1987-01-01

Preference Modelling: Conjoint Analysis and Multi-attribute Models

John Walsh
Cornell Univ

Patrick Roe
Technological University Dublin

Recommended Citation
Walsh, J., Roe, P.: Preference modelling:conjoint analysis and multi-attribute models. Irish Marketing Review, Vol. 2, 1987, pp.126-137.

PREFERENCE MODELLING: CONJOINT ANALYSIS AND MULTI-ATTRIBUTE MODELS

John W. Walsh, Cornell University,
Patrick Roe, College of Marketing & Design, Dublin Institute of Technology

Summary

While many multi-variate techniques have been applied frequently in Ireland, academics and practitioners alike have tended to shy away from incorporating conjoint analysis in their research. Conjoint analysis has been used extensively in other countries primarily to estimate consumers' preferences for products whereas a less complex, easily applied technique, compositional multi-attribute modelling, has been used in Ireland for the same task. This study sets out to compare the powers of estimation or prediction of preference of both techniques. Unresolved areas in preference modelling, namely attribute order bias, learning effects and hetrogeneity of preference structures, are addressed in an attempt to clarify the application and interpretation of both techniques under study. The analysis contradicts assumptions generally accepted and previous research work in relation to these areas. However, conjoint analysis is found to predict more accurately than compositional multi-attribute models. Those contemplating the application of either are advised not to always view them as interchangeable.

Numerous breakthroughs have occurred in market research by the application of techniques developed in other disciplines. Focus group interviews borrow from psychology as do measures of attitudes. Recently, however, statistical techniques have yielded a range of tools commonly termed multi-variate analysis. These analytical tools, as the name implies, consider many variables simultaneously. Therefore, intuitively they would appear appropriate in a market research context. Many multi-variate techniques such as multidimensional scaling, factor, cluster and discriminant analysis have been assessed and used by Irish academics. However, historically, market research practitioners in Ireland have been somewhat reluctant to accept developments, so the proliferation of commercial applications of multi-variate techniques perhaps serves as the most impressive indicator of their value. Despite this, conjoint analysis, a multi-variate technique which is primarily used to estimate and predict an individual's preference when given a choice among a different attribute combinations, has not been used extensively in Ireland to date.

Conjoint Analysis

In essence conjoint analysis may be used to supplement earlier, more simplistic ways of determining which attributes are more important to a consumer when purchasing a product. It may be thought artificial to ask a potential car purchaser, for example, "How important cost, and how important miles per gallon are" in separate questions when in fact he may really be seeking a certain combination/configuration or mix between these and other attributes. Conjoint analysis seeks to do just this, to analyse a range of product attributes conjointly.

Its first application in a marketing context was reported by Green and Rao[1]. What distinguishes conjoint analysis from the Fishbein/Rosenberg class of multi-attribute attitude models is that in conjoint analysis the weights assigned to any attribute are calculated by using statistical techniques whereas in the latter class of models the parameters are explicitly stated by the respondent[2]. Conjoint analysis takes the stated preferences of hypothetical attribute sets as given and then attempts to analyse and explain the implicit weighting system used by the individual to arrive at the stated preferences. Having thus derived the parameters of the model, the model is then used to simulate an individual's preference structure. In contrast, the Fishbein/Rosenberg class of models is confined to existing products and their inherent attribute profiles.

Various techniques have been employed in the past to perform a conjoint analysis and these are briefly described below. Similarly, numerous measures have been developed and tested to evaluate its performance. Basically there are three preference models which may be used in conjoint analysis: the vector, ideal point and part-worth models. The part-worth model was chosen in this study. It specifies a functional relationship between attribute level Y_{jp} and the preference as follows:

$$S_j = f_p\,(Y_{jp})$$

where f_p is the function for the pth attribute.

Basically there are two methods for collecting data to perform a conjoint analysis: a two factor evaluation method and a multi-factor evaluation method. The first approach uses a 'trade-off' procedure in which a respondent chooses between two sets of attributes at any one time. For example, a respondent's ranking of the combinations of the car attributes, top speed and price, might be as follows:

Table 1

	Top Speed		
Price	130 mph	140 mph	70 mph
£2,500	1	2	5
£4,000	3	4	6
£6,000	7	8	9

Other attributes may then be added and preferences expressed between say, price and seating capacity, and price and number of months warranty as illustrated below:

Table 2

	Top Speed			Seat Capacity			Months of Warranty		
	130	100	70	2	4	6	60	12	3
Price									
£2,500	1	2	5	2	1	3	1	3	4
£4,000	3	4	6	5	4	6	2	5	6
£6,000	7	8	9	8	7	9	7	8	9
Top Speed									
130 mph				2	1	3	1	2	5
100 mph				5	4	6	3	4	5
70 mph				8	7	6	7	8	9
Seat Capacity									
2							2	5	8
4							1	4	7
6							3	6	9

From this data, the utility or importance of each attribute level may be computed. While this procedure is relatively easy to administer a number of disadvantages are inherent in this two factor method.

