JOEL HUBER*

This study tests the ability of various models to predict individual preferences on stimuli defined by physical characteristics. Within all models, metric routines were found to be superior to nonmetric routines. Across models, differences in predictive ability were not found to be great compared to pragmatic or theoretical differences.

Predicting Preferences on Experimental Bundles of Attributes: A Comparison of Models

INTRODUCTION

In sharp contrast to the work that has been done on models relating preference to psychological variables [15], little work has been done to evaluate the relative effectiveness of preference models on objective attributes. In some instances physical attributes have the advantage of being more easily controlled than psychological attributes, so that less ambiguous determination of their causal effects can be made. This study compares the ability of various models to predict preferences for different bundles of objective attributes. In addition, the pragmatic and theoretic advantages of the models are explored.

The models are compared with respect to their effectiveness in predicting preferences on stimuli that differ over a restrictive set of physical dimensions. The stimulus set (iced tea) is varied in a balanced design over just two physical dimensions (sugar and tea). Individual preference scales are used to parameterize various models to make predictions of preferences on a different set of stimuli within the range of those used in the original parameterization. Such a research design provides a test of the minimal adequacy of various preference models That is, if a model does not perform well in this restricted context, it is even less likely to do so in a more complex marketing application. Furthermore, since the predictor variables (physical levels of sugar and tea) are

known exactly and are independent of one another, it is possible to make stronger statements about the determinants of the preference process under study. In particular, the study compares the predictive effectiveness of models that differ in the following ways:

1. different psychological assumptions as to how stimuli are processed by subjects. For example, additive models are compared with ideal point models.
2. the use of an intervening psychological variable, such as sweetness judgments, between the physical attributes and the preference judgments. These are related to the physical dimensions by psychophysical transforms
3. metric versus nonmetric treatment of the preference scales that serve as input to the models.

In summary, then, the study asks the following question: Given that we know an individual's preferences on one set of stimuli, which models or methods are most effective in predicting preferences on a second set of stimuli within the range of the first set? There are several things that should be noted about this formulation. First, since the no preference information on the validation set of stimuli is used in its prediction, predictive accuracy is not confounded with degrees of freedom. That is, a model that uses additional parameters will only have an advantage if those additional parameters register real information and not just noise. Second, all analysis is done at the level of the individual. Aggregation of preferences across individuals raises all the ambiguities and problems of interpersonal utility comparisons. In particular, insofar as interactions or nonlinearities tend to cancel across subjects, the model best fitting the average

*Joel Huber is Assistant Professor of Marketing, Purdue University.

290

subject may not be, on average, the best model for individual subjects. Thus, in this study, the success of a particular model is not an artifact of aggregation, so that statements can be made which refer to individual subjects rather than to a hypothetical average subject who may not adequately represent any subject. Finally, this study is concerned with the modeling of preference, not product choice. Product choice involves considerations of, among others, price, availability, and shelf position. It seems clear, however, that if a model is not able to predict preferences accurately, then it is not going to be able to predict choice. In this sense, then, the modeling of preferences is logically prior to the modeling of choice.

METHODOLOGY

The stimuli consisted of Lipton iced tea made with different levels of sugar and tea; water, lemon, and coloring are held constant. Respondents were required to make preference judgments on 16 stimuli in the calibration set and on 7 in the validation set. As illustrated in Figure 1, the physical dimensions of the validation set are within the range of the calibration set. The perceptual difference between adjacent stimuli was quite large. That is, subjects sometimes had difficulty determining which stimulus was preferred but almost never had difficulty perceiving the difference between stimuli.

A convenience sample of 22 students recruited from 1973 summer courses at the University of Pennsylvania was used. While such a sample is clearly invalid for the study of the preference levels of any defined universe, it is less so for the study of the preference *process*. For while different demographic groups may differ in their preference for tea mixtures, it is less likely that they will differ over the kinds of models

most appropriate to modeling that preference structure.

Subjects were required to make judgments on pairs of stimuli within each set and on the stimuli individually. The judgments on each pair consisted of (1) a similarity judgment on a five-point scale, (2) an indication of which is preferred, and (3) again on a five-point scale, the strength of that preference difference. The signed-preference differences were converted into one-dimensional preference scales using Bechtel's [2] modification of Scheffé's [12] method of analysis. This methodology assumes that the preference differences correspond to differences on a one-dimensional utility scale. It results in a scale that is determined up to a linear transformation and provides a test of whether the scale is different from one produced by random preference differences. This test indicated significant departures ($p > 0.10$) from a random model for all of the subjects in the study.

