THE VALIDITY OF ALTERNATIVE MEASURES OF ATTRIBUTE IMPORTANCE: THE ROLE OF PRODUCT PURCHASE FAMILIARITY

James H. Martin, John Carroll University
Cheryl Winteregg, John Carroll University

Abstract

Although product attribute importance is central to developing marketing efforts targeted to the consumer, there have been relatively few published studies assessing the "goodness" of alternative measures of attribute importance. The research that has been done on this issue finds significant inconsistencies among the various measures. This paper reviews the research that compares reliabilities and validities of alternative measures of attribute importance and offers a theoretical explanation for the apparent lack of convergence among the different measures using purchase familiarity as a key concept. Directions for further research on this issue are discussed.

Attribute importance is a critical concept in marketing strategy development. Features of the product important to consumers come into use in product development, product positioning and market segmentation activities. In fact, developing an effective marketing strategy directed toward the user/buyer is difficult without the knowledge of what is important to those potential buyers.

Due to the prominent role of attribute importance, measurement of attribute importance is of tremendous significance. Several different measurement methods have been proposed for ascertaining the level of importance of product features to buyers. There has been relatively little research comparing the validity of attribute importance measures. Although some studies have been encouraging, most of the research seems to indicate a lack of convergence in measurement methods for attribute importance. The lack of convergence in measurement methods, which may be the result of any number of factors, has led many researchers to conclude that the various measures are either not measuring the same construct or some measures are appropriate and some measures are simply bad measures.

The purpose of this paper is twofold. First, the purpose is to review the existing research that has investigated the "goodness" of attribute importance measures. Second, the purpose is to account for the low convergence between measures found in previous studies by using product purchase familiarity as a central explanatory device. This paper will proceed as follows: 1) define attribute importance and describe the alternative measurement methods currently in use, 2) examine the research on this issue focusing on issues of reliability and validity, 3) develop a theoretical rationale for why such inconsistencies exist in the research, and 4) suggest directions for future research and outline managerial implications.

Definition of Attribute Importance

Attribute importance is conceptualized in past research as having to do with the impact of changes in the amount of the attribute on product/brand preference. Several different conceptualizations have been used but all are relatively compatible with one another (e.g. Jaccard, Brinberg & Ackerman 1986; Sheluga, Jacoby, & Jaccard 1980; Danes, Johnston, Cattin, and Hunter 1980; Heller, Okechuku and Reid 1979; Leigh, Mackay & Summers 1984; Mackenzie 1986; Myers & Alpert 1968; Myers 1970; and Neslin 1981).

In keeping with previous definitions, attribute importance will be conceptualized as the degree to which the attribute can influence product preference. That is, the diagnostic value of a product attribute for judgments of preference is considered to be the attribute's importance, where the diagnostic value is based on a full range of attribute levels and not just the current levels of attributes in current alternatives.

Methods of Measuring Attribute Importance

Several methods have been suggested for measuring attribute importance. Some of the more commonly used methods include conjoint analysis, direct self-report rating scales, constant sum scales, paired comparisons, graded paired comparisons, open-ended elicitation, dollar metric scaling and the information display board. Each method purportedly measures attribute importance, yet as stated earlier, little convergence is found between measures. It may, therefore, be instructive to examine each measure with respect to the nature of the judgment task that each method presents to the respondent.

Open-ended elicitation asks the respondent to list the most important attributes in making a purchase decision. This method probably focuses on the respondent's salience of attributes (Jaccard, Brinberg, & Ackerman 1986) by forcing the subject to search his/her memory for product features. Paired comparisons, graded paired comparisons, and the constant sum scales require the respondent to identify the importance of an attribute relative to other attributes, hence, the judgment task is comparative in nature. Direct self-report rating scales are non-comparative in nature, requiring the respondent to identify the importance of the attribute to the purchase decision

isolated from other attributes. Conjoint analysis assesses the respondent's value for an attribute by having the respondent indicate his/her preferences for alternatives that differ in amounts of various attributes. The task essentially requires the respondent to tradeoff values of attributes against one another. The extent to which changes in an attribute contribute to changes in preferences across alternatives is the indicator for attribute importance. Dollar metric scaling requires to respondent to place a monetary value on changes in attributes. The different measurement methods require the respondent to engage in different judgment tasks or different types of cognitive processing to arrive at the judgment of attribute importance.