- The respondent is never asked to consider a *full* product profile at any point; they only consider pairs of attributes at a time. This could lead to confusion insofar as a consideration of price and say top speed might as far as the respondent is concerned inherently include seating capacity. However, seating capacity is not included in the evaluation.
- The model soon becomes unwieldy. If, for example, six attributes were being considered with four levels within each, the respondent would need to complete 15 matrices each containing 16 preferences, i.e. 240 in all.
- An underlying assumption also made is that there are no interactions between attributes. In some circumstances this might be unrealistic. For example, high top speed and good cornering ability may, when combined, be more attractive than when each is viewed individually. In this way there may be a Gestalt type 'whole greater than sum of the parts' effect.

The multi-factor evaluation method, pioneered by Green & Roa[3] is probably more widely used and accepted by academic and professional researchers. This method requires the respondent to evaluate all or a number of the attributes at once. For example, in evaluating a TV set a respondent might be asked to consider its price, brand name, size, colour reproduction, guarantee and design all together. The major problem with this method is, of course, information overload and in order to lessen this the number of attributes is usually kept to five or six. Also fractional factorial designs help to overcome this problem to a large extent as a parameter can still be estimated without the respondent having to allocate a score to each possible combination of attribute levels. They provide the respondent with a balanced set of attribute bundles or hypothetical products that is representative of all the possible hypothetical products. In

The Authors

John W. Walsh is pursuing doctoral studies at Cornell University, New York, on receipt of a fellowship from that College. A graduate of the College of Marketing and Design, Dublin Institute of Technology, he holds its Diploma in Advanced Marketing Techniques. He was awarded a First Class Honours B.SC. degree in management from Trinity College, University of Dublin. He has worked as an independent market research consultant and his research interests include preference modelling, multi-variate analysis, price sensitivity and perceived value pricing.

Patrick Roe lectures in Statistics and Quantitative Methods at the College of Marketing and Design, Dublin Institute of Technology. He worked for a number of years in the insurance industry and is presently on sabbathical with the Save & Prosper Group in London. His research interests are in financial and insurance services marketing.

addition, multi-factor evaluation allows interactions to be taken into account as it does not assume that attributes are independent of each other.

In estimating the parameters of a model one must decide on whether to use metric or non-metric scales as the dependent variable in the equation. If metric scales are used a respondent is required to indicate a score for any particular combination of attribute levels, whereas non-metric scaling is used in situations where the respondent merely ranks the combinations without indicating the degree of preference among different attribute bundles. Non-metric models, where the dependent variable is ordinally scaled have used MONANOVA[4], PREFMAP[5], LINMAP[6] and Johnson's non-metric trade-off procedure[7]. In those situations where the dependent variable is intervally scaled ordinary least squares regression has been successfully applied. LOGIT & PROBIT methods relate paired comparison data to a choice probability model. The study described in this article employs ordinary least squares which has been described as a "type of multiple regression with dummy variables"[8]. The dummy variable takes on the value 1 when a particular level of an attribute is present, otherwise 0. As an example, if three hypothetical price levels were £2, £4, and £6, two dummy variables could describe the three alternatives as follows:

Price	X_1	X_2
£2	1	0
£4	0	1
£6	0	0

Therefore, for a full product profile with four attributes, each with three levels, there would be eight dummy variables in the equation. In setting up the regression equation a typical observation could be:

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Concept A =	1	0	1	0	0	1	0	1

The full profile model would thus be expressed as:

$$\hat{Y}_i = B_0 + \sum_{i=1}^{n} B_i X_i$$

where n = the number of dummy variables

X_i = the dummy variables

B_i = the regression coefficients

$\hat{Y}$ = the estimated preference score.