After all of the judgments had been made on pairs, each stimulus was taken individually and evaluated in terms of whether more or less sugar or tea should be added to make the sample closer to the subject's ideal. The scores on these psychological attributes for the 7 validation stimuli were not used but were predicted from the scores on the 16 calibration stimuli.

The entire task took from three to four hours. It was spread over short morning and afternoon sessions covering a period of six days. While such a collection procedure might produce invalid scales to the extent that preferences are unstable over time, it has the advantage of minimizing the effects of short-term biases. That is, over the series of sessions, each stimulus appears at different stages of the session and in a different order relative to the other stimuli, with the result that short-term order and learning effects are approximately balanced over the entire task.

For each of the 22 subjects, the following data served as input to the models:

Figure 1
RELATIONSHIP OF 7 VALIDATION STIMULI (●) EMBEDDED
IN 16 CALIBRATION STIMULI (x)

Level of Sugar
(log scale)
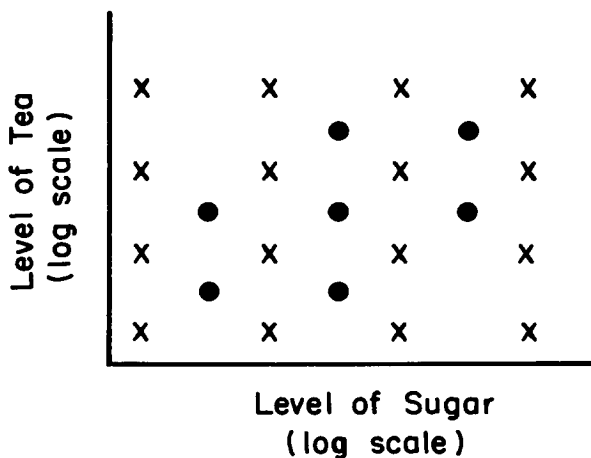
P_i = preference scale for stimuli i, $i = 1,16$ for the calibration stimuli and $i = 17,23$ for the validation stimuli. This scale was formed for each set from preference differences between stimulus pairs;

δ_{ik} = judgment as to the degree to which stimuli i has too much, or too little, sugar ($k = 1$) or tea ($k = 2$). These were coded on an integer scale from -3 to $+3$, negative numbers indicating too little, zero indicating optimum, and positive numbers indicating too much of the ingredient;

x_{ik} = objective level of sugar ($k = 1$) and tea ($k = 2$) for stimuli i;

y_{ik} = level of stimulus i on dimension k of a two-dimensional psychological space formed from proximity judgments on all pairs of the 16 calibration stimulus TORSCA [16] was used to form these individual psychological spaces

MODELS TO PREDICT PREFERENCE

Five models are considered, each reflecting a distinct strategy for predicting preferences on validation stimuli. First, the mathematical and theoretical bases of the models are considered; second, they are compared with respect to predictive effectiveness; and finally the models are evaluated from the perspective of the marketing researcher and practitioner.

The Naive Model

Each of the seven validation stimuli is surrounded on four sides by calibration stimuli for which preference values are known. A simple (minded) estimate of the affective value of the validation stimuli is the arithmetic average of these four calibration stimuli. This is really a "non-model" in that it assumes and explains very little. Essentially, it assumes that proximate points in physical space have preference values that are linearly related within the range of the four points. It does not attempt to reduce the amount of information needed to account for the preference surface. The adjacent points are the parameters of the model; they are all used in estimation so that there are no degrees of freedom for error.

This naive model does have some value, however. It is highly flexible and can be expected to provide reasonable results under a very wide set of assump-

Figure 2
A VISUAL COMPARISON OF THE PREFERENCE SURFACES FORMED BY DIFFERENT MODELS s=SUGAR, t=TEA, and p=PREFERENCE DIMENSION

tions. As such, it provides a benchmark against which the more restrictive models can be judged.

The Additive Model

The additive model assumes that the affective value of a stimulus can be decomposed into utility values for its components which add together to estimate the composite score. Thus, for stimulus i with sugar level j and tea level k the model seeks partworth utility values, a_j and b_k, such that

$$(1) \qquad P_{i(jk)} \cong a_j + b_k \qquad (j, k = 1, 4)$$

is a least squares fit.

