful information to the manager.

Efficiency

Retailers operate under strict budgets, especially for research. For a scale to be useful for the retailer, it must be cost effective. The researcher must be able to measure the store image construct, customer demographics, and other information of concern in a single, concise survey. Open-ended elicitation techniques obviously do not lend themselves to these restrictions. Based on discussions with retailers, it was decided the store image construct should consist of no more than a total of 15 to 20 questions. A survey of this length also appeals to respondents as they are less likely to suffer fatigue and more likely to maintain interest. Also, a parsimonious scale allows the retailer to ask other questions, such as demographics and attitudes towards specific operating procedures. It can be added to any customer survey or be administered at the point of purchase in a short time. Finally, the results of this type of survey allow for simple and meaningful interpretation of the results. Therefore, the retail manager need not apply advanced or complex statistical methods in order to measure deficiencies in the store's image.

Generalizability

The last priority used to develop the scale was that of external validity. In order to be widely accepted, a store image scale must be applicable for several different types of retailers in differing geographic locations. As indicated by Hirschman et al. (1978), store image issues need to be tested "among stores, among markets and within the measurement instruments themselves." Further, they state that "Few (store image studies) have been cross validated internally or replicated in a different setting or time" (p. 3). Most store image scales concentrate on specific types of retailers, such as department stores, in a single geographic area. However, the goal in the current research was to build a scale that could be used by retailers in diverse environments, not merely a single type of store.

Scale Development

Utilizing the four criteria, the store image scale was constructed and tested in several steps in accordance with Churchill's (1979) and Gerbing and Anderson's (1988) recommendations. Following Churchill (1979), the basic steps included 1) specifying the domain of the construct; 2) generation of sample items; 3) measurement purification; and 4) assessing reliability and validity. Techniques discussed by Gerbing and Anderson (1988) were included in the last step, assessing reliability and validity.

Specifying the Domain

Despite the importance of being able to determine store image, the theoretical construct of store image has not yet been firmly delineated. Some researchers contend that store image is a global impression (Zimmer and Golden 1988; Dichter 1985; Oxenfeldt 1974). Others address store image as a multi-attribute based construct (Menezes and Elbert 1979; Lutz and Bettman 1977; Hirschman et al. 1978; James et al. 1976). In either case, there is agreement that store image is essentially the way in which the retailer is viewed by its customers.

The definition of store image as a global impression that includes both judgments and affect has both advantages and disadvantages. The advantages pertain mainly to the usefulness of this definition for theoretical work by marketing scholars. However, this definition makes it difficult for both retailers and researchers to apply the construct in practice. As indicated by James et al. (1976), such studies present difficulties for retailers, however, because it is difficult to improve their marketing program based on the information provided by the data collected and because of the analysis requires unfamiliar skills. One of the purposes of this research was to develop a retail image scale that can be applied by retail managers at the store level.

A review of current store image scales indicated several possible dimensions (Berry, 1969; Lindquist, 1975; Zimmer and Golden, 1988). The various dimensions of previous store image can be divided into functional and symbolic (Sirgy and Samli 1985; Darden and Babbin 1994). Symbolic dimensions that were general or overall impressions were dropped as inactionable because of their intangible nature. For example, honesty could easily refer to pricing policies, service, or any management policies. Also, fixed assets, such as location, were deemed to be non-actionable because retail managers are unable to change or act on these in the short run (May 1981). Lastly, those dimensions that were deemed to apply to only specific stores were omitted in the interest of external validity. Six constructs remained for further analysis --merchandise/product, atmosphere, value, service level, clientele and convenience.

Generation of Sample Items

Sample items and dimensions from previously developed scales (Kelly and Stephenson, 1967; Kunkel and Berry, 1968; Berry, 1969; Lindquist, 1975; McDougall and Fry, 1975; Pathak et al., 1975; Hawkins et al., 1976; James et al, 1976; Hirschman et al., 1978; Hansen and Deutscher, 1978; Pessemier, 1980; Zelnio and Gagnon, 1981; Kasulis and Lusch, 1981; Malhotra, 1983; Mazursky and Jacoby, 1986;

Sirgy and Samli, 1985; Hildebrandt, 1988; Dickson and MacLachlan, 1990; Ohanian and Tashchian, 1992; Baker et al., 1994; Darden and Babbin, 1994) were examined by two researchers and two graduate teaching fellows. Items and dimensions were initially screened for duplicate items and ambiguity. Next, each judge labeled dimensions that each scale item was measuring. The group then met and further refined the labeling until a consensus was reached on which scale items were appropriate for which dimensions. Scale items on which there was not unanimous agreement were eliminated. Four items per construct remained before purification.

