ATTRIBUTE PRESENTATION ORDER BIAS AND NONSTATIONARITY IN FULL PROFILE CONJOINT ANALYSIS TASKS

Randall G. Chapman, University of Alberta
Ruth N. Bolton, University of Alberta

ABSTRACT

This paper describes an investigation of the incidence of attribute presentation order bias and nonstationarity in full profile conjoint analysis rating tasks. Conjoint tasks were administered with a microcomputer to permit randomization of attribute presentation order and serial profile presentation order across respondents. The results indicate that the attributes' order of presentation had a statistically significant effect on respondents' ratings of the profiles and on respondents' attribute relative importances. In particular, a pattern of recency effects was detected, where the attribute presented last had a disproportionately large influence on preference judgments. The actual size of the attribute order bias effect was not very substantial -- about 2%-5% -- when expressed in terms of the estimated attribute relative importances. The results suggest that nonstationarity is a potential problem in full profile conjoint rating tasks. Both novelty and fatigue effects were detected, with 32% of respondents exhibiting one of these nonstationary forms of behavior.

INTRODUCTION

Marketing scientists have devoted considerable attention to exploring a range of strategic and tactical issues associated with the conduct of conjoint analysis studies. Green and Srinivasan (1978) provide a detailed menu of relevant design and estimation issues. The scale of these research efforts may be attributed to the perceived value of conjoint analysis as a marketing research tool in product/service design, to the commercial application intensity of conjoint analysis (Cattin and Wittink 1982), and to the centrality of preference measurement in the study of buyer behavior.

This paper focuses on investigating the possible presence of attribute presentation order bias and nonstationarity in full profile conjoint analysis rating tasks. These two methodological concerns have received little attention in the conjoint analysis literature. Two specific questions are studied in this paper:

<u>Attribute Presentation Order Bias</u>: Does the order in which attributes are presented in full profile conjoint rating tasks influence the results obtained?

<u>Nonstationarity</u>: Do respondents to full profile conjoint rating tasks use the same judgment rules throughout the entire evaluation task?

These questions relate to the underlying assumptions that are implicitly made by conjoint analysis researchers. This paper seeks to test these assumptions empirically.

The existence of order bias and nonstationarity would have considerable significance for full profile conjoint research applications. Attribute presentation order bias would manifest itself in the following way. Attributes in certain positions, such as first or last, would receive extra attention and weight because of their position, and not due to their intrinsic importance to a decision maker. If respondents' judgments in conjoint tasks are influenced by the order of presentation of the attributes, then randomization tactics would have to be employed by conjoint researchers. This would make it very difficult to continue to use standard paper-oriented questionnaire formats, since too many different randomized orders would be required. Simple rotation schemes are unlikely to be successful in completely alleviating this kind of presentation order bias (Blunch 1984). If stationarity does not exist, a major underlying assumption of conjoint analysis -- that respondents employ the same decision rule throughout the task -- would be violated. With nonstationarity present, the researcher would be confronted with multiple sets of utility model weights, and a resolution as to which "best" describes the respondent's behavior would be required. Thus, these are important issues of interest to researchers involved in preference measurement.

The remainder of this paper is organized as follows. The next section reviews the published literature to provide the necessary theoretical background to the issues investigated in this paper. The following section presents the research methodology used in this study. Then, attribute presentation order bias and nonstationarity test results are described and interpreted. Some concluding remarks complete this paper.

THEORETICAL DEVELOPMENT

The existence of order bias effects is well established in the questionnaire design literature (c.f. Sudman and Bradburn 1982). These effects are more likely to occur in complex judgment tasks. A respondent may utilize a heuristic, such as a rule relying on attribute presentation order, to simplify the judgment task (Bettman 1979). For example, a respondent may pay more attention -- and, consequently, attach more importance -- to the attribute presented first (a primacy effect) or last (a recency effect). Attribute presentation order effects are characterized by the respondent weighing an attribute in a given presentation order more heavily. As a result, a respondent's judgments about conjoint profiles may reflect not only the attribute levels in a profile, but also the task characteristics.