Various studies have shown the predictive ability of ordinary least squares and MONANOVA to be very similar when the rank order has been taken as a pseudo-intervally scaled dependent variable[9, 10, 11]. Reliability and validity tests can be divided into two categories, first, those that test the reliability of the judgements of the respondent and, second, those that test the ability of the model to predict preference patterns when the stimulus is altered in some way, a test for what is called Structural Reliability. The first category is generally judged by a test-retest procedure whereby people's ability to state the same preference after a lapse of time is evaluated. Generally high levels of correlation have been found in these tests[12, 13, 14]. The second category is generally termed Structural Reliability tests and these have, on the whole, also yielded satisfactory results.

Compositional Multi-Attribute Models

Milton J. Rosenberg[15] and Martin Fishbein[16] are generally considered to be the pioneers of multi-attribute models. However, both developed their models for application in psychology. Rosenberg's model was developed from cognitive consistency theory and Fishbein's from behavioural learning theory. As neither model is readily applicable to market research in its original form most marketing applications have altered them in some way[17,18,19]. Out of these applications hybrid models were conceived. Hybrid models vary from application to application yet keep the mathematical form of both the Fishbein and Rosenberg models. An example is given here[20]:

$$A_b = \sum_{i=1}^{n} W_i B_{ib}$$

Where A_b = Attitudes towards a particular alternative b.

W_i = Weight or importance of evaluative criterion i

B_{ib} = Evaluative aspect or belief with respect to ability of alternative b to satisfy evaluative criterion i

n = Number of evaluative criteria important in selection of an alternative in category under consideration

This model comprises, for one alternative, an evaluation element for each salient evaluative criterion and a weight for each of these criteria. The attitude towards an alternative is the summation of each evaluation multiplied by its respective weight. The composition rule of the model is therefore identical to that typically employed in conjoint analysis procedures, being both compensatory and additive. Correlations between actual and predicted perceptions of 0.518 and 0.465 were found by Neslin[21] and correlations ranging from 0.287 to 0.231 were found by Akaah and Korgaonkar[22]. As only five and six attributes respectively were used in these studies the coefficients are extremely low.

Research Objectives

The basic objective of this study was to compare the predictive validity of conjoint analysis with that of the technique commonly used to measure preference in Irish market research, hybrid compositional multi-attribute models. Predictive validity is the term given to the measure of correlation between actual preference and those estimated by a model. As most academic studies reported are based on simulated data, results of predictive ability or the internal validity of the model are all that may be calculated. Obviously, predictive validity is a more rigorous and usually more reliable measure of a model's worth. A subordinate objective of the study was to investigate unresolved areas relating to the application of both above techniques insofar as it was possible without impairing the aforementioned comparison. By attempting to resolve these ideas the application of both research techniques may be rendered less ambiguous and more standardised.

Hypotheses

The following hypotheses were developed from the study objectives. The first hypothesis posits —

> That conjoint analysis has a predictive validity greater than that of compositional multi-attribute models.

However, some issues relating to the application of both techniques remain unresolved. Assumptions about attribute order bias, learning effects and preference structures are made by authors and researchers alike; yet, little proof is offered. It is important that steps be taken to resolve these issues for ease of application and comparision of preference modelling studies. The research required to test the first hypothesis facilitates their study. Acito[23] suggests that the order of attributes biasses their derived importances. It was decided to test this notion. This led to hypothesis two which states —

> That the order an attribute takes influences, either positively or negatively, the importance, derived or objectively stated, of that attribute.

However, should order be considered to affect importance, then surely by the same logic it could affect predictive validity. Consumers are believed to consider a product/service attribute by attribute. In both techniques being tested, respondents are forced to do so. Yet the order in which attributes are presented to respondents need not, and probably does not, resemble the order in which these individuals would evaluate a product/service if left to their own devices. Therefore, the order in which attributes are presented may affect the respondents' evaluation of a product which, in turn, would affect the accuracy of the model. This leads to the third hypothesis which postulates —

> That the order attributes take influences predictive validity.

Learning effects (i.e. a respondent, having completed a task, completes a second task differently from that had he/she not completed the first task, the effect being generally held to produce more accurate or congruent results) are considered to exist even over a six day period[24]. Therefore, they should be considered to exist if a product is immediately followed by another and both techniques relate to the same topic. It is important that learning effects are measured in this study as they may affect the testing of the first hypothesis. One technique may yield a higher predictive validity if it follows another. Therefore, to test for the existence of learning effects is not alone a worthy study in itself but also indicates the reliability of responses being used to test the first hypothesis. The fourth hypothesis thus states —

> That learning effects exists.