Measure Purification

A sample of 179 people in the south-central United States rated a large, regional department store chain on a seven-point scale marked strongly agree to strongly disagree for twenty-four items ranging from convenience to store personnel to parking ease. Respondents were first asked if they were familiar with the store and then given the questionnaire if they responded in the affirmative.

Three methods were used to purify the measures: exploratory factor analysis, item-to-total reliability, and covariance structure analysis. Reliability analysis, using Cronbach's alpha and item-to-total correlations, showed that the reliability could be improved on several dimensions by eliminating items. The items were tentatively dropped, pending further analysis. The exploratory factor analysis showed several items that loaded high (above .5) on more than one factor. Items dropped from further analysis were consistent with the items identified for removal by the reliability analysis. A measurement model using covariance analysis (Lisrel 8), was run on the remaining items. The modification indices and standardized residuals were used to purify the scale and achieve unidimensionality of constructs (Gerbing and Anderson 1988).

Assessing Reliability and Validity

Data was collected, using the purified scale, from random samples of 156 respondents in the south-central United States and 130 respondents in the Pacific Northwest. The wide geographic distribution was used to reduce the potential for geographic bias (Hirschman et al. 1978) and to increase external validity. To test generalizability, a grocery store (south-central) and a department store (Pacific Northwest) were examined. These two store types were chosen because they represent the majority of retail sales in the United States. Also, they represent different types of product classes. Grocery stores supply essential items that are consumable, whereas department stores tend to offer more discretionary items and durable goods. Moreover,

Table 1
Scale Purification

DIMENSION =>>	MERCHANDISE	ATMOSPHERE	VALUE	SERVICE	CLIENTELE	CONVENIENT
Final Items	High quality products	Pleasant atmosphere	Good value for the money	Fast Checkout	Where my friends shop	Good parking availability
Seven point scale - Strongly Agree to Strongly Disagree	Never out of stock	Comfortable to shop there	Affordable	Many extra services	Where my peers shop	Not crowded
Highest item-to-total Alpha coefficient	0.8545	.9339	0.8884	0.8764	0.8458	0.6608
Final Reliability	0.7400	0.9120	0.8278	0.60685	0.8727	0.6198

Confirmatory Factor Analysis	Original Model	Purified Model
Chi-sqr	512.78 p=0.00	49.16 p=0.13
RMSEA	0.11 p<.0001	0.038 p=0.71
ECVI (Model)	3.44	0.71
ECVI (Saturated)	2.36	0.88
RMR	0.100	0.037
GFI	0.77	0.96 adj=0.91

virtually everyone has had experience with each type of store, and has formed an image related to these stores.

The results of the store image scale were subjected to reliability and validity tests. Because traditional reliability measures are negatively correlated with the number of items (Parameswaran et al. 1979), reliability was calculated using rho (r), as recommended by Joreskog (1971). The validity and unidimensionality (Gerbing and Anderson 1988) of the scale was tested with confirmatory factor analysis using LISREL 8 (Joreskog and Sorbom 1993). Convergent validity was tested by examining the t-values of the Lambda-X matrix (Bagozzi 1981). Discriminant validity was tested by setting the individual paths of the Phi matrix to 1 and testing the resultant model against the original (Gerbing and Anderson 1988; Nirmalya et al. 1992) using the D statistic (Joreskog and Sorbom 1993). Face validity was established earlier when generating the scale items through a matching technique.

Results

Scale Purification

The results for the purified dimensions are shown below in Table 1. The highest item-to-total Cronbach's Alpha coefficients ranged from .9339 to .6608 for the original

indicators. However, these dropped dramatically when the number of indicators per construct was reduced to two. Joreskog's measure of reliability (1971) was used for final reliability measurement because of its relative insensitivity to the number of indicators. Most of the original dimensions showed high reliability. However, they also exhibited a lack of unidimensionality. Reliability dropped slightly, with the exception of the CLIENTELE dimension, from the deletion of indicators. On the other hand, the dimensionality was greatly improved. This tradeoff was necessary to achieve an accurate measure of the individual dimensions. The Chi-squared on the confirmatory factor analysis dropped from 512 to 49. Changes in the other goodness-of-fit indicators are given in Table 2. All indicate a poor fit for the original model and an acceptable fit for the purified model. This indicates that unidimensionality was achieved in the second model.

