A Test of a Decompositional Method of Multiattribute Perceptions Measurement

Author(s): R. Kenneth Teas and Andrea L. Perr

Source: *Journal of Consumer Research* , Dec., 1989, Vol. 16, No. 3 (Dec., 1989), pp. 384-391

Published by: Oxford University Press

Stable URL: https://www.jstor.org/stable/2489519

A Test of a Decompositional Method of Multiattribute Perceptions Measurement

R. KENNETH TEAS
ANDREA L. PERR*

The purpose of this research is to test the temporal stability and validity of decompositional perceptions measures. The test is based upon a perceptions measurement experiment focusing on the perceived visual congruence between product concept statements and product prototypes that were designed by manipulating four product characteristics that affected product appearance. The results indicate the perceptions measures are characterized by considerable temporal stability, discriminant validity, and predictive validity.

Decompositional multiattribute measurement is a method used to examine the structure underlying individuals' responses to multiattribute stimuli, such as estimating utility functions for some product category. Although many decompositional multiattribute measurement procedures have been developed (Green and Srinivasan 1978; Lynch 1985), they all generally involve the decomposition of individuals' overall responses (such as preference ratings) to factorially designed stimuli so that the effects of the stimulus attributes on the responses can be systematically examined.

Most decompositional measurement in consumer research has focused on the measurement of preferences, but the approach shows promise in the measurement of perceptions, such as perceived convenience (Louviere and Gaeth 1987) and perceived complexity (Holbrook 1981). For example, Green and DeSarbo (1978) and Holbrook and Moore (1981) outline several advantages associated with decompositional perceptions measures when compared to direct scaling of perceptions. In addition, researchers have successfully incorporated decompositional perceptions measures into frameworks that specify perceptions as intervening variables between stimulus attributes and respondents' judgments (Holbrook 1981; Louviere 1984; Louviere and Gaeth 1987; Louviere and Meyer 1981).

The purpose of this research is to outline a procedure for decompositional perceptions measurement in the context of product concept testing and to test the procedure for temporal stability, discriminant validity, and predictive validity. The perceptions measurement method is based upon a conjoint measurement approach, but the multiattribute stimuli are not verbal product descriptions. They are actual product prototypes manufactured by blending attribute treatment levels based upon a factorial design. The research, therefore, extends the evidence of stability and validity of conjoint measurement of individuals' preferences for verbal product descriptions (Jain et al. 1979; McCullough and Best 1979; Teas 1985) into the arena of conjoint measurement of individuals' perceptions of multiattribute stimuli that consist of holistic blends of attributes rather than perceptually transparent verbal descriptions.

PERCEPTIONS AS INTERVENING VARIABLES

Most of the research in decompositional measurement in marketing, such as conjoint analysis applications, involves linking product attributes directly to affective responses, such as preferences. However, some evidence suggests that it may be more useful and valid to assume that perceptions mediate this linkage. For example, an important element in Louviere's (1984) hierarchical information integration method for the decompositional analysis of complex multiattribute judgment problems is the concept of higher order constructs, which are functions of the respondent's integration of information on antecedent variables. An example of a higher order construct is perceived convenience in retailing with antecedent variables involving such issues as travel time, parking, and check cashing (Louviere and Gaeth 1987). Other models that incorporate decompositional perceptions measurement include Hammond et al.'s

*R. Kenneth Teas is Professor of Marketing, College of Business Administration, Iowa State University, Ames, IA 50011. Andrea L. Perr is Project Manager—Consumer Research, Ralston Purina Company, St. Louis, MO 63164. The research was supported by a grant from the Iowa State University College of Business Administration. The authors wish to thank Sridhar N. Ramaswami, Iowa State University, and three anonymous reviewers for their helpful comments.

(1977) social judgment theory and policy capturing paradigms and Holbrook's (1981) model of evaluative judgment.

