

# Extended Axiomatic Conjoint Measurement: A Solution to a Methodological Problem in Studying Fertility-Related Behaviors

Carol A. Nickerson and Gary H. McClelland  
University of Colorado

This paper has three purposes: (1) to develop a methodology based on axiomatic conjoint measurement to accompany the McClelland (1980) fertility decision-making model; (2) to demonstrate the use of the methodology by applying it to the study of a specific fertility-related behavior (contraceptive choice); and (3) to evaluate the usefulness and validity of the model and its methodology for studying fertility-related behaviors. The within-persons methodology developed is contrasted with the across-persons methodology customarily used to test expectancy-value models of behavior. *Index terms: Conjoint measurement, Contraception, Decision making, Expectancy-value models, Preference models, Within-persons analysis.*

As effective means of fertility regulation become more available and more acceptable, fertility-related behaviors (e.g., deciding to have another child, or selecting a birth control method) will increasingly become matters of personal decision (Bourgeois-Pichat, 1967). It therefore seems appropriate to study these behaviors using a model of individual decision making such as that proposed by McClelland (1980). This model depicts decision making as a cognitive process whereby a person evaluates a variety of behavioral alternatives and selects from among them. The purpose of the present paper is to develop a methodology to accompany this

model and to demonstrate its use by applying it to the study of contraceptive choice.

According to the McClelland (1980) model, the decision-making process consists of four major stages: judgment, valuation, integration, and choice. In the *judgment* stage, the person making the decision identifies the possible consequences of performing some behavior. A woman may believe, for example, that the possible consequences of using an intrauterine device (IUD) include effective protection against unwanted pregnancy, convenience of use, low cost, and a slight danger of uterine perforation. These believed consequences may or may not be veridical. In the *valuation* stage, each consequence identified in the judgment stage is assigned some value according to its desirability. Thus, the same woman is likely to view the consequences of effectiveness, convenience, and low cost as very desirable and the consequence of uterine perforation as quite undesirable. Whereas beliefs about consequences resulting from the performance of a behavior may be correct or incorrect in an objective sense, values assigned to consequences are more subjective. In the *integration* stage, the values of the consequences are combined in some manner to form an overall evaluation of the behavior. For example, this woman's values may be combined in such a way that the IUD's effectiveness, convenience, and low cost more than compensate for its health risks, and hence it receives a high overall evaluation. In the *choice* stage, the person com-

---

APPLIED PSYCHOLOGICAL MEASUREMENT  
Vol. 12, No. 2, June 1988, pp. 129-153  
© Copyright 1988 Applied Psychological Measurement Inc.  
0146-6216/88/020129-25\$2.50

compares the overall evaluations of a number of behavioral alternatives (e.g., use of the IUD, pill, diaphragm, etc.) and selects the behavior with the highest overall evaluation. The decision-making model assumes that a person attempts to select a behavior that will maximize benefits (desirable consequences) relative to costs (undesirable consequences); however, decisions may be suboptimal because of cognitive errors made at one or more of the four decision-making stages.

As McClelland (1980) noted, this model has important implications for the study of fertility-related decision making. The emphasis on believed consequences, subjective values, and personal integration rules suggests the existence of significant individual differences at the judgment, valuation, and integration stages of the decision sequence. That is, persons are likely to differ in their judgments of the consequences that may result from a given behavior, and even persons in agreement on the possible consequences of a behavior may hold very different opinions of their desirability. Two women may agree, for example, that the IUD is inexpensive to use, but the woman on a limited budget may value this consequence more than would a wealthier woman. Furthermore, persons with both similar beliefs and similar values may integrate this information differently in forming their overall evaluations. For the woman described above, the effectiveness, convenience, and low cost of the IUD more than compensate for its health risks. Another woman with the same beliefs and values regarding the IUD may combine this information in such a way that the risk of uterine perforation overrides all other considerations and the method receives a low overall evaluation.