In this paper, nonstationarity is defined to occur when a respondent's preference explication process changes during the conjoint task. A respondent's decision rule may exhibit temporal instability due to task novelty or fatigue effects. Novelty effects reflect the respondent's difficulty in initially adjusting to the unfamiliar nature of the conjoint judgment task. These effects may occur during the initial stages of a conjoint task. Fatigue effects reflect a respondent's difficulty in making a large number of repetitive judgments in a short period of time. These effects may occur during the latter stages of a conjoint task. In either situation, the respondent may choose to utilize a simplifying heuristic which does not accurately reflect his/her preferences. For example, a respondent may weigh a specific attribute or a specific presentation order position more heavily during one stage of the conjoint task. Or, a respondent may assign higher (or lower) ratings to the profiles presented early in the conjoint task. (Note that the first heuristic is more complex than the second and, consequently, more difficult to detect.) These serial profile presentation order effects would imply the existence of nonstationarity or temporal instability.

Given the practical importance of order bias in conjoint analysis, it is surprising to find that so little attention has been devoted to it in the literature. Only two order bias studies appear to exist. Acito (1977) and Johnson (1982) both found that relative importance is influenced by attribute presentation order, and specifically that attributes appearing first on a full profile stimulus card tend to have higher relative importances. The literature with regard to nonstationarity is even more sparse. In an information processing assessment of conjoint decision making behavior, Olshavsky and Acito (1980) detected some evidence of shifts in choice rules across two different conjoint card-sorting tasks. These studies involved relatively small numbers of respondents, simple conjoint contexts, and limited efforts to completely study order biases and nonstationarity (since, in each case, the order bias/nonstationarity effects were by-product "findings" and not the main focus of the researchers' efforts).

To summarize this discussion, the traditional conjoint model represents a respondent's profile rating, or overall evaluative judgment, as a function of attribute levels. If there are order bias effects, the measurement model would have to be expanded as follows:

RATING = f(Attribute-Level Effects, Attribute Order Position Effects, Serial Profile Presentation Effects).

Thus, ratings would be seen to depend on the usual attribute-level effects, but also (potentially) on attribute presentation order and serial profile presentation order. This study will investigate whether ratings are affected by attribute order and serial profile presentation order, and measure the size of these two order biases.

RESEARCH METHODOLOGY

The goal of this paper is to conduct a large-scale thorough investigation of order bias and nonstationarity. The study utilizes a substantial number of respondents to benefit from the generalizability advantages inherent in large samples. Order bias and nonstationarity are examined in four, five, and six attribute tasks, within the same conjoint context. This research design was chosen to attempt to avoid the possibility that the findings might be idiosyncratic to a specific degree of task complexity (as represented by number of attributes to evaluate).

The research methodology adopted in this study involved choosing a suitable product/service context for a convenience sample of respondents. The convenience sample of respondents, students, dictated the conjoint context, since the context had to be meaningful to the decision makers. The subjects in this experiment were undergraduate and MBA students in marketing management and marketing research courses at the University of Alberta. The students were not told anything about the experiment -- other than the fact that it would involve the completion of a questionnaire regarding their preferences for a new product/service -- prior to completing the questionnaire task. The conjoint analysis context used was long-distance discount telephone services. These services do not yet exist in Canada (at the time of this study), although they are plentiful in the United States.

Based on existing services of this kind in the United States, the following six attributes were employed in this conjoint analysis task:

- Discount From Regular Telephone System Daytime Rates
- Hours of Operation During Which the Service is Available
- Monthly Fixed Service Charge (in Addition to Long Distance Calls)
- Extent of Access
- Coverage in Canada and the U.S.
- Telephone Type Required to Use the Service.

The specific levels associated with each of these six attributes are displayed in Exhibit 1. To provide for some generalizability of the findings, four, five, and six attribute versions of the same product/service evaluation task were developed. Appropriate experimental designs were employed so that all respondents faced the same number of full profile judgments. For the four attribute task, a full factorial design was used. For the five and six attribute tasks, Latin Square designs were used to retain the same number of profiles -- 24 -- for each respondent.

The ability to administer questionnaire tasks on microcomputers makes it possible to examine the existence of attribute presentation order bias and nonstationarity in conjoint analysis tasks. Microcomputer administration permits the convenient possibility of randomization of both attribute and profile presentation order across respondents, both essential components of an empirical study to investigate order bias and nonstationarity effects.

A recently developed interactive conjoint analysis administration program, CJ (Chapman and Cochrane 1985), was used in this study. Students arrived at a computer lab containing 12 IBM PCs. The conjoint analysis questionnaire administration programs had already been allocated in a random fashion to each of the 12 machines. Four machines had each of the three different tasks. Upon entering the computer lab, students were free to choose any available microcomputer. Hence, students were randomly assigned to the three different attribute tasks by their seating position. A total of 60, 45, and 59 students completed the four, five, and six attribute tasks, respectively.