It might be possible to use direct scaling to measure perceived linkages between stimulus attributes and higher order constructs. An example of direct scaling would be to use a five-point importance rating scale to measure the relative importance of various product features as determinants of perceived product quality or styling attractiveness. However, decompositional perceptions measurement may be more appropriate. First, decompositional perceptions measurement provides information concerning the way attributes are integrated in perceptual processes by facilitating tests for interaction and curvilinearity effects (Louviere 1984; Louviere and Gaeth 1987). Second, it may be difficult for respondents to directly assess the role of specific attributes in perceptual judgments (see Green and DeSarbo 1978 for a general discussion), because it is often the holistic blend of attributes that creates the perceptual impression rather than the component parts. For example, Holbrook and Moore (1981) note:

> many products involving aesthetic, sensory, or symbolic benefits must be experienced to be judged adequately. Music, haute cuisine, or fashion designs must be heard, tasted, or seen to be properly appreciated.

In these situations, measurement approaches that present multiattribute stimuli in representational form, such as product prototypes, may be preferable to direct ratings of individual attributes. In fact, concealing the actual treatments being examined to avoid biasing the response is often desirable.

PERCEPTIONS MEASURES IN CONCEPT TESTING

A potential use of decompositional perceptions measurement is to use procedures based upon conjoint measurement (Green and De Sarbo 1978; Holbrook 1981) in product concept tests. For example, a product concept based upon consumer research could take the form of a verbal concept statement, descriptions of consumer benefits, or advertising copy. As in traditional conjoint measurement research, multiattribute product descriptions or prototypes are constructed using factorial designs. However, the respondent ranks or rates each stimulus in terms of its similarity to the concept statement rather than in terms of preference. Decomposition of the similarities data produce "part-similarity functions" (Green and DeSarbo 1978) that reflect the impact of each attribute on the degree to which the product stimuli are perceived by the respondent as similar to the concept statement.

Although research evidence indicates decompositional *preference* measures such as conjoint measurement tend to be stable and valid (Green and Srinivasan 1978; Jain et al. 1979; McCullough and Best 1979; Teas 1985), it is dangerous to generalize this evidence to perceptions measurement applications. For example, one important question is the degree to which decompositional *perceptions* measurement is characterized by discriminant validity with respect to decompositional *preference* measurement. For example, discriminant validity would be lacking in a situation in which respondents instructed to evaluate stimuli in terms of similarity to some concept merely rate the stimuli in terms of preference.

It is not known whether the findings of stability and validity of conjoint measurement using perceptually transparent, verbal listings of attributes can be extended to perceptions measurement using stimuli consisting of less transparent, nonverbal blends of attributes (e.g., product prototypes). For example, Holbrook and Moore (1981, p. 104) discuss important differences between evaluative judgment processes involving verbal descriptions (verbal processing) and those involving pictorial descriptions (imagery). Information is processed simultaneously, or more gestalt-like, when nonverbal multiattribute stimuli are evaluated, but sequentially when verbal multiattribute stimuli are evaluated. And, the "left and right (cerebral) hemispheres play dominant roles in the verbal and imagery (nonverbal) systems, respectively."

Empirical tests of some of the implications of these potential judgment process differences indicate decompositional measurement involving nonverbally described (pictorial) multiattribute stimuli may produce more main effects and more interaction effects than decompositional measurement involving verbally described multiattribute stimuli (Holbrook and Moore 1981). An important question that remains, however, is whether these judgment process differences result in differences in the inherent stability and validity of decompositional multiattribute measures.

METHOD

This study consisted of a perceptions measurement experiment focusing on the perceived visual congruence between product concept statements (dog food descriptions) and 27 prototypes (dog food products). Temporal stability (the degree to which respondents replicate their perceptions judgments of a set of multiattribute stimuli over time), predictive validity (the degree to which the estimated perception model is predictive of another variable), and discriminant validity (the degree to which the measure has low correlation with variables from which it is supposed to differ) were the decompositional perceptions measurement issues examined. Since actual product prototypes that blend various combinations of four attributes were used, the transparency of the attribute manipulations was reduced, which resulted in a more severe test of these stability and validity issues when compared to previous research on conjoint measurement.