Previous Research Comparing Measurement Methods

Evaluating the "goodness" of a measure of attribute importance requires examination of the measure's reliability and validity.

Reliability

Although there have been only a limited number of studies comparing alternative measures of attribute importance, there have been even fewer studies that examined the reliability of the different measurement forms. Two studies that reported reliability coefficients for alternative measures reported conflicting findings. Heeler, Okechuku, and Reid (1979), using a form of inter-rater reliability, found reliability coefficients of .81, .71, and .76 for their self report, conjoint measurement, and information display board measures of attribute importance, respectively. Leigh, Mackay and Summers (1984), assessing test-retest reliability, reported much lower reliability coefficients for their measures of attribute importance. Those authors reported reliability coefficients of .41 - .54 for various methods of conjoint attribute weight measures and coefficients of .38 - .51 for various forms of self-explicated attribute weight measures. Leigh, Mackay and Summers (1984) reported much higher reliability coefficients for estimated stimulus utilities based on these measures (e.g. .79 and .85 for conjoint analysis and self reports, respectively), however as measures of attribute weights, the measures did not appear to perform well.

Validity

Two types of validity, convergent and predictive, have been assessed in past research to evaluate the "goodness" of alternate measures of attribute importance. This past research shows conflicting results. For example, in assessing convergent validity, Danes, Johnston, Cattin, & Hunter (1980) found extremely high convergence between conjoint analysis utility estimates and their measure of psychological relevance (subjective conditional probability based on Jaccard, Knox & Brinberg (1977)), reporting a validity coefficient between those two measures of attribute importance of approximately .95. Conversely, Heller, Okechuku, and Reid (1979) found relatively low convergence between conjoint analysis and the information display board ($r=.32$) and only moderate convergence between a self report measure and the information display board and conjoint analysis (coefficients were .527 and .588). Sheluga, Jaccard and Jacoby (1979) examined the convergence between direct rating scales, conjoint analysis, and graded pared comparisons. Sheluga et al. (1979) reported low convergence between direct ratings of attribute importance and their other two measures of attribute importance ($r=.38$ for both cases) and higher convergence between the conjoint analysis and the graded paired comparison methods ($r=.67$). Finally, Jaccard, Brinberg and Ackerman (1986) compared six measures of attribute importance (open ended elicitation, paired comparisons, direct ratings, conjoint analysis, information display board, and psychological relevance as subjective conditional probability) for two product categories (cars and birth control). This last study reported extremely low levels of convergence between all six methods (validity coefficients ranged from .15 to .68 with only 4 of the 30 coefficients above .50). The highest correlations occurred between direct ratings and the subjective conditional probability measures leading those authors to conclude method variance accounted for the small amount of convergence found between the alternative methods.

Predictive validity (using attribute importance as measured by different methods to predict product preference) has been assessed in at least two studies (Leigh, Mackay & Summers, 1984; Neslin 1981). Neslin (1981) compared the predictive validity of two statistical methods for assessing attribute importance and a direct, self report rating scale. He found moderate validity coefficients for the direct ratings method ($r=.47$ to .52) and reasonably high coefficients for the statistical models (.56 to .67). However, in examining each method's ability to recover the original rankings of attributes, none of the methods were able to predict over 66% of the original rankings. Leigh, Mackay and Summers (1984) compared several methods of conjoint analysis and self explicated weights with respect to their predictive validities. Predictive validity was measured using the percentage of subjects whose choices were predicted correctly based on estimated attribute weights. Those authors reported finding no meaningful differences between conjoint tasks and self report measures in terms of predictive validity. They also reported low predictive validities for both conjoint and self explicated weights (36% and 35% correct predictions, respectively).

Previous research was conflicting in finding convergence at all, and when convergence between measures was found, research seemed to be inconsistent in which types of measures produced similar results. For example, Sheluga et al. (1979) found the graded paired comparisons and conjoint analysis methods to be highly

related but neither were related to the direct self report ratings. Jaccard et at. (1986) found paired comparisons and conjoint analysis to be only slightly correlated but found the subjective probability and direct ratings to be highly related, and Heeler et al. (1979) found moderate convergence between a self report direct rating scale and conjoint analysis. Reliability and predictive validity studies also reported inconsistent results.