There are good theoretical reasons for believing that the main effect additivity assumption is false, i.e., that it is not reasonable to assume that, for example, the contribution of a given level of sugar is the same regardless of the level of tea. But the additive model has flexibility (wrought by 6 degrees of freedom in the present case) that may allow it to closely approximate idiosyncratic preference surfaces even in the face of such expected interaction.

A disadvantage with the additive model is that it is only defined at the component levels represented in the input data. To make predictions on the validation stimuli, the partworth functions are interpolated to produce continuous functions. Panel 1 of Figure 2 provides a visual illustration of the additive model. Notice that there is relatively great freedom in the definition of the marginal effect of, for example, tea. Once determined, however, this cross section must have the same shape regardless of the level of sugar.

Ideal Point in Physical Space

The ideal point model postulates a point, I, with values I_k on physical dimension k such that the distances from I to the stimuli correspond inversely to their preference values. The particular model used allows an orthogonal rotation and differential expansion of the physical space (shown in Figure 1). If T is an orthogonal transformation with elements t_{jk}, the model transforms I and the physical space by:

$$(4) \qquad I_j^* = \sum_k t_{jk} I_k$$

and

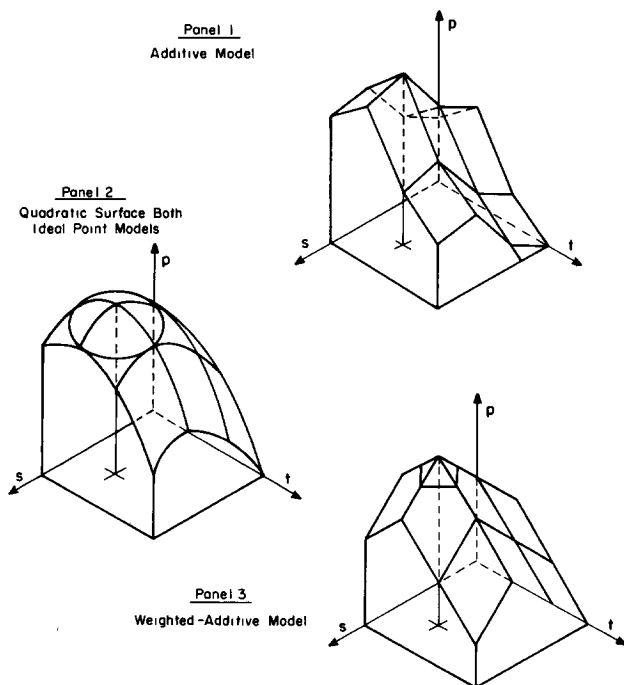$$(5) \qquad x_{ij}^* = \sum_k t_{jk} x_{ik}.$$

The parameterization process essentially determines those values of T and I and weights a_k such that

$$(6) \qquad P_i^2 \cong \sum_k a_k (x_{ik}^* - I_k^*)^2$$

is the best least-squares fit. PREMAP, Phase I [3] is used to parameterize the metric and nonmetric versions of this model.

The ideal point model in physical space is mathematically equivalent to the fitting of a quadratic response surface to the physical space. Geometrically, a representation of the model is given in Panel 2 of Figure 2. The ideal point represents the peak of the quadratic surface while conic sections with axial symmetry through the peak make up the isopreference contours. Generally, these conic sections take the form of ellipses whose axes are not parallel to the original physical dimensions. Except for the restriction of having a single peak, this is a very flexible model having six degrees of freedom in estimation.

Ideal Point in Psychological Space

This two-stage model uses a perceptual space as an intervening variable between the objective space and the preference values. The model fits an ideal point to each subject's two-dimensional psychological space constructed from similarity judgments. Psychophysical transforms are then needed to position the validation stimuli in the derived space and thus predict preferences.

In the first stage, the ideal point is fitted to the psychological space using the same technique as was used to fit points to the physical space except that the model is restricted to the circular isopreference contours of PREMAP, Phase III rather than the general ellipses of Phase I. The justification for this restriction derives from the notion that an individual psychological space already possesses the orientation and dimensionality suitable for modeling that individual's preferences.