EXHIBIT

PRODUCT CONCEPT STATEMENTS

1. Young dogs need extra protein because they are so active. This new dry dog food, which is made from *dried fish*, provides all the protein your dog needs.
2. Now there's a dog food that not only looks and tastes like real *uncooked meat*, but chews like real meat too. Moist and chewy pieces will give your dog all the eating pleasure he gets from uncooked real meat.
3. Give your dog that outdoor flavor of *barbecued ribs* in a dry dog food. This new dog food has the spicy taste of barbecued ribs *cooked* on an open fire.
4. Outdoor dogs have special nutritional needs. They need more fat in their diet than indoor dogs. This new dry dog food is a complete and balanced diet with *extra fat* to keep your outdoor dog healthy.
5. Give your dog a special treat at meal time. This new snack for dogs looks like *tablescraps*, but has the vitamins and minerals a dog needs.
6. Some dogs eat more food than their bodies require. By feeding your dog this new, *high fiber* dog food, he'll eat less, feel full, and lose those extra pounds.
7. Invite your dog to breakfast! Now there's a new dog food that combines the taste and nutrition of *eggs and sausage* for a breakfast made just for dogs.
8. Give your dog nutritious tasting pork flavored dry dog food. This new dog food looks, smells, and tastes like real *cooked pork*.
9. Give your dog that tasty flavor of *eggs* in a dry dog food. This new dog food formula has the taste and full nutritional value of real eggs.

The Product Concept Statements

Based upon research experience of the second author (one of her primary responsibilities as a consumer research project director during the last three years involved dog food concept development and testing) and from ideas generated from previous focus groups, nine dry dog food concept statements, which were hypothetical yet realistic, were developed. These concept statements are presented in the Exhibit. The first four statements were used in the perceptions experiment. The remaining five were created so that a sufficient number were available to use multidimensional scaling to examine the differentiation among the four test concepts.

Product Prototypes

Twenty-seven dog food prototypes were produced by a pet food manufacturing company by manipulating four product attributes according to a $3 \times 3 \times 3 \times 2$ fractional factorial design. The attributes and corresponding treatment levels were (1) color—red, brown, and clear glaze; (2) coating[1]—thin, medium, and thick; (3) chunk size—small, medium, and large; and (4) coating layers—one sided or two sided. (The factorial design was based on a "Class One Compromise Plan" with 27 trials using three interacting factors from Addelman 1962.)[2] The effects of these attributes were expected to be different across concept statements. For example, the red treatment level of the color attribute was expected to produce a positive part-similarity value when associated with the uncooked meat concept statement, whereas it was expected to produce a negative part-similarity value when associated with the fish concept statement.

Measures

Data were collected from each respondent in three testing sessions. Usable response sets were obtained from 52 respondents: 41 undergraduate marketing majors and 11 employees of a local business organization.

Test 1 Data. Test 1 consisted of group interviews. The following data were collected.

1. Concept Statement Preference Rankings. The nine concept statements were ranked by each respondent in terms of "appeal to you as a dog owner."
2. Concept Statement Dissimilarity Ratings. All possible pairs of the nine concept statements were rated on a nine-point dissimilarity scale ranging from 1 ("The two concepts would be very similar in appearance") to 9 ("The two concepts would be very different in appearance").

Test 2 Data. Test 2 consisted of personal interviews with the same respondents in which the following data were collected.

1. Product Prototype/Concept Statement Congruence Ratings. Four sets of product prototype/concept statement congruence ratings were obtained from each respondent (one set for each of the first four concepts in Table 1). The procedures consisted of placing a concept statement in front of the respondent and having the respondent rate each of the 27 product prototypes on a scale ranging from 0 ("a dog food that is completely unlike the statement") to 100 ("a dog food that looks exactly like the statement"). Presentation of the concept statements and product prototypes was randomized.
2. Product Prototype Preference Ratings. Each respondent rated the 27 product prototypes on a preference scale ranging from 0 ("a dog food that is completely unappealing—that is, a dog food for which you would have extremely low preference") to 100 (" a dog food that is extremely appealing—that is, a dog food for which you would have ex-

[1]This attribute is a light coating material hypothesized to make the prototype look more like fish when the coating is thick and like beef when the coating is thin.