In examining past research, what conclusions can be drawn? Some authors have suggested that, in the face of inconsistent results, one should rely on the face validity of the measure and suggested which they felt was the best measure based on this criterion (Heeler, Okechuku & Reid 1979; Neslin 1981). Jaccard, Brinberg and Ackerman (1986), when confronted with quite low validity coefficients for all six measures, suggested that the different methods of measuring attribute importance may be measuring different aspects of the importance domain. Those authors suggested attribute importance might be a multidimensional concept and that future research should examine how the several dimensions of importance map into product evaluations. Leigh, Mackay and Summers (1984) suggested situational variations may produce convergence under some conditions and inconsistencies in other situations. One of the factors influencing the validity of the alternative measures suggested by Leigh et al. (1984) was whether the product and product attributes are familiar to respondents.

The logical direction to pursue for a resolution of the inconsistency in measures is to further develop theoretical explanations for the inconsistencies and test those theoretical explanations. The concept of product familiarity will be pursued further in the present paper as having a role in influencing the equivalence of alternative measures of attribute importance.

The Role of Product Purchase Familiarity

There is a significant amount of research from a variety of perspectives suggesting that familiarity with an object influences judgments and judgment processes about that object (e.g. Hoyer 1984; Loken & Hoverstad 1985; Mackenzie 1986; Marks & Olson 1981; Rao & Monroe 1988). If we assume attribute importance judgments occur prior to the product evaluation process, and then are adjusted based on the feedback received from the actual product purchase, a fairly simple theoretical explanation for inconsistencies in results from alternative methods for measuring attribute importance may be suggested.

It is hypothesized that familiarity with the purchase of the product (not just familiarity with the product) helps the consumer to develop firm, concrete judgments about the importance of product features as they are related to the purchase decision process. That is, by making the purchase decision over several trials and experiencing the consequences of those purchase decisions, consumers establish a knowledge structure pertaining to those product features that are actually important. For example, Marks and Olson (1981) and Rao and Monroe (1988) (among others) suggest familiarity leads to a more well developed cognitive structure about the product. In addition, Loken and Hoverstad (1984) suggest that the judgment task influences the way in which information is coded and stored (i.e. its concreteness and organization) and that the manner in which the information is stored influences subsequent judgments. Therefore, familiarity with actually purchasing the product should enhance the concreteness of judgments of attribute importance (for a similar argument regarding the polarity of judgments, see Mackenzie (1986).

That is, when an individual first purchases a product, that individual should be fairly uncertain about what features are important. As the individual increases the experiences with purchasing the product he/she receives more feedback about the results of the purchases and uncertainty about attribute importance in the purchase process should decrease. Therefore the role of product purchase familiarity is suggested to be that of reducing uncertainty about what product features are important relative to the purchase decision. As stated earlier, product purchase familiarity is considered to be the important factor, not just product familiarity. This is because the concern is with understanding and measuring attribute importance as it relates to the purchase decision.

How does this proposed theoretical process help in resolving the inconsistencies across alternative measures of attribute importance? If an individual is uncertain about the importance of an attribute, then requiring him/her to identify the level of importance under different judgment tasks (e.g. direct isolated judgments, relative judgments and tradeoff judgments) should produce inconsistent responses because the different judgment tasks require the respondent to think differently about the product. Conversely, if the respondent has purchased the product frequently, his/her knowledge structure and accompanying attribute weights are firmly established. In the latter situation, the different judgment tasks required by the alternative measures should have limited impact on the resulting judgments of importance. So, for example, Jaccard et al. (1986) used automobiles as their product category and found virtually no convergence between the different types of measures. It seems unlikely that college sophomores (the population base in that study) had purchased few, if any, automobiles. Although respondents were familiar with the product, they were not familiar with the purchase of that product and were inconsistent in starting what was important in <u>purchasing</u> the product. That sample may not have had a well developed knowledge schema for <u>purchasing</u> automobiles. Hence, when asked to make judgments about what is important in purchasing automobiles, respondents were

inconsistent in their answers. Danes et al. (1980), using choice of spring break location as the product and college students as the sample found high convergence. College students may have thought more about spring break locations and have chosen locations in the past, such that they would have developed the requisite schemes for making consistent importance judgments across types of measures.