Learning effects may operate in either of two ways. First, the respondent may become familiar with the subject as a result of a task and therefore give more learned, congruent or accurate responses to a subsequent task. Alternatively, the respondent may become bored with the topic and concentrate less on the subsequent task thus yielding less learned, incongruent or inaccurate responses. The former is usually held to be the case. Hypothesis five states —

> That learning effects are positive in nature.

Should learning effects be considered to be positive in nature, then logically, the more involved the first task, the more the responses to the second task will be positively influenced. Of the two tasks in this study conjoint analysis is undoubtedly the more complex. Therefore, hypothesis six posits —

> That the learning effect of conjoint analysis on a compositional multi-attribute model is greater than the learning effect of a compositional multi-attribute model on conjoint analysis.

Preference modelling is carried out at the individual level because the structure of preferences is thought to vary from individual to individual[25]. However, very little research has been performed on this issue. Researchers have simply assumed variation in preference structure. The seventh hypothesis, therefore, states —

> That the structures of preferences are similar across individuals.

Methodology

In this study one was faced with a procedure totally the reverse of a typical market research project. Usually the area on which research is to be performed is selected or given and then the appropriate technique nominated. However, the opposite applies here; the techniques were given and an appropriate subject had to be found. Such a subject would influence the performance of both techniques. So it was imperative that this effect would not be biased towards either. The subject of the research finally chosen was lager. Lager consumers were considered to be familiar with its attributes in general and those of specific brands. Consumers were widespread and as such sample wastage would not be a significant problem. It was also felt that lager had sufficient salient evaluative criteria to facilitate a test but not such an excessive amount as to cause information overload. Equally, there were sufficient products on the market to offer a test of predictive validity.

Following consultation with prominent members of the industry, the following six criteria were considered the most salient: taste, price, image, alcohol content, aftertaste, quality. The sample was to be forty students. Although hardly representative of the lager drinking population, they would facilitate 'theory application' research of this kind. "Theory applications are used to test a theory by creating a context and measuring effects with that context that have the potential to disprove or refute the theory"[26]. Any group of respondents can provide a valid test if the theoretical framework under examination is universal in scope[27]. To facilitate the correlation of actual and estimated preferences, six brands on sale in Dublin in draught form were chosen. These were *Fürstenberg, Heineken, Steiger, Harp, Tennents, Carlsberg.*

Conjoint Analysis Procedure: The part-worth model was selected as the preference model to be incorporated in the conjoint analysis procedure primarily because its flexibility would be required in the calculation of attribute utilities. The composition rule for attribute utilities was compensatory and additive. Of the two data collection techniques, multiple factor evaluation and two factor evaluation, the shorter but cerebrally more demanding multi-factor evaluation was chosen. The following attribute levels were decided upon:

Table 3

Price	£1.35	£1.45	£1.55	£1.65
Alcohol content	High	Medium	Low	
Taste	Pleasant	Average	Plain	
Quality	Superior	Good	Average	
Aftertaste	Present	Absent		
Image	Distinctive	Not Distinctive		

A fractional factorial design was employed to obtain a balanced set of concepts that represented four hundred and thirty two possible attribute combinations derivable from these attribute levels. Twenty four hypothetical products, as illustrated in Table 4 were generated and randomly labelled from A - X. Each of these concepts was presented on a 9cm x 14cm card. A non-metric scale (rank ordering) was used although a metric estimation technique was all that was at the authors' disposal. However, such hybridisation is justifiable as metric estimation techniques have been found to perform well, even when non-metrically scaled data has been treated as pseudo-intervally scaled. It is the authors' experience with both academic and commercial applications of conjoint analysis that metric scales are difficult to apply. First, the overall procedure may be too lengthy if the cue cards are ordered before being rated. Secondly, inconsistent preferences or evaluations are commonly found if no ordering procedure takes place before rating, as the respondent is not aware of the inherent parameters of the concept set when the rating starts. In other words, he/she does not know the description of the cards he/she will like most and least while evaluating/rating intermediate cards. In this way ratings may be skewed to either extreme.