Reliability and Validity

The results of the covariance structure tests of validity for the second and third samples are shown in Tables 2 and 3. Reliability was again calculated using the formula provided by Joreskog (1971). The table indicates relatively high reliability for all but the SERVICE dimension. The confirmatory factor analysis shows an overall good fit of both models. The Chi-Squared is insignificant in the

Table 2
Validity Tests - South Central Grocery Store

DIMENSION ==>>	MERCHANDISE	ATMOSPHERE	VALUE	SERVICE	CLIENTELE	CONVENIENT
Scale mean and Std Dev (7 point)	3.13 1.08	3.09 1.36	3.57 1.23	3.47 1.29	4.01 1.51	3.29 1.35
Reliability	0.8160	0.9280	0.8848	0.6072	0.9173	0.7928
Lambda-X t-values (Convergent Validity)	11.91 , 11.14	13.99 , 15.26	13.01 , 11.64	7.36 , 7.53	13.77 , 12.58	11.84 , 8.75
D^2 for df=5 (Discriminant Validity)[1]	99.67	110.49	58.24	128.10	50.48	76.93
Chi-sqr	50.39	p = 0.10				
RMSEA	0.043	p = 0.60				
ECVI (Model)	0.83					
ECVI (Saturated)	1.01					
RMR	0.041					
GFI	0.95					

[1] The constrained models failed to converge after 20 iterations. The summary statistics of these models were used to calculate D^2 and should be interpreted with care.

Table 3
Validity Tests
Pacific Northwest Department Store

DIMENSION ==>>	MERCHANDISE	ATMOSPHERE	VALUE	SERVICE	CLIENTELE	CONVENIENT
Scale Mean and Std Dev (7 point)	2.90 1.15	2.44 1.29	3.92 1.40	3.41 1.32	3.46 1.45	3.09 1.20
Reliability	.7261	.8852	.8858	.6904	.8511	.7139
Lambda-X t-values (Convergent Validity)	9.45 , 9.55	13.67 , 11.36	9.77 , 11.31	8.98 , 7.71	10.91 , 9.90	7.68 , 7.83
D^2 for df=5 (Discriminant Validity)	137.51	136.38	35.47	75.67	49.69	39.63
Chi-sqr	58.81	p = 0.022				
RMSEA	0.063	p = 0.25				
ECVI (Model)	1.06					
ECVI (Saturated)	1.21					
RMR	0.040					
GFI	0.93					

grocery store model while significant in the department store model. The significance of the Chi-Square statistic is likely a function of the sample size. The RMSEA in the grocery store model is well below the 0.05 cutoff and the upper confidence interval is also below the 0.08 cutoff values suggested by Browne and Cudeck (1993). The department store model shows statistics only slightly beyond these values. However, the p-value for "test of a close fit (RMSEA: 0.05)" is not significant in either model (p = 0.25). Another indication that the model fits well is that the ECVI for the model (0.83 and 1.06) is less than for the saturated model (1.01 and 1.21) (Joreskog and Sorbom 1993). Finally, the GFIs (.95 and .93) are well above the commonly recommended 0.90 limit (Lichtenstein et al. 1992). Overall, the fit measures indicate a unidimensional model in both the grocery and department store models.

Convergent validity was tested by the examination of the t-values of the Lambda-x matrix. These values range form 7.36 to 15.26, all well above the 2.00 level specified by Kumar et al. (1992). This indicates high convergent validity of the store image scale in both models. The tests for discriminant validity all indicate that the model for the store image scale fits significantly better than the constrained models for each dimension, as indicated by the high D statistics. This indicates that the dimensions of the store image construct display high discriminant validity.

Scale Discussion

The store image scale developed in this research offers an actionable, cost efficient, and minimal effort approach to determining and altering, if necessary, individual store image. It is limited to those aspects of image a manager can modify and utilize as a competitive tool. The scale provides a succinct instrument that enables determination of, with some precision, exactly what aspects/dimensions of the store are appealing to clientele and which ones generate negative reactions. As such, the scale provides a technique that is actionable by the individual store manager, unlike more global store image scales. The scale will be more useful than scales that fail to specify what about the image is faulty or that have elements such as location or reputation that are beyond the control of the store manager. Six aspects or constructs were identified in the purified models. These constructs, viewed as being under the control of the store manager, are merchandise, atmosphere, value, service level, clientele, and convenience.