Future Research and Managerial Implications

The issue of measuring attribute importance correctly is vital to developing marketing strategy. Given its significance, research is needed to extend our knowledge of and ability to measure attribute importance. Several directions of research are suggested based on the previous discussion:

1) More research is needed comparing the reliability and validity of several alternative measures of attribute importance under differing levels of purchase familiarity. Under low purchase familiarity one would expect to find low convergent validity for the alternative measures. Under conditions of high familiarity, a high degree of convergence among the measures should result.

2) To further test the notion of the relationship between purchase familiarity and attribute importance uncertainty, "before and after" experiments should be conducted in which familiarity is increased over time and uncertainty in importance judgments is measured.

3) From a managerial perspective, knowing how much purchase familiarity is necessary to reduce uncertainty in importance judgments would be valuable in helping identify whether the method of attribute importance measurement should be a concern.

4) Managers must make decisions regarding the level of complexity for a market study involving attribute importance. Determining if there are situational variations in the validity of measures of attribute importance due to product purchase familiarity would help the manager assess the value of increasing the complexity of the measurement process (e.g. using conjoint analysis rather than direct rating scales).

5) For those situations in which there is low convergence between measurement methods, one of two managerial implications exist. First, it may be that, in such situations, one type of measure may be "better" than another, implying research must determine for the manager, which method is "best." Second, following the suggestion of Jaccard, Brinberg and Ackerman (1986), low convergence may mean that different measures assess different dimensions of attribute importance, implying research must determine for the manager the nature of the multidimensional importance construct in that setting such that a comprehensive assessment of attribute importance could be made.

Attribute importance is used for developing several components of a marketing strategy. Therefore, accurate measurement of attribute importance is critical to effective marketing. Previous research indicates significant inconsistencies in the appropriateness of alternative measurement methods. A theoretical explanation was provided for the inconsistencies using the concept of product purchase familiarity. Much research is needed to further examine the extremely important issue.

References

Danes, J., Johnson, J., Cattin, P., & Hunter, J. (1980). "Evaluating Marketing Communication Strategies: Attribute Importance versus Transfer Discrepancy," Advances in Consumer Research, 7, Jerry Olson (ed.), Ann Arbor MI: ACR, 561-565.

Heeler, R., Okechuku, C., & Reid, S. (1979). "Attribute Importance: Contrasting Measurements," Journal of Marketing Research, 16, 60-63.

Hoyer, W. (1984). "An Examination of Consumer Decision Making for a Common Repeat Purchase Product," Journal of Consumer Research, 11, 822-829.

Jaccard, J., Brinberg, D. & Ackerman, L. (1986). "Assessing Attribute Importance: A Comparison of Six Methods," Journal of Consumer Research, 12, 463-468.

Jaccard, J., Knox, R. & Brinberg, D. (1979). "Prediction of Behavior from Beliefs: An Extension and Test of a Subjective Probability Model," Journal of Personality and Social Psychology, 37, 1239-1248.

Leigh, T., Mackay, D. & Summers, J. (1984). "Reliability and Validity of Conjoint Analysis and Self Explicated Weights: A Comparison," Journal of Marketing Research, 21, 456-462.

Loken, B. & Hoverstad, R. (1985). "Relationships Between Information Recall and Subsequent Attitudes: Some Exploratory Findings," Journal of Consumer Research, 12, 155-167.

Mackenzie, S. "The Role of Attention in Mediating the Effect of Advertising on Attribute Importance," Journal of Consumer Research, 13, 174-195.

Marks, L. & Olson, J. (1981). "Toward a Cognitive Structure Conceptualization of Product Familiarity," in Advances in Consumer Research, 8, ed. Kent Monroe, Ann Arbor MI: ACR, 145-150.

Myers, J. (1970). "Finding Determinant Buying Attitudes," Journal of Advertising Research, 10, 9-12.

Myers, J. & Alpert, M. (1968). "Determinant Buying Attitudes: Meaning & Measurement," Journal of Marketing, 32, 13-20.

Neslin, S. (1981). "Linking Product Features to Perceptions: Self-Stated versus Statistically Revealed Importance Weights," Journal of Marketing Research, 18, 80-86.

Rao, A. & Monroe, K. (1988). "The Moderating Effect of Prior Knowledge in Cue Utilization in Product Evaluation," Journal of Consumer Research, 15, 253-264.

Sheluga, D., Jaccard, J. & Jacoby, J. (1979). "Preference, Search & Choice: An Integrative Approach," Journal of Consumer Research, 6, 1979.