Respondents were asked to separate the card set into three groups approximately equal in size — those liked, those not liked and those in-between. Each group was then rank ordered internally. The extremes of consecutive groups were then compared until the rank order of all twenty four cards was agreed upon. Ordinary least squares regression was the estimation technique incorporated. Because rankings served as the dependent variable, the most preferred option was given to the lowest dependent variable score, 1. Therefore, the intercept/base value was typically greater than 24 and the B coefficients negative. The utility scores were yielded by multiplying the B value by -1. Attribute importances were equated to the difference between the largest and smallest utility scores. Dummy variables were assigned binary values; effects coding was not

Table 4
Listing of 24 Hypothetical Products

Concept Name	Price	Taste	Alcohol Content	Quality	Image	Aftertaste
A	£1.35	Poor	Low	Superior	Not Distinctive	Present
B	£1.65	Average	High	Superior	Distinctive	Present
C	£1.35	Poor	Medium	Average	Not Distinctive	Absent
D	£1.45	Pleasant	Low	Average	Not Distinctive	Absent
E	£1.45	Average	Low	Superior	Not Distinctive	Absent
F	£1.55	Poor	Low	Good	Not Distinctive	Absent
G	£1.55	Poor	High	Superior	Not Distinctive	Present
H	£1.45	Average	Low	Average	Distinctive	Absent
I	£1.65	Poor	High	Good	Distinctive	Absent
J	£1.55	Pleasant	Medium	Good	Distinctive	Present
K	£1.55	Pleasant	High	Average	Distinctive	Absent
L	£1.55	Average	Medium	Average	Distinctive	Present
M	£1.65	Pleasant	Low	Good	Not Distinctive	Present
N	£1.65	Poor	Medium	Superior	Distinctive	Absent
O	£1.45	Pleasant	High	Good	Not Distinctive	Present
P	£1.35	Average	Low	Good	Distinctive	Absent
Q	£1.35	Pleasant	Medium	Average	Distinctive	Absent
R	£1.55	Average	Medium	Superior	Not Distinctive	Absent
S	£1.35	Pleasant	High	Superior	Not Distinctive	Present
T	£1.45	Poor	High	Average	Distinctive	Present
U	£1.65	Pleasant	Low	Superior	Distinctive	Absent
V	£1.35	Average	High	Average	Not Distinctive	Present
W	£1.65	Average	Medium	Good	Not Distinctive	Present
X	£1.45	Poor	Medium	Good	Distinctive	Absent

used. These steps yielded utility scores for the designated attribute levels. However, in order to calculate preference scores for the six actual brands each respondent was asked to describe each actual brand in terms of the given attribute levels. These descriptions were then combined with the utility scores for the relevant attribute levels and an overall score of preference found for each actual product.

Hybrid Compositional Multi-Attribute Model Procedure: A five point scale ranging from +2 to -2 was used to measure both attribute desirability and the estimated ability of the given brands to satisfy on each attribute. This was carried out for each actual brand. The respective 'desirability' and 'ability' of each attribute was multiplied and summed for all attributes. This summation represented the preference score. Table 5 illustrates, in part, this approach.

Table 5
Attribute Desirability Rating
Please indicate the desirability of each of the following characteristics to you when choosing a lager:

	Very Undesirable	Undesirable	Neither	Desirable	Very Desirable
Moderate Price	-2	-1	0	+1	+2
Taste	-2	-1	0	+1	+2
High Alcohol Content	-2	-1	0	+1	+2
Quality	-2	-1	0	+1	+2
Image	-2	-1	0	+1	+2
Absence of Aftertaste	-2	-1	0	+1	+2

Each respondent was required to perform both the research techniques. To test for learning effects the sample was divided in two. One half of the group performed the conjoint procedure first and then the hybrid model. The order was reversed for the second group. To test for attribute order effects the sample was again divided in two. One half had an attribute sequence as follows: aftertaste, image, quality, alcohol content, taste, price. The sequence was reversed for the other half. This process of sample segmentation resulted in four independent groups:

- conjoint analysis second — aftertaste first
- conjoint analysis second — aftertaste last

- conjoint analysis first — aftertaste first
- conjoint analysis first — aftertaste last.

Each group comprised ten respondents which was considered sufficient to provide statistically significant results. Before either technique was administered actual rank order preference of each individual respondent was found in relation to each brand.