So that all attribute-levels might be available at all times during the judgment task, respondents were provided with an instruction sheet to which they could refer throughout the questionnaire administration. This instruction sheet listed all attribute-level possibilities.

The study participants were guided through the data collection task by a specially designed program developed for use on personal computers. This program randomizes the

EXHIBIT 1

Attributes and Levels in the Conjoint Analysis Study

DISCOUNT (Discount From Regular Telephone System Daytime Rates)
- One-Third
- One-Half
- Two-Thirds

OPERATION HOURS (Hours During Which Service is Available)
- All the Time
- 5PM-8AM Only

FIXED CHARGE (Monthly Charge in Addition to the Cost of Long Distance Calls)
- No Charge
- $10 Per Month

EXTENT OF ACCESS
- Everywhere (All Telephone Numbers)
- Major Cities (Telephone Numbers in Major Cities Only)

CANADA/U.S.? (Coverage in Canada Only or Canada and U.S.)
- Canada Only
- Canada & U.S.

TELEPHONE TYPE (Required Telephone Equipment)
- Pushbutton (Pushbutton Telephone is Required to Use the Service)
- Any Phone (The Service May Be Used With Any Telephone)

order of presentation of the attributes across respondents. (However, all profiles presented to a given respondent had a fixed attribute order.) This program also randomizes the serial presentation order of the 24 profiles across respondents.

Respondents evaluated each profile on a 0-100% probability-of-purchase rating scale. The end points of this scale were anchored by the phrases "no chance" and "certain." Intermediate 10%-point increments were described in terms of the phrases "very slight possibility," "slight possibility," "some possibility," "fair possibility," "fairly good possibility," "good possibility," "probably," "very probably," and "almost sure."

Post-administration in-class debriefings of the respondent groups did not reveal any systematic difficulties either with the conjoint task context or with the computer program used to administer the conjoint tasks.

EMPIRICAL RESULTS

This section begins by displaying some descriptive statistics to characterize the basic conjoint model results. Then, the attribute presentation order bias phenomenon and the nonstationarity problem are examined. Since the order of presentation of the attributes and the profiles were randomized across respondents, the potential influence of these effects on the decision makers' ratings can be measured separately. Nonstationarity, or serial profiles presentation order bias, is examined at the aggregate and individual levels.

Preliminary Results

The average part-worth utility weights for the respondents who completed each of the four, five, and six attribute tasks are displayed in Table 1. These weights were estimated via OLS for each respondent. (The weights were estimated for each respondent separately; then, average weights were formed as the arithmetic means of the individual respondents' weights.) An additive main-effects model form was assumed. The part-worth utility model weights were re-scaled to lie in the 0-1 interval for the purposes of interpretation.

TABLE 1

Average Estimated Part-Worth Utilities

Attributes and Levels	Four Attribute Task	Five Attribute Task	Six Attribute Task
DISCOUNT:			
One-Third	0.146	0.197	0.076
One-Half	0.604	0.486	0.612
Two-Thirds	0.755	0.816	0.844
OPERATION HOURS:			
All the Time	0.666	0.617	0.884
5PM-8AM Only	0.338	0.383	0.137
FIXED CHARGE:			
No Charge	0.988	0.995	0.980
$10 Per Month	0.016	0.005	0.042
EXTENT OF ACCESS:			
Everywhere	0.799	0.720	0.871
Major Cities	0.205	0.280	0.151
CANADA/U.S.?:			
Canada Only		0.506	0.397
Canada & U.S.		0.494	0.625
TELEPHONE TYPE:			
Pushbutton			0.430
Any Phone			0.592
Sample Size	60	45	59

The results appear to pass some very basic tests of face validity. With regard to long-distance discount telephone services, these respondents prefer a larger discount to a smaller discount, more operation hours to less, no fixed charge, complete access to all telephones, extensive geographical coverage, and no restrictions on equipment. The one anomalous result is for the five attribute task where, on average, respondents slightly prefer access to "Canada Only" as opposed to "Canada & U.S." The two least important attributes were the fifth ("Canada/U.S.?") and the sixth ("Telephone Type") used in this study. Thus, the omitted attributes (selected on the basis of preliminary exploratory research) in the four and five attribute tasks appear to have been correctly chosen because the omission of determinant attributes would have had potentially serious consequences.