[2] This design permits the estimation of all main effects and the color-by-coating and color-by-size interaction effects. However, the color-by-coating layers and coating-by-size interaction effects are confounded, and no other interaction effects can be estimated.

TABLE 1

AVERAGE WITHIN-SUBJECT RESULTS USING PRETEST-RETEST DIFFERENCE SCORES[a] AND PRETEST SCORES[b]

	Mean part-similarity estimates							
	Fish model		Meat model		Ribs model		Fat model	
Treatment variable	Difference scores	Pretest scores	Difference scores	Pretest scores	Difference scores	Pretest scores	Difference scores	Pretest scores
X1								
Clear color	−0.53	11.11	−0.87	−19.12	−0.67	−22.59	1.87	8.50
	(−1.03)	(8.19)[d]	(−1.34)	(−13.38)[d]	(−1.19)	(−17.36)[d]	(3.22)[d]	(7.79)[d]
X2								
Brown color	1.17	2.21	−0.27	0.02	−0.65	0.49	−0.52	−1.41
	(3.34)[d]	(3.67)[d]	(−0.58)	(0.02)	(−1.40)	(0.61)	(−1.74)	(−3.11)[d]
Red color	−0.64	−13.32	1.14	19.10	1.32	22.10	−1.35	−7.09
	(−1.18)	(−10.36)[d]	(1.90)	(12.93)[d]	(1.97)	(20.22)[d]	(−2.70)[d]	(−7.18)[d]
X3								
Thin coating	−0.14	5.60	−0.14	−5.38	.10	−4.42	0.97	−12.03
	(−0.30)	(11.30)[d]	(−0.50)	(−7.07)[d]	(.33)	(−8.82)[d]	(2.13)[c]	(−12.54)[d]
X4								
Medium coating	−0.59	−1.32	0.98	0.94	0.34	1.50	−0.07	3.72
	(−1.32)	(−3.07)[d]	(2.97)[d]	(1.83)	(−1.25)	(3.88)[d]	(−0.20)	(5.81)[d]
Thick coating	−0.73	−4.28	0.84	4.44	−0.44	2.92	−0.90	8.31
	(1.55)	(−8.44)[d]	(−1.78)	(5.18)[d]	(−1.19)	(5.89)[d]	(−1.83)	(8.76)[d]
X5								
Large chunk	−0.99	−7.43	0.22	4.78	−0.24	−0.61	−0.03	6.17
	(−2.88)[d]	(−10.40)[d]	(0.60)	(7.22)[d]	(−0.79)	(−0.82)	(−0.10)	(8.04)[d]
X6								
Medium chunk	0.21	−0.65	−0.40	−1.13	0.58	−1.74	−0.22	−0.94
	(0.62)	(−1.14)	(−1.16)	(−2.89)[d]	(1.96)	(−3.74)[d]	(−0.79)	(−1.78)
Small chunk	0.78	8.08	0.18	−3.65	−0.34	2.35	0.25	−5.23
	(1.60)	(10.13)[d]	(0.42)	(−4.71)[d]	(−0.83)	(2.32)[c]	(0.79)	(−6.13)[d]
X7								
Coating on one side	−0.74	0.24	0.69	1.52	0.79	1.74	−0.05	−6.05
	(−1.86)	(0.45)	(2.33)[c]	(2.80)[d]	(3.35)[d]	(−2.25)[c]	(−0.26)	(−12.00)[d]
Intercept	−0.42	29.68	−0.85	46.89	−1.57	47.40	−0.35	56.94
	(−0.80)	(14.91)[d]	(−1.22)	(36.66)[d]	(−2.74)[d]	(39.82)[d]	(−0.55)	(44.00)[d]

[a] Temporal stability tests are based upon estimates using these difference scores as the dependent variable. Under the assumption of temporal stability, all part-similarity values are hypothesized to be equal to 0.
[b] Part-similarity estimates using pretest data are provided so that the relative magnitude of non-zero part-similarity values estimated using the pretest-retest difference scores can be assessed; t values are in parentheses.
[c] $p < 0.05$ for a two-tail test.
[d] $p < 0.01$ for a two-tail test.

tremely high preference"). The order of presentation was randomized.