ANALYSIS

Hypothesis One

Both Kendall's rank correlation coefficient (tau) and Spearman's rank correlation coefficient (rho) utilises the same data to equal results[28] and have been applied equally in preference modelling. The latter was incorporated in this study, however, to measure the correlation between the actual and predicted preference rank for both research techniques. Naturally, as the distribution of the Spearman's coefficients is not normal a Student's t test could not be performed without normalising the data via Fisher's Z_r transformation, which is given by:

$$Z_r = \tfrac{1}{2}\ln(1 + r) - \tfrac{1}{2}\ln(1 - r)$$

where r is the correlation coefficient

Obviously, when $r = 1$, $Z_r = \infty$ so an arbitrary Z_r value had to be used when perfect correlation was found. A test of the difference between two means for correlated samples was carried out on the transformed data. The t value was found to be 3.50596 which is significant at the 0.005 level. However, it may be argued that this is not an adequate test, i.e., that a more appropriate test would be to compare only the prediction of first preferences. Studies to date have tended not to concentrate on first preferences alone. Nonetheless, it may indeed be possible that conjoint analysis only predicts more accurately in the lower or intermediate ranks of preference. A chi-square test was performed on the first data alone resulting in $X^2 = 4.0125$ which is significant at the 0.05 level. Therefore, under both tests the hypothesis may be accepted.

It is the authors' opinion that the difference in the modelling or predictive powers of the two techniques lies in the greater ability of conjoint analysis to predict intermediate preferences. Respondents, generally, are quite clear as to their first, second, and last preferences. Both techniques find these easiest to predict. However, conjoint analysis facilitates the prediction of the order of other preferences as it acquires more information on the structure of preferences.

The Spearman's rhos for the correlation between both techniques were quite low. The predictions of conjoint analysis correlated to a higher degree with the actual ranks than with the hybrid compositional multi-attribute model predictions. This would appear to support the notion that respondents give conflicting information on their preference structure to both techniques. The absolute and relative attribute importances of both techniques should be compared for each individual in order to test this.

Hypotheses Two and Three

If order effects were to exist, the importance of the attributes when presented first would be significantly different to the importance when presented last. It need not, of course, be a positive effect. As 'price' and 'aftertaste' were at the extremes of the presentation sequence, X^2 tests were performed on their importance scores. It was found that the importance scores of 'price' and 'aftertaste' did not differ significantly due to the order or presentation for either technique. The hypothesis was, therefore, rejected.

A test of the significance of the difference between two means for independent samples was used, in this instance, on the Fisher's Z_r transformations of the Spearman's rhos. The two samples were, 1) those with aftertaste presented first and 2) those with price presented first. A t-test gave $t = -0.07191$, which is not significant at the 0.05 level even with 78 degrees of freedom. Using only first preference data to test the same hypothesis yielded $X^2 = 1.0031$. This is not significant at the 0.05 level either. Therefore, hypothesis three was rejected.

The test for order effects was comprehensive and conclusive in this study. They were found to affect neither attribute importance nor predictive validity. However, it may be logical to consider product familiarity to play a role in the effect of order of attributes. Should an individual be unfamiliar with the product and the attributes that pertain to it, he/she should be influenced more by order effects. Should products be put on a continuum of familiarity to their consumers, lager would not approach either extreme. Given this, we may assume that for products quite familiar to respondents, such as industrial goods, order effects would be negligible or non-existant. Further study of products and respondents with varying degrees of familiarity could prove fruitful in this respect.

As a result of this finding, less hesitancy may be expected in comparing studies in which the order of attributes is different. Equally, the validity of studies in which the order of attributes has not been varied across respondents need not be questioned automatically. In observing respondents performing evaluation tasks in both research techniques, the authors noted that some respondents did not tackle the task in the order in which the attributes were presented. Instead, the fourth attribute for example was dealt with first, then the sixth, then the third and so forth, each individual apparently following an order unique to himself/herself. This may indicate that an individual is unaffected by the order in which attributes are presented to him/her as he/she evaluates in the order he/she wishes. Alternatively, one could argue that those respondents evaluating in the order in which attributes were presented had this order forced upon them. The validity of their result could thus be said to be reduced. However, this was not the case. Predictive validity was found not to vary with variations in the order of presentation of attributes. This may indicate either that individuals evaluate in the order they wish or that in evaluation the order of attributes is irrelevant to predictive validity. Although the former would appear intuitively more attractive, both suggest that order of *presentation* is of little importance.

Hypothesis Four, Five and Six

Two approaches could be applied here. A test of the significance of the difference between two means for correlated samples could be performed. This would require combining the data of those techniques performed first and comparing it with the combined data of those techniques performed last. The advantage of this analysis is that the compared samples would be of maximum size (i.e. 40), given the total sample size (i.e. 40). However, combining the data in this case need not necessarily be correct. In testing hypothesis three the combination of results from both research techniques was permissible because order effects are not considered to vary across the techniques. However, it is plausible to suppose that the learning effects of each technique on the other are not equal. Should this be the case, it may be argued that it is technically more appropriate to perform two tests, to compare the data of one research technique when performed first with the data of that same technique when performed second. These tests would be tests of the difference between two means for independent samples. They would, however, be less powerful than the singular test mentioned above, as the compared samples would be smaller (i.e. 20). However, this three test approach should improve the reliability of any results.