With regard to statistical fit of the individual regression models, the average R^2 values were 0.75, 0.73, and 0.63 for the four, five, and six attribute tasks, respectively. It is interesting to note the generally satisfactory levels of these values, although there is a detectable drop in fit from the four and five attribute tasks to the six attribute task. Apparently, task complexity increases quite substantially after five attributes. Based on the pattern of these statistical fit levels, it is possible to hypothesize an additional drop in average statistical fit for a similarly constructed seven attribute task. In addition, one might also hypothesize that the length of the conjoint task also influences statistical fit. The 24 profile tasks used here probably represent about the limit to which respondents may be pushed in full profile conjoint rating/ranking tasks before refusals, early terminations, and very high levels of nonstationarity appear with great regularity.

Attribute Presentation Order Bias

This experiment is a between-subjects design, in which a subject is exposed to a single attribute presentation order and a single serial profile presentation order. While a within-subjects design (involving complete re-randomization of attribute presentation order across profiles presented to a given respondent) would have been preferable, a between-subjects design is necessary to reduce potential respondent confusion. Thus, it is not possible to isolate a separate respondent effect in this study.

If the ratings are not influenced by the attribute presentation order, then the mean ratings associated with each order position of each attribute should be identical, except for random sampling variation. One-way ANOVAs were conducted to test the hypothesis that all mean ratings are equal for the order position of each attribute. The results of these ANOVA assessments are displayed in Table 2. (Each row of Table 2 represents a separate one-way ANOVA.) In general, the null hypotheses of equal mean ratings across various presentation order positions are rejected at conventional levels of statistical significance -- all F statistics for the five and six attribute tasks are statistically significant at the 0.01 level, while two of the four tests for the four attribute case are statistically significant at the 0.01 level. This finding is consistent with the notion that as the task complexity increases, respondents are more susceptible to attribute order bias effects.

TABLE 2

ANOVA Results For Attribute Presentation Order Bias Analysis

Task & Attributes	Sample Size	F-Test Stat.	Calculated Significance Level of This F Statistic Value
Four Attribute Task			
Attr. #1: Discount	1440	8.43	0.000
Attr. #2: Operation Hours	1440	2.09	0.100
Attr. #3: Fixed Charge	1440	15.02	0.000
Attr. #4: Extent of Access	1440	0.51	0.679
Five Attribute Task			
Attr. #1: Discount	1080	14.22	0.000
Attr. #2: Operation Hours	1080	14.18	0.000
Attr. #3: Fixed Charge	1080	16.62	0.000
Attr. #4: Extent of Access	1080	6.88	0.000
Attr. #5: Canada/U.S.?	1080	5.61	0.000
Six Attribute Task			
Attr. #1: Discount	1416	18.73	0.000
Attr. #2: Operation Hours	1416	8.81	0.000
Attr. #3: Fixed Charge	1416	3.58	0.003
Attr. #4: Extent of Access	1416	3.30	0.006
Attr. #5: Canada/U.S.?	1416	6.58	0.000
Attr. #6: Telephone Type	1416	7.56	0.000

Notes: (1) Each respondent was presented with 24 profiles, so the actual number of respondents corresponding to each sample size equals the reported sample size divided by 24.

The average (across respondents) conjoint model relative importances for the various attributes in the four, five, and six attribute tasks are reported in Table 3. These relative importances are the usual normalized ranges of the estimated part-worth utility weights. If there is no attribute presentation order bias, the relative importances should be the same across all positions (i.e., columns).

If no order bias is present, the sum of the relative importances across all attributes for a given order position and task should be 100.0%. If respondents are weighting attributes in a given order position and task more heavily, the sum will be greater than 100.0%. The one clear result apparent here is a recency effect. Attributes listed in last position receive a greater weight than in other positions. However, the magnitude of the difference is not that large. The total relative importances for the last position attributes sum to 102.0%, 104.0%, and 104.8%, respectively, for the four, five, and six attribute tasks. This suggests that last position attributes receive about 2%-5% extra importance, not an overly troublesome amount relative to sampling variation. It is interesting to note that the total relative importances across the various attributes and tasks all lie in the range 97.1%-104.8%, thus supporting the assessment that the order bias effect is not very large.