The second stage determines the non-orthogonal psychophysical transformation that predicts the psychological space as a function of the physical space. Each of the psychological dimensions is predicted as a linear combination of the known physical dimensions via multiple regression. Thus, the psychophysical transforms determine the orientation and weights of the axes, while the ideal point is positioned by the preference function.

The ideal point in psychological space uses that space as an intervening variable between the physical space and the preference scores. In its two stages, it is fully as general as the ideal point in physical space in the sense of being capable of fitting elliptical isopreferences in any orientation. The difference is that in this two-stage model, the transformation of the space and the fitting of the ideal point are done separately while in the single-stage model they are fitted simultaneously.

The Weighted-Additive Model

The weighted-additive model is a two-stage model using the δ_{ik}'s as intervening variables between the physical dimensions and the preference scores of the stimuli. The first stage predicts preference as a function of subjective sugar and tea (δ_{ik}'s), while the second stage relates subjective sugar and tea to the actual levels of these variables. The second stage is needed to provide estimates of the δ_{ik}'s for the validation stimuli. δ_{ik} is the degree to which stimulus i needs more or less sugar ($k = 1$) or tea ($k = 2$) relative to the subject's ideal. The absolute value of this signed quantity is an estimate of the psychological distance between the stimulus and the ideal along the particular psychological dimension. The model postulates that preference scores are an additive combination of these weighted psychological distances.

The δ_{ik}'s are related to preferences by linear regression. PREFMAP [3] Phase IV determines those weights (a_0, a_1, a_2) such that:

$$(2) \qquad P_i \cong a_0 + a_1 |\delta_{i1}| + a_2 |\delta_{i2}|$$

is a least squares fit.

Since the arguments of the preference function (δ_{ik}'s) are along subjective dimensions, psychophysical transforms are necessary to predict the δ_{ik}'s for the validation stimuli. This is done by assuming that the δ_{ik}'s are a linear combination of the physical dimensions. This assumption is justified if the preference function is single peaked with respect to the physical space. A different transform is necessary for sugar and tea. For particular dimension k it is:

$$(3) \qquad \delta_{ik} \cong b_0 + b_1 x_{i1} + b_2 x_{i2}.$$

Here, x_{i1} is the level of stimulus i along physical dimension 1 (sugar) and x_{i2} is its level along physical dimension 2 (tea). The metric version of PREFMAP Phase IV is used to parameterize (3).

There is an interesting geometric interpretation of this two-stage model. The two psychophysical transforms effectively perform a non-orthogonal transformation of the physical space. The psychophysical transforms define the ideal point and rotate the axes of the space. The preference function (2) then defines the relative importance of these new axes as one moves away from these ideal points. As illustrated in Panel 3 of Figure 2, this model assumes that preference diminishes linearly and symmetrically as one moves from the ideal point. Unlike the additive model, however, interaction between tea and sugar is allowed so that the isopreference contours are parallelograms with any orientation.

PREDICTIVE EFFECTIVENESS

This section assesses the predictive effectiveness of each model by comparing the root mean square across subjects of the correlation between the actual and predicted preference scales. The statistical significance of this difference given any two models is determined by the following procedure: (1) for each subject the difference between the Fisher [5] Z-transform of the correlations is computed; (2) a t-test is then used to test the null hypothesis that the average difference for the 22 subjects is zero. First the effec-

tiveness of the nonmetric options is assessed, followed by an evaluation of the different functional forms.

The Predictive Effectiveness of the Nonmetric Options

With the exception of the naive model, each of the models described is run in its metric and nonmetric formulation. The nonmetric routines allow the preference scales to undergo a monotone transform, one that preserves the order of, but not the intervals between, the scores. The routines generally proceed iteratively by monotonically adjusting the preference scores to be as close as possible to the metric solution and then using these scores to produce a new metric solution. The end result is a model that provides a very good fit to a monotone transform of the original preference scores. While the result is a local optimum, it may, or may not, be a global one. However, since the methods used all begin with the metric solution as the starting configuration, the test of predictive accuracy essentially asks whether the nonmetric adjustments improve prediction within the region of the metric solution.