Test 1: The average of the transformed correlation coefficients of those techniques performed first (i.e. 20 conjoint and 20 hybrid) was compared with the average of the transformed correlation coefficients of those techniques performed second.

Test 2: The result from the hybrid data when it was performed first was compared with the result of the hybrid data when it was performed second.

Test 3: The conjoint results when the conjoint was performed first were compared with the conjoint results when it was performed last.

The t values for the three tests were -0.0539, -0.5201 and 0.2729 respectively, each too low to be significant. Hence, the hypothesis was rejected. Thus, learning effects were not found. As a consequence hypothesis five and six must also be rejected.

However, two types of learning may take place. Firstly, the respondent may learn about his/her preferences and structure of preferences. Secondly, he/she may learn about the research techniques. These two forms of learning are obviously very distinct. Only the former was tested here. Perhaps market researchers do an injustice to respondents in presuming that they may effectively learn about their preferences solely as a result of a half hour study. Indeed, it may be a pretentious notion that by merely providing an individual with tasks to perform he/she is being taught about his/her preferences. It is the authors' opinion that respondents are far more capable and informed than is generally considered. The results of this study support this view. Respondents did not learn about their preferences or preference structure.

Hypothesis Seven

The order of the concept cue cards was used to provide an indicator of respondents' personal preference structure. Kendall's coefficient of concordance (W) was used to measure the degrees of association among these ranks. The W score was found to be 0.3749. This is quite a low degree of association. However, it is significant at the 0.0001 level. Siegel[29] states that highly significant values of W indicate that the same standards have been applied in ranking the objects. In this case, it would indicate similar preference structures among the respondents. Therefore, the hypothesis may be accepted. Thus it was found that the structures of preferences were similar across individuals.

The belief that preference structures vary from individual to individual has been generally held by academics and researchers alike. However, the opposite has been proven in this study with a high level of confidence. Therefore, the acceptance of individualism of preference structures may be questioned. The most important application of this is for research design. A representative sub-sample of the total sample may be selected in preference modelling studies, and their average attribute importances may be universally applied to the total sample. Obviously, much time effort and expense would be saved. Equally, studies could become more attractive to respondents because of the reduction in the time requirements. What would distinguish individuals, therefore, would be their evaluation of actual products. In observation of respondents, the authors noted that evaluations were quite divergent across respondents.

Conclusion

First, it was found that conjoint analysis predicts preferences more accurately than the hybrid compositional multi-attribute model. Great variation in time, effort and expense was found between these techniques and as such they should not be considered interchangeable in market research. Further research needs to be executed to reveal those areas in which the difference in predictive validity is significant and those in which compositional self-explicative models suffice.

Second, the order in which attributes were *presented* were found not to influence predictive validity or attribute importances. Whether the order in which attributes are *evaluated* affects predictive validity or attribute importance must be determined by further research.

Third, respondents were found not to learn of their preferences/preference structures by performing research tasks. More accurate results may be found if the respondent becomes more efficient with the research technique through learning, but this was not studied here.

Fourth, preference structures were found not to vary substantially across individuals. In order words respondents viewed attribute importances similarly. Intuitively, this rejects the notion of market segmentation using utility functions as a base. However further research could confirm this theory or prove it to be an isolated case.

References

1. Paul E. Green and Vithala Rao, "Conjoint measurement for quantifying judgmental data", *Journal of Marketing Research*, 8, 1971, pp. 355-363.

2. William L. Wilkie and Edgar A. Pessemier, "Issues in marketing's use of multi-attribute attitude models", *Journal of Marketing Research*, 10, 1973, pp. 428-441.

3. See Paul E. Green and Vithala Rao, *loc. cit.*

4. J.B. Kruskal, "Analysis of factorial experiments by estimating monotone transformations of the data", *Journal of the Royal Statistical Society*, Series 3, 27, 1965, pp. 251-263.

5. J. Douglas Carroll, "Individual differences and multidimensional scaling", in R.N. Shepard *et al* eds., *Multidimensional Scaling: Theory and Applications in Behavioural Sciences*, Vol 1, New York, Seminar Press, 1972 pp. 105-155.