Retailers are under increasingly tight budget constraints. The scale can be implemented within parameters which are cost efficient and require minimal time and effort, therefore meeting budget constraints. The scale is operationalized through a questionnaire that can be added to any customer

survey or be administered at the point of purchase in a short time. This provides a cost efficient, minimal effort technique for use in a variety of types of retail operations. It also allows continuous monitoring and examination of store image without conducting extensive research for each iteration.

Store image is of critical concern in a wide variety of retail stores. Though the testing of the image scale involved a department store and a grocery store, it is applicable for a broad category of different types of retailers in differing geographic locations. Therefore, the scale can be used in a broad range of retailing to tailor image to effectively reach targeted customers in a variety of geographic settings. This characteristic of generalizability, combined with actionability, affords the retail manager with a dynamic scaling technique which can be applied at the individual store level.

Scale Limitations

The store image scale that was developed is not without limitations. It suffers the same limitations of other multi-attribute scales that measure global constructs. Specifically, this scale may fail to capture the fullness of store image construct. As mentioned earlier, image variables were limited to those that are actionable by the manager. It may be that the items omitted can, in some cases, affect store image. The scale was shortened for ease of use by managers. The resulting tradeoff between scope for more parsimony and greater specifics was necessary due to the target use of the scale.

Another potential problem arises from the use of two item dimensions. In this instance, some robustness of measurement was sacrificed for parsimony. However, both the reliability and validity of the scale appear to be relatively high.

Suggestions For Future Research

It is incumbent upon academics who have the ability and tools to focus on the issue of applied scale development. Not only does the manager need scales to be developed for use, but also the underlying theory of applied scale development needs to be further addressed. This paper provides a starting point and suggestions for developing applied scales. However, it is just that: a starting point. Researchers need to concentrate effort on further developing the methods and norms for improvement of this type scale. After this, the work begins in earnest. A few of the areas in desperate need of applied scales include: customer satisfaction, employee evaluation, and general image.

Conclusion

The development of the applied store image scale differed from what would have been undertaken for scientific research. As compared to traditional store image scales, developed for scientific research, the applied scale is much shorter and the items measured are more actionable for store managers. The applied store image scale allows managers to determine, with some precision, exactly what characteristics of the store are appealing to clientele and which ones generate negative reactions. Testing indicated the scale is useful for determining customer's perceptions of store image across a broad range of retail stores in different geographic areas. In addition, a more rigorous test of scale validity and reliability has been provided. ▱