Since the fit of the nonmetric options is to a monotone transform of the original data, predictions can be expected to be accurate only up to a monotone transform. Accordingly, it is reasonable to hypothesize that nonmetric methods will predict the order of preferences better than metric methods, while the reverse will occur when a linear criterion is used as the index of predictive accuracy. The table compares the metric and nonmetric formulation of the models using product moment and Spearman rank correlations, respectively, as the linear and ordinal measures of predictive accuracy.

The results are consistent and quite clear. Using either the metric or the ordinal criterion, the metric

THE PREDICTIVE EFFECTIVENESS OF THE MODELS IN THEIR METRIC AND NONMETRIC FORMULATIONS: PRODUCT MOMENT (ABOVE) AND SPEARMAN RANK CORRELATION (BELOW)

Model	Root mean square of correlation between predicted and actual		Significance of difference	Correlation of metric to nonmetric
	Metric	Nonmetric	Metric— nonmetric	nonmetric
	0 82	0 79	$p < 05$	0 97
Additive	0 79	0 74	$p < 05$	0 94
Weighted-	0 77	0 74	$p < 05$	0 97
additive	0 74	0 72	$p < 05$	0.97
Ideal point	0 83	0 80	$p < 005$	0 98
in physical	0 79	0 76	$p < 05$	0 97
space				
Ideal point in	0 78	0 77	$p < 10$	0 99
psychological	0 74	0 73	n s	0 97
space				

and nonmetric versions of each of the models are very similar. They have correlations between their respective predictions in the high nineties. However, this small difference in their predictions consistently favors the metric version over the nonmetric version. Furthermore, while the individual statistics may themselves be marginally significant, their conjunction is strongly significant.

Especially surprising is the fact that these results are relatively invariant over the two criteria for accuracy. The reduced significance of the Spearman rank criterion can be attributed to the lesser power of this measure. The differences, however, are quite similar whichever is used.

The high correlation between the two predictions can be explained by the fact that all of the nonmetric routines use the metric solution as the first approximation and then iteratively change the solution to satisfy the ordinal constraints. Given, then, that the preference data form a reasonable interval scale, the arbitrary monotone transform amounts to a loss of information on the preference differences between stimuli. Thus, the relative ineffectiveness of the nonmetric version can be accounted for by the fact that it simply uses less of the available information.

For the marketing manager the implication should be clear. Most of his data could be best classified as being "weakly interval" rather than purely ordinal. That is, there are some monotone transformations that are unreasonable given the data. Consider a 10-point preference scale. A logarithmic or monotone quadratic transformation might be reasonable, but a large step function doubling the utility between scale values "5" and "6" would not. Thus, the intervals do provide information, and it would be ill-advised for a researcher to forsake the robustness and economy of the metric models until further research is able to delineate those situations where nonmetric routines are superior to their metric counterparts. This research would take the form of predictive studies similar to the one presented here, as well as simulations to determine how radical the optimal monotone transform must be so that nonmetric routines outperform metric ones. It is the opinion of this author that the transformations would have to be very radical, indeed, to compensate for the loss of information of the nonmetric version and enable it to compete successfully with the robust metric model.
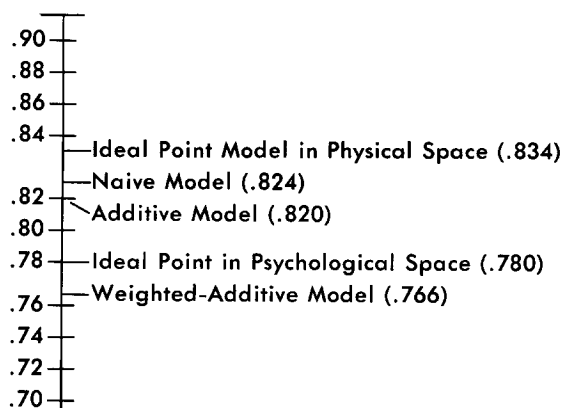
The Predictive Effectiveness of Different Functional Forms

The various functional models can be compared in much the same manner. Since the metric versions have been shown to be superior, only they will be considered. Additionally, since their output should be of interval quality, the product moment correlation of predicted with actual preferences will be the primary criterion of accuracy.

Figure 3
THE PREDICTIVE ACCURACY OF THE METRIC VERSIONS OF FIVE MODELS

Root mean square (22 subjects) of correlation coefficient between actual and predicted preferences.