There may be a primacy effect, as well as a recency effect, in the four attribute task. However, this effect is very small and may be due to sampling variation.

Nonstationarity

If decision makers use the same judgment process (the same part-worth utility weights) throughout a conjoint questionnaire task, then they adhere to the temporal stability assumption inherent in all conjoint analyses. Two violations in the assumption of temporal stability were investigated. First, systematic nonstationarity across respondents was examined by testing for changes in mean ratings across profile order positions. Then, nonstationarity unique to individual respondents was examined by testing for changes in attribute importance weights across profile order positions.

The first test for serial profile presentation order bias investigated systematic changes in profile ratings across respondents. Profiles were characterized as occuring in one of three order positions: the first one-third (profiles 1-8), the second one-third (profiles 9-16), or the last one-third (profiles 17-24) of the conjoint task. One-way ANOVAs were conducted to test the hypothesis that respondents' mean ratings are equal across these three groups of profile order positions. The relevant F-Test Statistic values (and their corresponding levels of statistical significance) were 0.77 (0.46), 2.30 (0.10), and 0.41 (0.62) for the four, five, and six attribute tasks, respectively. The null hypothesis of equal means cannot be rejected at conventional levels of statistical significance for all three task conditions. Thus, in aggregate, these respondents' mean ratings were statistically identical throughout the conjoint judgment tasks.

The second, and more important, test of serial profile presentation order bias investigated changes in importance weights for individual respondents. Two pooling tests were conducted. The first pooling test investigates novelty effects. The null hypothesis is that a respondent's importance weights for the first one-third of the profiles are equal to his/her importance weights for the remaining two-thirds of the profiles. The second pooling test investigates fatigue effects. The null hypothesis is that a respondent's importance weights for the last one-third of the profiles are equal to his/her importance weights for the first two-thirds of the profiles. These kinds of pooling tests were originally developed by Chow (1960) and Fisher (1970). See also Johnston (1972, p. 207). The relevant test statistic for both of these pooling tests is calculated as follows:

$$\frac{[SSE(24)-SSE(16)]/8}{SSE(16)/(16-K)}$$

where:

K = the number of independent variables in a regression model (including the constant term)

SSE(24) = the residual sum-of-squared errors from the regression on all 24 of the profiles

SSE(16) = the residual sum-of-squared errors from a regression with only 16 of the profiles (either the first 16 of the last 16).

This test statistic is distributed as F with 9 and 16-K degrees of freedom, where K is the number of variables. For the four, five, and six attribute tasks, there were 5, 6, and 7 variables (plus constant terms), respectively.

Note that this pooling test is not the familiar Chow-test for assessing whether two sample populations are characterized by the same underlying regression model. The testing procedure used here involves splitting the available data for each respondent into groups of 8 and 16. Since it is not possible to estimate a preference regression model for this long-distance discount telephone service example with only 8 data points, it was necessary to use a variant of the familiar Chow-test. This variant involves estimating the regression model with 16 data points, and then adding in the remaining 8 data points (for a total of 24) and observing whether the incremental sum-of-squares increases more than would be expected.

TABLE 3

Estimated Relative Importances (%)

Task Situation and Attributes	Overall	Estimated Relative Importance When the Attribute Was Listed in Position 1	2	3	4	5	6
Four Attribute Task							
Attribute #1: Discount	24.9	26.3	23.9	24.3	24.3		
Attribute #2: Operation Hours	13.0	12.6	12.7	12.4	14.4		
Attribute #3: Fixed Charge	38.6	39.0	38.1	38.7	38.8		
Attribute #4: Extent of Access	23.4	23.1	22.4	24.0	24.5		
	100.0	101.0	97.1	99.4	102.0		
Five Attribute Task							
Attribute #1: Discount	26.3	25.4	26.6	25.1	26.4	27.4	
Attribute #2: Operating Hours	10.0	9.4	9.4	10.9	9.0	11.3	
Attribute #3: Fixed Charge	42.5	42.5	40.9	42.1	44.9	42.3	
Attribute #4: Extent of Access	18.4	19.3	17.7	19.1	16.2	20.7	
Attribute #5: Canada/U.S.?	2.7	1.9	2.5	3.5	3.1	2.3	
	100.0	98.5	97.1	100.7	99.6	104.0	
Six Attribute Task							
Attribute #1: Discount	21.7	19.9	22.7	22.8	21.4	22.1	22.1
Attribute #2: Operation Hours	20.9	22.3	21.0	20.5	19.8	20.8	21.5
Attribute #3: Fixed Charge	26.3	26.4	25.8	27.2	25.5	25.0	28.8
Attribute #4: Extent of Access	20.0	17.6	19.5	20.3	20.5	20.7	20.8
Attribute #5: Canada/U.S.?	6.6	7.1	5.5	6.1	7.3	6.1	7.3
Attribute #6: Telephone Type	4.5	3.9	4.7	3.9	5.1	5.2	4.3
	100.0	97.2	99.2	100.8	99.6	99.9	104.8