Test 3 Data. Test 3 consisted of personal interviews with the same respondents one to two weeks after the Test 2 data were collected. A second set of product prototype/concept statement congruence ratings was obtained. The procedures were identical to the procedures used in Test 2 except that, for each respondent, the order of presentation of the concept statements in the pretest was eliminated as a possible order of presentation in the retest.

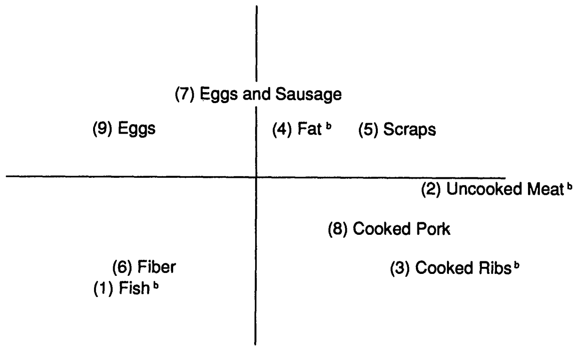
Perceptual Model

The multiattribute perceptual model[3] was:

$$Y_j^a = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7$$

[3]The question of interaction effects was examined by estimating regression models containing interaction terms. The within-subject analysis, which was based upon procedures used by Louviere and Meyer (1981), indicated little evidence of interaction effects. These findings were not surprising given other research on conjoint

FIGURE
MULTIDIMENSIONAL SCALING[a] OF THE NINE PRODUCT CONCEPT STATEMENTS



[a] Multidimensional Scaling Results (SPSSX ALSCAL); Stress = 0.199 and R^2 = 0.807.
[b] This concept statement was used in the perceptions measurement experiment.

where Y_j^a equals the congruence rating for the jth dog food product prototype with respect to concept a; β_0 equals the grand mean of the dependent variable; β^s equals the partial regression coefficients representing part-similarities; and X^s equals a vector of effects coded variables; that is, two variables for the color (X_1 and X_2), coating (X_3 and X_4), size (X_5 and X_6), and one variable for layers (X_7).

FINDINGS

Multidimensional Scaling Results

The ratings of the paired comparison concept statements on the nine-point dissimilarity scale were analyzed via multidimensional scaling. First, the 52 sets of dissimilarity data were averaged across respondents to produce one matrix of mean dissimilarities data. Second, the SPSSX version of ALSCAL was used to produce two-, three-, and four-dimensional solutions assuming metric measurement. The solutions produced similar results concerning the degree to which the four focal descriptions were differentiated. The two-dimensional solution, which is presented in the Figure, indicates the meat and ribs descriptions were perceived as similar by the respondents. The expected appearances of other focal product descriptions appear to be differentiated.

Temporal Stability and Discriminant Validity Tests

Pretest-Retest Difference Scores. One test of temporal stability was based on within-subject analysis of pretest-retest difference scores. First, each respondent provided four sets of ratings (27 product prototype/concept congruence ratings per set) in two separate interview settings. Based on these ratings, four sets of congruence rating difference scores were calculated for each respondent (one set of 27 pretest-retest difference scores for each of the four concepts). Second, these sets of difference scores were used as dependent variables in individual subject regression estimates of Equation 1. This procedure produced four estimated equations for each respondent (one estimate for each of the four concepts). Under the assumption of temporal stability, the null hypothesis is that all effects are zero for all respondents, and this was tested by calculating the mean of each parameter estimate across respondents and testing against its standard error.

The results, which are presented in Table 1, indicate nine estimated coefficients are statistically significant ($p < 0.05$). One of these significant estimates is an intercept term. In addition, the significant estimated mean part-similarities are small (when compared to the corresponding part-similarities estimated using the pretest data) for the large chunk treatment in the fish model and the thin coating treatment in the fat model. However, other significant coefficients are relatively large when compared to the corresponding pretest data estimates, and the total number of significant tests is larger than what would be expected by chance.