Figure 3 shows the models on a scale of predictive accuracy. Most of the differences between the models are *not* significant. The three models that do not use intervening variables are each significantly different from the two models that do use intervening variables, but only at a 0.10 alpha level. Certainly, a joint test could be formulated which would show these two groups to be *statistically* different at a greater level of significance, but it is questionable whether such a difference in predictive accuracy is of practical significance.

What does this mean with respect to theoretical conclusions about the process that underlies the preference judgments? Since they cannot be all theoretically correct, must we conclude that they are all incorrect, by approximately the same margins? Although the models are theoretically quite different, comparison of the three panels of Figure 2 indicates that the predicted preference surfaces are not very dissimilar. Perhaps the robustness of these models allows them to approximate preference surfaces in spite of theoretical invalidity. Inspection of the approximated surfaces supports this view.

Figure 4 provides a SYMAP approximation to the preference surface of a typical subject. SYMAP [14] is a computer routine designed to produce geophysical contour maps. It estimates the height (preference) at each point as a function of the level and trend of the surrounding points. The effect is not unlike a two-dimensional exponential smoothing adjusted for trend. In this application, higher numbers reflect greater preference and the lines between the regions approximate can then be considered isopreference contours.

In the example given, there is an ideal region at about the fifth level of sugar comprising all levels of tea below level five. The level of tea does not affect preference except at very low levels of sugar or very high levels of tea. From the shape of the isopreference contours there appears to be an anti-ideal point at high tea and low sugar, but its effect applies only at low levels of sugar. The point here is that no simple model will account for this preference surface. While one might fit in a region of the space, it will not do so globally. This subject is typical of other subjects with respect to such interactions and threshold effects. Furthermore, these effects appear to differ across subjects in such a way as to frustrate a general model to account for them all. One gets the impression of surveying land rather than preferences.

The models can be said, then, to approximate the preference surfaces but not explain them. The choice between models in a practical situation becomes largely pragmatic and situational rather than absolute. In such

Figure 4
SYMAP APPROXIMATION TO PREFERENCE SURFACE FOR SUBJECT NUMBER 1. HIGHER NUMBERS INDICATE GREATER PREFERENCE AND LINES ARE ISOPREFERENCE CONTOURS
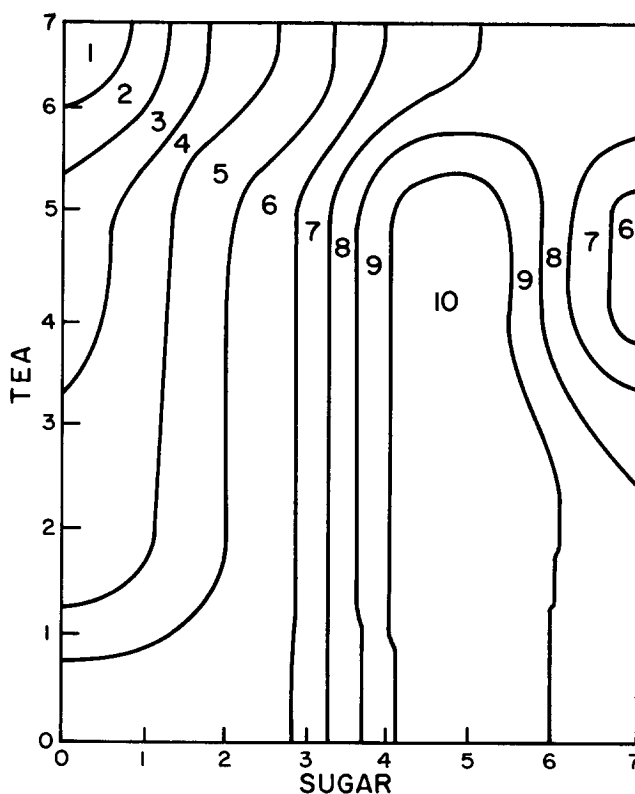
a context, the value of the various models is considered for marketing managers.

Additive Models

The additive model derives partworth contributions for the four levels of sugar and tea and predicts preference as the sum of these. Its predictive power is as good as any of the models in spite of the fact that the expected interaction between sugar and tea implies the inadequacy of an additive representation.

For the marketing practitioner such additive models provide a good way to approximate consumers responses, given a wide variety of underlying processes. Of course, factors must be divided into a relatively small number of levels, and interpolation between these levels might be perilous, but with fractional factorial designs [1, 4, 6, 8] much information can be derived from relatively few data points. What must not be done is to assume from a relatively good fit that the real decision process has been uncovered. Real interactions are often lost in fallible data [11], and the standard tests appear to lack the power to weed them out.