Notes: (1) "Overall" refers to the estimated relative importances across all of the respondents, regardless of the position order of an attribute.

This pooling test was conducted for each of the respondents in each of the tasks. The results are displayed in Table 4. There is considerable evidence of nonstationarity in judgment behavior: about 32% of the respondents (51 of 160 respondents) exhibit nonstationarity. Novelty effects were detected in approximately 21% of respondents (34 of 160). These novelty effects occurred for all three levels of task complexity. This suggests that novelty is important in simple tasks (four attributes), as well as in more complex tasks (six attributes). Fatigue effects were detected in approximately 11% of respondents (17 of 160). These effects occurred in all three levels of task complexity, as well. This finding is, of course, not surprising since the number of profiles was held constant across tasks. It is interesting to note that novelty occurs with about twice the frequency as fatigue.

TABLE 4

Pooling Test Results For Nonstationarity

Task	Total Number of Respondents	Number of Respondents Exhibiting Nonstationarity (at the 0.05 Level of Statistical Significance): Novelty Effects	Fatigue Effects
Four Attributes	60	10	8
Five Attributes	43	12	4
Six Attributes	57	12	5

Notes: (1) These statistical tests could not be performed for one of the respondents in each of the five and six attribute tasks. These respondents had provided constant ratings for all profiles they judged.

CONCLUDING REMARKS

The findings in this study demonstrate the clear presence of both attribute presentation order bias and nonstationarity in full profile conjoint rating tasks. However, the actual magnitude of the attribute presentation order bias appears to be relatively small. The existence of a statistically detectable order bias effect suggests that order bias effects may be important in some experimental contexts. The nonstationarity problem appears very serious -- almost one-third of this study's respondents exhibited nonstationarity.

To ensure that attribute presentation order bias and nonstationarity are absent in full profile conjoint analysis rating tasks, a researcher must explicitly conduct tests such as those used in this paper. Such tests can become routine for conjoint tasks administered by microcomputers, since attribute presentation order and serial profile order can be conveniently randomized. When a conjoint task is administered in a paper-questionnaire format, the presence of order bias and temporal instability cannot be analyzed unless costly "split-ballot" questionnaire forms are employed. In a conjoint ranking task (involving iterative sorting of profiles from most to least preferred), temporal instability would normally be undetectable, even if it exists, due to the logistics associated with the ranking task.

A number of additional research inquiries and statistical analyses in the area of attribute presentation order bias and nonstationarity are merited. First, it would be appropriate to try to separate the order bias and nonstationarity issues. Perhaps novelty and fatigue are really just artifacts of the attribute presentation order bias phenomenon, or vice versa. For example, in a within-subjects design (involving varying attribute presentation orders for a single subject), attribute presentation order bias would imply nonstationarity.

Second, in addition to positional order bias, adjacency effects may also exist. These adjacency effects would exist if the relative importance of an attribute was influenced by the other attributes adjacent to it in the listing. One might, for example, hypothesize that attributes of minor importance have their apparent worth increased when they are positioned next to a very important attribute. Adjacency effects could be assessed by further pooling tests. Tests would be required for each pair of attributes. The two data groups in such tests would be the adjacent and the nonadjacent observations. Pooling tests would then reveal whether similar or different decision rules were used under adjacency and nonadjacency conditions.

Third, these nonstationarity findings merit further detailed investigation to assess where novelty ends fatigue sets in. Knowing where novelty effects end would guide the researcher in determining an appropriate number of "warm-up" profiles (to be discarded for the purposes of estimation). More specificity in the onset of fatigue would provide useful information to researchers regarding the length of full profile conjoint evaluation tasks, which would influence the experimental designs used in conjoint studies.