The pattern of significant mean estimates based upon the difference scores does not suggest obvious explanations for the results. Statistically significant coefficients were found in each of the four estimated models and were scattered across all four attributes. Moreover, identical procedures were used in both interviews, and the order of stimulus presentation was randomized. One possible explanation for the unexpected results is maturation error. Each respondent participated in three interviews and provided a total of 216 product prototype/concept statement congruence ratings. This data collection process may have induced maturation error in the form of learning, fatigue, or boredom effects that subtly altered the way some respondents processed the information.

Prediction Accuracy. A second test of temporal stability and discriminant validity was based upon the 52 sets of individual respondent perceptions model estimates using the pretest data. These estimates were used to predict each respondent's retest concept congruence ratings and prototype preference ratings. High prediction accuracy for identical concepts (e.g., using the meat part-similarity functions estimated from the pretest data to predict the retest meat congruence ratings) is evidence of temporal stability. Low prediction accuracy for different concepts (e.g., using the meat part-similarity functions to pre-

analysis (Green and Srinivasan 1978) and linear models in decision making (Anderson 1981, 1982).

TABLE 2

TESTS BASED UPON WITHIN-SUBJECT PREDICTION ACCURACY (r VALUES)[a]

Predicted rating	Fish retest congruence fish r	Meat retest congruence meat r	Ribs retest congruence ribs r	Fat retest congruence fat r	Preference r
Fish	.71 (.53)[b]	−.55	−.49	−.15	.09
Meat	−.59	.74 (.65)[b]	.65	−.02	.19
Ribs	−.55	.68	.84 (.78)[b]	−.13	.26
Fat	−.07	−.11	−.27	.67 (.22)[b]	−.31

[a] The r values are average within-subject correlations between (1) predicted values based upon the pretest congruence estimates and (2) retest congruence and preference ratings. The within-subject correlations were Z-transformed prior to calculating the mean correlation coefficient.
[b] Predictive validity using the color attributes (e.g., X_1 and X_2 in Equation 1) as the predictor model.

dict the retest fish congruence ratings) is evidence of discriminant validity.

These prediction accuracy tests produced 52 within-subject correlation matrices. The average correlation matrix, which is presented in Table 2, indicates the prediction accuracy for identical concept statements ranging from 0.67 to 0.84.[4] These results, which indicate relatively high temporal stability, are similar to the results of previous research testing the stability of conjoint preference measurement (Jain et al. 1979; McCullough and Best 1979; Teas 1985). In addition, with the exception of the correlation between the ribs and meat ratings,[5] the other correlations are low or negative, indicating high discriminant validity across different concepts and between congruence and preference ratings.[6]

The Part-Similarity Functions. The within-subject analysis produced 208 regression estimates based upon the pretest perceptions data (four regression estimates for each respondent). The findings, which are presented in Table 1, provide additional evidence of discriminant validity. First, the fish concept estimates indicate the less red (or more clear) a color treatment, the more the sample products resembled fish. In contrast, the meat concept estimates indicate the red color treatment tended to cause the sample to appear to be more meat-like. Second, as would be expected based upon the multidimensional scaling results, the meat and rib concept test results were similar. Third, the thick coating, clear color, large chunk, and two-sided coating treatments increased the degree to which the respondents perceived the prototypes to be similar to the fat concept statement.

Within-Subject Predictive Validity

The issue of predictive validity was examined by analyzing the within-subject correlations between the predicted retest concept congruence ratings and product preference ratings using the following procedures.

1. Each respondent's four part-similarity functions (estimated from the pretest data) were used to calculate four sets of 27 predicted product prototype/concept statement congruence scores—one set corresponding to each concept statement.
2. Each respondent's 27 prototype preference ratings were correlated with each of the four sets of 27 predicted product prototype/concept statement congruence scores (described in Point 1) producing 208 (52 × 4) within-subject correlation coefficients.
3. Under the assumption of predictive validity, the correlations between predicted prototype/concept

[4]Since the color attribute was an extremely important attribute in the fish, meat, and ribs models, a single color attribute model was estimated and prediction accuracy tests were made using the simplified model. As noted in Table 2, the prediction accuracy dropped for all of the models.