Ideal Point Models

The ideal point models fit a quadratic function to the preference surface, with the ideal point being defined as the modal value of the fitted surface. Although the differences between it and either the additive or the naive model are not significant, there are good reasons for a manager or researcher to use the ideal point model. First, it is a continuous model allowing both interpolation and extrapolation. Second, fewer parameters are needed to represent the preference surface so it is to be preferred by the principle of parsimony. Finally, in situations such as the case presented here, the free orientation of the isopreference ellipses takes account of the expected interaction between the dimensions.

The ideal point model should, however, be used with caution. There are cases where the model simply does not apply, such as when the preference surface is expected to have multiple peaks. It is also possible that the space into which the ideal points are being positioned is missing a critical dimension. This would cause two stimuli positioned closely in the space to have very different preference values. For example, while in the study reported here physical attributes were quite effective in predicting preference for iced tea mixtures, the physical space could be expected to be quite ineffective in predicting preference on *brands* of iced tea. It would be possible, and not unlikely, to find physically *identical* products with very different preference scores.

Two-Stage Models

Two-stage models use a psychological variable as an intervening variable between the physical space and the preference judgments. Their predictive effectiveness is somewhat less than the direct models (0.75 versus 0.80), but this is compensated for by the fact that the intervening variable and the psychophysical transforms reveal a great deal about the way subjects perceive the stimulus set. In the study presented here, estimates of subjective sugar and tea provide evidence for mutual suppression of sugar and tea (see [7]). This could lead to further experiments where the physical dimensions are chosen to minimize interaction, which in the case shown here would lead to the dimensions of "balance" (sugar/tea) and "strength" (sugar × tea). Furthermore, in a world where advertising is of such importance, perceived sweetness might well be more important than actual sweetness. This distinction is far from trivial as the former may be changed by decreasing the level of tea. Thus, it is probably as important that the marketing manager understand how the consumer perceives his product as it is for him to be able to predict what his preference will be with a new combination.

The psychological space formed from proximities judgments holds great promise since the dimensions need not be specified beforehand. This is a necessity in dealing with conceptual data such as brands or package combinations where the physical dimensions are numerous and confounded. Hence, the psychological space can be used to determine which of the perceptual and physical dimensions are salient by using these dimensions to predict the various psychological dimensions. More work, however, needs to be done to validate this methodology. While it works in the study reported here with minimal loss of predictive power, few studies have shown similar results with conceptual data.

To summarize, the various preference models all have valid, if different, functions for the marketing researcher and practitioner. First, the inherent flexibility of the additive model renders it appropriate for approximating response surfaces where the underlying functional is unknown or expected to vary across subjects. Second, the ideal point model, with its intuitive appeal to the marketing manager and parsimonious representation of the response surface, provides a methodology that worked very well here. This model, however, is more sensitive to violations of its assumptions (such as the requirement of a single peak) than is the additive model, and should be used with caution when such uncertainty exists. Third, the two-stage models are valuable where the loss of some predictive power is justified by the need to gain information on the consumer's mediating perceptual mechanism. Finally, across all of these models the metric options were found to produce significantly better predictions than their nonmetric counterparts, indicating that if one has scales that can be classified as "weakly interval," the use of the nonmetric option throws away some of this information.

IMPLICATIONS TO MANAGEMENT AND FUTURE RESEARCH

Suppose a manager collects individual preferences on bundles of physical attributes for a representative sample of an appropriate market segment. The models discussed here provide reasonable ways to interpolate preferences on bundles not in the original sample. But the manager's problem is to determine modifications in an existing product or the composition of a new product that in some way optimally alters group choice in the marketplace.

Since the leap from preferences to choice in the marketplace is very great, a typical response is to formulate a decision rule on the preferences. For example, the manager might search the physical space for a product that maximizes the number of individuals predicted to prefer that product over a competitive brand. This provides no way to predict eventual sales but it does provide some assurance that the product specifications are reasonable for a given marketing strategy.