❈❈❈ References ❈❈❈

1. Baggozi, Richard P., "Evaluating Structural Equation Models With Unobservable Variables and Measurement Error: A Comment," *Journal of Marketing Research*, Vol. 18 (August), 375-81, 1981.
2. Baker, Julie, Dhruv Grewal, and A. Parasuraman, "The Influence of Store Environment on Quality Inferences and Store Image," *Journal of the Academy of Marketing Science*, Vol. 22, No. 4, 328-9, 1994.
3. Berry, Leonard L., "The Components of Department Store Image: A Theoretical and Empirical Analysis," *Journal of Retailing*, Vol. 45, No. 1, 3-20, 1969.
4. Browne, M.W. and R. Cudeck, "Alternative Ways in Assessing Model Fit", in *Testing Structural Equations Models*, ed. Bollen and Long, Sage Publications, 1993.
5. Campbell, Donald R. and Donald W. Fiske, "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin*, Vol. 56, 81-105, 1959.
6. Calder, Bobby J., Lynn W. Phillips and Alice M. Tybout, "Designing Research for Application," *Journal of Consumer Research*, Vol. 8 (Sept), 197-207, 1981.
7. Calder, Bobby J., Lynn W. Phillips and Alice M. Tybout, "The Concept of External Validity" *Journal of Consumer Research*, Vol. 9 (Dec), 240-3, 1982.
8. Calder, Bobby J., Lynn W. Phillips and Alice M. Tybout, "Beyond External Validity" *Journal of Consumer Research*, Vol. 10 (June), 112-4, 1983.
9. Churchill, Gilbert, "A Paradigm for Developing Better Measures of Marketing Constructs", *Journal of Marketing Research*, Vol. 16 (Feb), 64-73, 1979.
10. Cronbach, Lee J., "Test Validation," in *Educational Measurement*, 2nd ed., R.L. Thorndike, ed. American Council on Education: Washington D.C., 443-507, 1971.
11. Darden, William R. and Barry J. Babin, "Exploring the Concept of Affective Quality: Expanding the Concept of Retail Personality," *Journal of Business Research*, Vol. 29, No. 2, 101-9, 1994.
12. Dichter, Ernest, "What's in an Image?", *Journal of Consumer Marketing*, Vol. 2 (Winter), 75-81, 1985.
13. Dickson, John P. and Douglas L. MacLachlan, "Social Distance and Shopping Behavior," *Journal of the Academy of Marketing Science*, Vol. 18, No. 2, 153-61, 1990.
14. Donovan, Robert J. and John R. Rossiter, "Store Atmosphere: An Environmental Psychology Approach," *Journal of Retailing*, Vol. 58, No. 1, 34-57, 1982.
15. Doyle, Peter and Ian Fenwick, "How Store Image Affects Shopping Habits in Grocery Chains," *Journal of Retailing*, Vol. 50, No. 4, 39-52, 1975.
16. Fitzgerald, Kate, "The Latest Side of Sears," *Advertising Age*, Vol. 65, No. 19, 1-48, 1994.
17. Gerbing, David and James Anderson, "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment", *Journal of Marketing Research*, Vol. 25 (May), 186-92, 1988.
18. Hansen, Robert A. and Terry Deutscher, "An Empirical Investigation of Attribute Importance in Retail Store Selection," *Journal of Retailing*, Vol. 53, No. 4, 54-72, 1978.
19. Hawkins, Del I., Gerald Albaum, and Roger Best, "Reliability of Retail Store Images as Measured by the Stapel Scale," *Journal of Retailing*, Vol. 52, No. 4, 31-8, 1976.
20. Heeler, Roger M. and Michael L. Ray, "Measure Validation in Marketing," *Journal of Marketing Research*, Vol. 9 (Nov), 361-70, 1972.
21. Hemmasi, Masoud, Kelly C. Stron, and Steven A. Taylor, "Measuring Service Quality for Strategic Planning and Analysis in Service Firms," *Journal of Applied Business Research*, Vol. 10, No. 4, 24-33, 1994.
22. Hildebrandt, Lutz, "Store Image and the Prediction of Performance in Retailing," *Journal of Business Research*, Vol. 17, No. 1, 91-100, 1988.
23. Hirschman, Elizabeth C., Barnett Greenberg, and Dan H. Robertson, "The Intermarket Reliability of Retail Image Research: An Empirical Examination," *Journal of Retailing*, Vol. 54, No.1, 3-12, 1978.
24. Jacoby, Jacob, "Consumer Research: Telling it Like it Is," in *Advances in Consumer Research*, eds. B.B Anderson, Association for Consumer Research: Chicago, 1-11, 1976.
25. James, Don L., Richard M. Durand, and Robert A. Dreves, "The Use of a Multi-Attribute Attitude Model in a Store Image Study," *Journal of Retailing*, Vol. 52, No. 2, 23-32, 1976.
26. Joreskog, Karl and Dag Sorbom, *LISREL8: Structural Equation Modeling with the SIMPLIS Command Language*, Chicago: Scientific Software, 1993.

27. Joreskog, Karl, "Statistical Analysis of Sets of Codgen-ric tests," *Psychometrica*, Vol. 36 (June), 109-33, 1971.

28. Kassarjian, Harold H., "Personality and Consumer Behavior: A Review," *Journal of Marketing Research*, Vol. 8 (Nov), 409-18, 1971.

29. Kasulis, Jack J. and Robert F. Lusch, "Validating the Retail Store Image Concept," *Journal of the Academy of Marketing Science*, Vol. 9, No. 4, 419-35, 1981.

30. Keaveney, Susan and Kenneth Hunt, "Conceptualization and Operationalation of Retail Store Image: A Case of Rival Middle-Level Theories," *Journal of the Academy of Marketing Science*, Vol. 20:, No. 2, 165-75, 1992.

31. Kelly, Robert F. and Ronald Stephenson, "The Semantic Differential: An Information Source for Designing Retail Patronage Appeals," *Journal of Marketing*, Vol. 31 (October), 43-7, 1967.

32. Kumar, Nirmalya, Louis Stern, and Ravi Achrol, "Assessing Reseller Performance From the Perspective of the Supplier," *Journal of Marketing Research*, Vol. 29 (May), 238-53, 1992.

33. Kunkel, John H. and Leonard L. Berry, "A Behavioral Conception of Retail Image," *Journal of Marketing*, Vol. 32 (Oct), 21-7, 1968.

34. Lichtenstein, Donald, Nancy Ridgway, and Richard Netemyer, "Price Perceptions and Consumer Shopping Behavior: A Field Study," *Journal of Marketing Research*, Vol. 30 (May), 234-45, 1992.