[5]Based upon the multidimensional scaling results, the correlations between the ribs and meat concepts were expected to be relatively high.

[6]The results of factor analysis of the combined within-subject correlation matrices also indicate considerable discriminant validity among the perceptions measures and between the preference and perceptions measures. The loadings on Factor 1 (eigenvalue = 2.67) were −0.96 for fish, 0.91 for meat, 0.85 for ribs, −0.04 for fat, and 0.15 for preference. The loadings on Factor 2 (eigenvalue = 1.24) were 0.12 for fish, 0.22 for meat, 0.28 for ribs, −0.84 for fat, and 0.75 for preference.

statement congruence scores and prototype preference ratings should be higher for concept statements ranked high in preference (identified in the preference rankings of the nine original concept statements) than for concept statements ranked low in preference. Prototypes predicted to be less congruent with preferred concept statements or highly congruent with non-preferred concept statements would be expected to be rated low in preference. Consequently, the concept statement preference rankings were used to group the 208 within-subject correlation coefficients (described in Point 2) into nine categories—coefficients associated with number one ranked concept statements, number two ranked concept statements, and so on.

4. The average within-subject correlation coefficients were calculated for each of the nine ranking categories.

The results provide considerable evidence of predictive validity.[7]

Preference rank	*Frequency of rank*	*Average within-subject correlation*
1	34	.56
2	24	.52
3	18	.46
4	19	.09
5	17	−.03
6	22	−.10
7	14	.11
8	20	−.22
9	36	−.54

The average within-subject correlations between the predicted congruence ratings and the prototype preference ratings were 0.56, 0.52, and 0.46 for concept statements that were ranked first, second, and third, respectively, in terms of preference. The remaining average within-subject correlations indicate that product samples predicted to be congruent with less preferred concepts were rated considerably lower in preference.

CONCLUSIONS

The issues examined in this study concern the degree to which measures of decompositional perceptions are characterized by temporal stability, predictive validity, and discriminant validity. These were examined in an experiment in which product prototypes were manufactured specifically for this study by blending various combinations of four attributes. Although the findings generally support the conclusion that the decompositional perceptions measurement method is characterized by considerable temporal stability and measurement validity, three important limitations of this study must be mentioned. First, a limited sample of respondents participated in the study. The findings, therefore, may be unique to the sample. Second, the perceptions measurement experiment was limited to perceptions measurement in the context of product concept testing. Clearly, additional research involving other perceptual issues is needed before general conclusions about decompositional perceptions measurement can be made. Third, the research design resulted in each respondent's participating in three interviews, which may have resulted in some maturation error in the form of learning, fatigue, or boredom effects.

Within-subject regression analysis was used to examine the issues of temporal stability and discriminant validity. With the exception of scattered pretest-retest instability, the results indicate the decompositional perceptions measurement method tested is characterized by considerable temporal stability and discriminant validity. Additional research, however, is needed to determine the degree to which the scattered instability found in this research is due to experimental error, such as maturation effects, or to a tendency for decompositional perceptions measures involving nonverbal, holistic blends of attributes (i.e., product prototypes) to be less stable than more traditional decompositional measures involving verbal stimulus descriptions. For example, blocked designs (Green 1974) or designs that do not require pretest and retest data from each respondent might be used in future research to reduce the likelihood of maturation error.

Predictive validity was examined at the within-subject level. In these tests, the correlations between the predicted concepts congruence ratings and preference ratings were higher for concepts ranked high in preference than for concepts ranked low in preference. However, this test of predictive validity is limited to the prediction of preference ratings. Consequently, additional research is needed that examines this predictive validity issue in the context of models that specify perceptions as intervening variables linking product attributes with affective responses and, perhaps, with choice behavior. Models proposed by Louviere (1984), Louviere and Gaeth (1987), and Holbrook (1981) represent promising frameworks for these more thorough predictive validity tests.