6. Venkataraman Srinivasan and Allan D. Shocker, "Linear programming techniques for multidimensional analysis of preferences", *Psychometrika*, 38, 1973 pp. 337-369.

7. Richard M. Johnson, "Varieties of conjoint measurement", Working Paper, Market Facts Inc., Chicago, 1973, cited in Paul E. Green and Venkataraman S. Srinivasan, "Conjoint analysis in consumer research: issues and outlook", *Journal of Consumer Research* 1978, pp. 103-123.

8. Paul E. Green and Donald S. Tull, *Research for Marketing Decisions*, 4th ed., Englewood Cliffs, New Jersey, Prentice Hall Inc., 1978, p. 31.

9. Frank J. Carmone, Paul E. Green and Arun, K. Jain, "Robustness of conjoint analysis: some Monté Carlo results", *Journal of Marketing Research*, 15, 1978, pp. 300-303.

10. See Philippe Cattin and Dick R. Wittink, "A Monté Carlo study of metric and non-metric estimation methods for multi-attribute models", Research Paper No. 341, Graduate School of Business, Stanford University, 1976.

11. Dick R. Wittnik and Philippe Cattin, "Alternative estimation methods for conjoint analysis: a Monté Carlo study", *Journal of Marketing Research*, 18, 1981, pp. 101-106.

12. R. Kenneth Teas, "An analysis of the temporal stability and structural reliability of metric conjoint analysis procedures", *Journal of the Academy of Marketing Science*, 13, 1, 1985, pp. 122-142.

13. Barnett R. Parker and Venkataraman Srinivasan, "A consumer preference approach to the planning of rural primary health care facilities", *Operations Research*, 24, 1976, pp. 991-1025.

14. Franklin Acito, "An investigation of some data collection issues in conjoint measurement", in 1977 Educators' Proceedings, Chicago, American Marketing Association, 1977, pp. 82-85.

15. Milton J. Rosenberg, "Cognitive structure and attitudinal effect", *Journal of Abnormal and Social Psychology*, 53, 1956, pp. 367-372.

16. Martin Fishbein, "A behaviour theory approach to the relations between beliefs about an object and the attitude towards the object", in M. Fishbein, ed., *Readings in Attitude and Theory Measurement*, New York, John Wiley & Sons Inc., 1967, pp. 388-399.

17. Richard J. Lutz and James R. Bettman, "Multi-attribute models in marketing: a bicentennial review", in Arch G. Woodside et al, eds., *Consumer and Industrial Buying Behaviour* New York, North-Holland, 1977, pp. 137-149.

18. Fleming Hansen, "Consumer choice behaviour, an experimental approach", *Journal of Marketing Research*, 4, 1969, pp. 436-443.

19. See William L. Wilkie and Edgar A. Pessemier, *loc. cit.*

20. James F. Engel and Roger D. Blackwell, *Consumer Behaviour*, Hinsdale, Illinois, The Dryen Press, 1982, 4th ed., p. 447.

21. Scott A. Neslin, "Linking product features to perceptions: self-stated versus statistically revealed importance weights", *Journal of Marketing Research*, 18, 1981, pp. 80-86.

22. Ishmael P. Akaah and Pradeep K. Korgaonkar, "An empirical comparison of the predictive validity of self-explicated, huber-hybrid, traditional conjoint and hybrid conjoint models", *Journal of Marketing Research*, 20, 1983, pp. 187.197.

23. Franklin Acito, *op. cit.*, p. 85.

24. J. McCullough and R. Best, "Conjoint measurement, temporal stability and structural reliability", *Journal of Marketing Research*, 16, 1979 pp. 26-31.

25. Paul E. Green and Venkataram Srinivasan, "Conjoint analysis in consumer research: issues and outlook", *Journal of Consumer Research*, 1978, p. 104.

26. For a justification of using unrepresentative samples in 'theory application' research see Bobby J. Calder, L.W. Phillips and A.M. Tybout, "Designing research for application", *Journal of Consumer Research*, 8, 1981, pp. 197-207.

27. Arie Kruglanski, "Much ado about the 'volunteer artifacts'", *Journal of Personality and Social Psychology*, 28, 1973, pp. 348-354.

28. Sidney Siegel, *Non-Parametric Statistics for the Behavioural Sciences*, Tokyo, McGraw-Hill Kogkusha Ltd., 1956, p. 219.

29. *Ibid.*, p. 237.