A more intellectually satisfying way to deal with the same data is to convert individual preferences to individual choice probabilities and then determine expected purchases over the group of individuals. Assuming independence across subjects, probabilities of purchase can be aggregated across subjects in a way that utilities cannot. Pessemier et al. [10] has suggested a way to rescale preferences into choice probabilities. These can then serve as input to comprehensive models [9, 13], which evaluate and search for product bundles. These comprehensive models deal with ideal points in psychological space. The study presented here would suggest that other models might be appropriate to approximate the preference surface, and that the underlying space might in some cases be better represented by physical rather than psychological attributes. In any event, the idea of moving from physical attributes through preferences to choice is very exciting and deserves further study.

These conclusions and recommendations have dealt with the pragmatic and predictive aspects of preference modeling. One disappointment in this study has been the inability to account for preferences in a theoretical way. That is, the models, while providing with varying degrees of accuracy approximations to the preference surfaces, fail to provide a theoretical base or explanation. Preferences can be approximated quite well, but a simple look at an individual's response function should convince most people that a closer approximation would be more complex and less understandable. Perhaps we in marketing are moving in the direction of the physicists with respect to the gas laws. For understanding Boyle's Law, $(PV = nRT)$ works very well. For prediction, however, one must use more and more adjusting coefficients to achieve a desired level of accuracy. One can only hope that this apparent dilemma is a result of our particular perspective and methodology, and that future generations will see our confusion as an artifact of our present methods.

REFERENCES

1. Addelman, Sidney. "Recent Developments in the Design of Factorial Experiments," *Journal of the American Statistical Association,* 67 (March 1972), 103-11.
2. Bechtel, Gordon G. "The Analysis of Variance and Pairwise Scaling," *Psychometrica,* 39 (March 1967), 47-65.
3. Carroll, J. D. "Individual Differences in Multidimensional Scaling," in R. N. Shepard, A. K Romney, and S. Nerlove, eds , *Multidimensional Scaling: Theory and Application in the Behavioral Sciences.* New York: Academic Press, 1972, 105-55.
4. Cochran, W G. and G. M. Cox. *Experimental Designs,* second edition. New York: Wiley, 1966.
5. Fisher, R. A. *Statistical Methods of Research Workers.* Edinburgh: Oliver and Boyd, 1925.
6. Green, Paul E. "On the Analysis of Interactions in Marketing Research Data," *Journal of Marketing Research,* 10 (November 1973), 410-20.
7. Huber, Joel. "The Psychophysics of Taste: Perceptions of Bitterness and Sweetness in Iced Tea," *Proceedings.* Association for Consumer Research, 1973, 166-81
8. Holland, C. W. and David W. Cravens. "Fractional Factorial Experimental Designs in Marketing Research," *Journal of Marketing Research,* 10 (August 1973), 270-6
9. Pessemier, E. A. "Stratop A Model for Designing Effective Product and Communication Strategies," Institute Paper No. 470, Krannert Graduate School of Industrial Administration, Purdue University, 1974.
10. _____ , P. Burger, R. D. Teach, and D. J. Tigert "Using Laboratory Brand Preference Scales to Predict Consumer Brand Purchases," *Management Science,* 17 (February 1971), 371-85.
11. Rorer, L. G. "A Circuitous Route to Bootstrapping," in H. B. Haley, A. G. D'Costa, A. M. Schafer, eds., *Conference on Personality Measurement in Medical Education.* Washington, D C.: Association of American Medical Colleges, 1971.
12. Scheffé, H. "An Analysis of Variance for Paired Comparisons," *Journal of the American Statistical Association,* 47 (September 1952), 381-400.
13 Shocker, Allan D. and V. Srinivasan. "A Consumer-based Methodology for the Identification of New Product Ideas," *Management Science,* 20 (February 1974), 921-37
14 Shepard, Donald S. "SYMAP Interpolation Characteristics," *Computer Mapping as an Aid in Air Pollution Studies,* Volume 2 Cambridge, Mass. Report L, Lab for Computer Graphics, Harvard University, 1970.
15 Wilkie, William L. and Edgar A Pessemier. "Issues on Marketing's Use of Multiattribute Attitude Models," *Journal of Marketing Research,* 11 (May 1974), 164-71.
16. Young, F. W. and W. S Torgerson. "TORSCA, A FORTRAN IV Program for Shepard-Kruskal Multidimensional Scaling Analysis," *Behavioral Science,* 12 (November 1967), 498-9.