The results of this study indicate decompositional perceptions measures involving stimuli consisting of nonverbal, holistic blends of attributes (e.g., product

[7]Thirty-four respondents ranked one of the four test product concepts "number one." The remaining 18 respondents ranked one of the five other concept statements number one. The correlation value of 0.56 was calculated as follows. First, for each respondent, the number one ranked concept statement was identified. Second, for each respondent, the part-similarity regression estimate associated with this top ranked concept was used to forecast the 27 prototype preference ratings. Forecast accuracy was evaluated by calculating the within-subject correlation between the forecasted score and the preference score. The average correlation of 0.56 equals the average of these 34 within-subject correlations.

prototypes) are characterized by considerable temporal stability, discriminant validity, and predictive validity. The findings, therefore, offer preliminary support for the use of decompositional perceptions measures in future studies testing perceptual and information processing models. In addition, the results indicate decompositional perceptions measures may be useful in concept tests designed to examine the congruence between actual product prototypes and product concept statements.

[*Received July 1988. Revised April 1989.*]

REFERENCES

Addelman, Sidney (1962), "Symmetrical and Asymmetrical Fractional Factorial Plans," *Technometrics,* 4 (February), 47–58.

Anderson, Norman H. (1981), *Foundations of Information Integration Theory,* New York: Academic Press.

——— (1982), *Methods of Information Integration,* New York: Academic Press.

Green, Paul E. (1974), "On the Design of Choice Experiments Involving Multifactor Alternatives," *Journal of Consumer Research,* 1 (September), 61–68.

——— and Wayne S. DeSarbo (1978), "Additive Decomposition of Perceptions Data Via Conjoint Analysis," *Journal of Consumer Research,* 5 (June), 58–65.

——— and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research,* 5 (September), 103–123.

Hammond, Kenneth R., John Rohrbaugh, Jeryl Mumpower, and Leonard Adelman (1977), "Social Judgment Theory: Applications in Policy Formation," in *Human Judgment and Decision Making in Applied Settings,* eds. Martin F. Kaplan and Steven Schwartz, New York: Academic Press, 1–29.

Holbrook, Morris B. (1981), "Integrating Compositional and Decompositional Analysis to Represent the Intervening Role of Perceptions in Evaluative Judgments," *Journal of Marketing Research,* 18 (February), 13–28.

——— and William L. Moore (1981), "Feature Interactions in Consumer Judgments of Verbal Versus Pictorial Presentations," *Journal of Consumer Research,* 8 (June), 103–113.

Jain, Arun K., Franklin Acito, Naresh K. Malhotra, and Vijay Mahajan (1979), "A Comparison of the Internal Validity of Alternative Parameter Estimation Methods in Decompositional Multiattribute Preference Models," *Journal of Marketing Research,* 16 (August), 313–322.

Louviere, Jordan J. (1984), "Hierarchical Information Integration: A New Method for the Design and Analysis of Complex Multiattribute Judgment Problems," in *Advances in Consumer Research,* Vol. 11, ed. Thomas C. Kinnear, Provo, UT: Association of Consumer Research, 148–155.

——— and Gary J. Gaeth (1987), "Decomposing the Determinants of Retail Facility Choice Using the Method of Hierarchical Information Integration: A Supermarket Illustration," *Journal of Retailing,* 63 (Spring), 25–48.

——— and Robert J. Meyer (1981), "A Composite Attitude-Behavior Model of Traveler Decision Making," *Transportation Research,* 15B (December), 411–420.

Lynch, John G. (1985), "Uniqueness Issues in the Decompositional Modeling of Multiattribute Overall Evaluations: An Information Integration Perspective," *Journal of Marketing Research,* 22 (February), 1–19.

McCullough, James and Roger Best (1979), "Conjoint Measurement: Temporal Stability and Structural Reliability," *Journal of Marketing Research,* 16 (February), 26–31.

Teas, R. Kenneth (1985), "An Analysis of the Temporal Stability and Structural Reliability of Metric Conjoint Analysis Procedures," *Journal of the Academy of Marketing Science,* 13 (Winter), 122–142.