Fourth, subjects in this study were provided with an instruction sheet to which they could refer throughout the conjoint evaluation task. This instruction sheet contained all of the available levels of all attributes. See Exhibit 2 for a sample instruction sheet for the four attribute task. It would be instructive to explore whether the presence of the instruction sheet reduces or increases the attribute presentation order bias and nonstationarity.

Replications of these findings -- in different conjoint contexts and for different respondent groups -- would be highly desirable. Attribute presentation order bias and nonstationarity are serious problems for conjoint analysis research efforts. If the findings in this paper are supported in replications, substantial changes in the way in which conjoint data are collected and analyzed may be required. The presence of attribute presentation order bias, however small, implies that presenting respondents with a single fixed ordering of the attributes may distort their responses. Nonstationarity directly attacks a fundamental assumption in conjoint analysis studies, that respondents use the same decision rule throughout the conjoint judgment task.

EXHIBIT 2

Sample Instruction Sheet

In this questionnaire, you will be asked to provide your opinions about the attractiveness (to you!) of some possible new product concepts.

The particular new product of interest in this study is LONG DISTANCE DISCOUNT TELEPHONE SERVICES. Such a service would operate only from your home telephone. In return for a fixed monthly service charge, you would be able to obtain a discount from regular Telephone System daytime rates on your long distance calls from your home during certain times of the day and/or evening. Such a service might operate to all telephone numbers, or perhaps be more limited in scope and only reach telephone exchanges in major cities.

In this questionnaire, you will be asked to make a number of judgments about the desirability of various kinds of posible long distance discount telephone services. These services will be described in terms of four particular features:

- Discount From Regular Telephone System Daytime Rates
- Hours of Operation During Which the Service is Available
- Monthly Fixed Service Charge (in Addition to Your Long Distance Calls)
- Extent of Access.

In making these judgments, we are interested in your opinions as a potential purchaser and user of such a long distance discount telephone service.

For your reference, the particular features to be used in this study are shown below. Feel free to refer to this list as you work through the questionnaire.

DISCOUNT (Discount From Regular Telephone System Daytime Rates)
- One-Third
- One-Half
- Two-Thirds

OPERATION HOURS (Hours During Which the Service is Available)
- All the Time
- 5PM-8AM Only

FIXED CHARGE (Monthly Charge in Addition to the Cost of Long Distance Calls)
- No Charge
- $10 Per Month

EXTENT OF ACCESS
- Everywhere (All Telephone Numbers)
- Major Cities (Telephone Numbers in Major Cities Only)

REFERENCES

Acito, Franklin (1977), "An Investigation of Some Data Collection Issues in Conjoint Measurement," in Contemporary Marketing Thought, Barnett A. Greenberg and Danny N. Bellenger, eds., Chicago: American Marketing Association, 82-5.

Bettman, James R. (1979), An Information Processing Theory of Consumer Choice, Reading, Massachusetts: Addison-Wesley Publishing Company.

Blunch, Niels J. (1984), "Position Bias in Multiple-Choice Questions," Journal of Marketing Research, 21 (May), 216-20.

Cattin, Philippe and Dick R. Wittink (1982), "Commercial Uses of Conjoint Analysis: A Survey," Journal of Marketing, 46 (Summer), 44-53.

Chapman, Randall G. and Ian Cochrane (1985), "CJ: An Interactive Conjoint Analysis Questionnaire Administration Program For Microcomputers," Faculty of Business, University of Alberta (January).

Chow, Gregory C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," Econometrica, 28, 591-605.

Fisher, Franklin M. (1970), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions: An Expository Note," Econometrica, 38, 361-6.

Green, Paul E. and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," Journal of Consumer Research, 5 (September), 103-23.

Johnston, J. (1972), Econometric Methods, Second Edition, New York: McGraw-Hill Book Company.

Johnson, Richard M. (1982), "Problems in Applying Conjoint Analysis," in Analytical Approaches to Product and Marketing Planning: The Second Conference, Rajendra K. Srivastava and Allan D. Shocker, eds., Cambridge, Massachusetts: Marketing Science Institute, 154-64.

Olshavsky, Richard W. and Franklin Acito (1980), "An Information Processing Probe Into Conjoint Analysis," Decision Sciences, 11, 451-70.

Sudman, Seymour and Norman M. Bradburn (1982), Asking Questions: A Practical Guide to Questionnaire Design, San Francisco: Jossey-Bass Publishers.