35. Lindquist, Jay D., "Meaning of Image", *Journal of Retailing*, 50 (Winter), 29-38, 1974-75.

36. Lutz, Richard, and James Bettman, "Multiattribute Models in Marketing: a Bicentennial Review," In *Consumer and Industrial Buying Behavior*, Eds. Arch Woodside, Jagdish Sheth, and Peter Bennett. New York: Elsevier-North Holland, 137-49, 1977.

37. Lynch, John J. Jr., "On the External Validity of Experiments in Consumer Research," *Journal of Consumer Research*, Vol. 9 (Dec), 225-39, 1982.

38. Lynch, John J. Jr., "The Role of External Validity in Theoretical Research," *Journal of Consumer Research*, Vol. 10 (June), 109-114, 1982.

39. Malhotra, Naresh K., "A Threshold Model of Store Choice," *Journal of Retailing*, Vol. 59, No. 2, 3-21, 1983.

40. Martin, Ingrid M. and Sevgin Eroglu, "Measuring a Multi-Dimensional Construct: Country Image," *Journal of Business Research*, Vol. 28, 191-210, 1993.

41. May, Eleanor, "Product Positioning and Segmentation Strategy: Adaptable to Retail Stores?", In *Theory in Retailing: Traditional and Nontraditional Sources*, Eds. R. Stampfl and E. Hirchman, Chicago, IL: American Marketing Association, 144-153, 1981.

42. McDougall, G.H. and J.N. Fry, "Combining Two Methods of Image Measurement," *Journal of Retailing*, Vol. 50, No. 4, 53-61, 1975.

43. Menezes, Dennis and Norbert Elbert, "Alternative Semantic Scaling Formats For Measuring Store Image: An Evaluation," *Journal of Marketing Research*, Vol. 16 (Feb), 80-7, 1979.

44. Nunnally, Jum C., *Psychometric Theory*, 2nd ed. McGraw-Hill: New York, 1978.

45. Ohanian, Roobina and Armen Tashchian, "Consumers' Shopping Effort and Evaluation of Store Image Attributes: The Roles of Purchasing Involvement and Recreational Shopping Interest," *Journal of Applied Business Research*, Vol. 8, No. 4, 40-9, 1992.

46. Oxenfeldt, Alfred, "Developing a Favorable Price-Quality Image," *Journal of Retailing*, Vol. 50 (Winter), 8-14, 115, 1974.

47. Parameswaran, Ravi, Barnett A. Greenberg, Danny N. Bellenger and Dan H. Robertson, "Measuring Reliability: A Comparison of Alternative Techniques," *Journal of Marketing Research*, Vol. 16 (Feb), 18-25, 1979.

48. Pathak, Dev S., William J. E. Crissy, and Robert W. Sweitzer, "Customer Image Versus the Retailer's Anticipated Image," *Journal of Retailing*, Vol. 50, No.4, 21-8, 1975.

49. Pessemier, Edgar A., "Store Image and Positioning," *Journal of Retailing*, Vol. 56, No. 1, 94-106, 1980.

50. Peter, J. Paul, "Reliability: A Review of Psychometric Basics and Recent Marketing Practices," *Journal of Marketing Research*, Vol. 16 (Feb), 6-17, 1979.

51. Peter, J. Paul, "Construct Validity: A Review of Basic Issues and Marketing Practices," *Journal of Marketing Research*, Vol. 18 (May), 133-45, 1981.

52. Rogers, Everett, "New Product Adoption and Diffusion," *Journal of Consumer Research*, Vol. 2 (Mar), 290-301, 1976.

53. Sirgy, M. Joseph and A. Coskun Samli, "A Path Analytic Model of Store Loyalty Involving Self-Concept, Store Image, Geographic Loyalty, and Socioeconomic Status," *Journal of the Academy of Marketing Science*, Vol. 13, No. 3, 265-91, 1985.

54. Steenkamp, E. Jan-Benedict and Michel Wedel, "Segmenting Retail Markets on Store Image Using a Consumer-Based Methodology," *Journal of Retailing*, Vol. 67, No. 3, 300-20, 1991.

55. Zelnio, Robert N. and Jean P. Gagnon, "The Construction and Testing of an Image Questionnaire," *Journal of the Academy of Marketing Science*, Vol. 9, No. 3, 288-99, 1981.

56. Zimmer, Mary and Linda Golden, "Impressions of Retail Stores: A Content Analysis of Consumer Images," *Journal of Retailing*, Vol. 64, No. 3, 265-93, 1988.