The existence of individual differences poses a dilemma for the study of fertility-related decision making. These differences suggest the possibility that several persons may arrive at the same choice for different reasons; that is, the same choice can result from different sets of beliefs, values, and integration rules. For example, one woman may decide to use the IUD because she believes that it is very convenient and she places high value on convenience, whereas another may decide to use the IUD because she believes that it is very effective

in preventing pregnancy and she places high value on effectiveness. It is impossible, therefore, to infer beliefs, values, or integration rules from the observation of a single chosen behavior.

The presence of individual differences also implies that important details of the decision-making process are lost by aggregating or averaging data over persons. For example, suppose that one-third of a study's respondents value highly the convenience of a contraceptive method, another third place moderate value on convenience, and the remaining third are not at all concerned about this consequence. In this case, averaging the data across all respondents would lead to the conclusion that convenience is a moderately-valued contraceptive consequence, a summary that does not reflect the values of two-thirds of the respondents.

The decision-making model thus implies that inferences about fertility-related decisions cannot be made from the observation of single behaviors, but it also rules out the aggregation of data across persons as a means of observing a number of behaviors. The resolution of this dilemma clearly requires the within-persons analysis of a number of behaviors performed by the same person. Unfortunately, however, a person usually makes only a few fertility-related decisions in an entire lifetime, seemingly eliminating the possibility of employing this strategy.

The dilemma can be resolved if it is possible to obtain *preference* information about the entire set of behavioral alternatives under consideration by the person. This approach simulates the observation of a number of behaviors by requiring the person to answer a series of "what-if" questions (e.g., "What is your most preferred contraceptive method?", "What contraceptive method would you want to use if for some reason you were unable to use your most preferred method?", etc.) or, equivalently, to rank-order according to preference the set of behaviors (e.g., use of various contraceptive methods) under consideration.

If the consequences associated with the entire set of behavioral alternatives are considered as a whole, very often it will become evident that those consequences are in fact *variables*. That is, a consequence may constitute one category or level of

a more global, multilevel consequence. For example, three different contraceptive methods may be 75%, 85%, and 95% effective, respectively. These percentages may be viewed as three different levels of a single effectiveness consequence rather than as three separate consequences. The "levels of a consequence" are analogous to the levels of an independent variable in analysis of variance.

If the behavioral consequences can be viewed as variables, and if the set of behavioral alternatives to be ranked can be created by the factorial (or fractional factorial) combination of the varying levels of those consequences, then axiomatic conjoint measurement (Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971) can be used to determine a person's integration rule and to derive his or her values for each consequence, each consequence level, and each behavioral alternative. A person's beliefs about the behavioral alternatives are then the levels of the consequences describing those alternatives. Note, however, that those alternatives are necessarily hypothetical, however realistic their descriptions might be (Wallsten & Budescu, 1981). Real-world behaviors are not naturally described by their consequences but are intact, wholistic, *named* behaviors (Anderson, Deane, Hammond, McClelland, & Shanteau, 1981, pp. 79–80, 257–258).

Use of axiomatic conjoint measurement is thus problematic when the interest is in studying real-world behaviors, as is usually the case in fertility research (and in many other areas of research as well). The problem can be circumvented, however, if the preference information obtained about hypothetical behavioral alternatives is supplemented by belief and preference information about real-world behavioral alternatives. The person's beliefs about these alternatives are used to establish a correspondence between the hypothetical and the real-world behaviors. The preference ordering across real-world behaviors can then be used to determine the degree to which that ordering is predicted by the person's beliefs, values, and integration rule.

The present study demonstrates how this extension of axiomatic conjoint measurement might be used to study a real-world fertility-related decision. The methodology is used to study the process by

which a person who wishes to delay pregnancy selects a contraceptive method from a set of possible alternative methods. Empirical applications of axiomatic conjoint measurement are not common in psychological research. Furthermore, with few exceptions (e.g., Wallsten, 1972, 1976), those applications that do exist have employed the technique for determining integration rules and have de-emphasized or neglected the derivation of values (e.g., Coombs, 1979; Coombs & Bowen, 1971; Coombs, Coombs, & McClelland, 1975; Coombs & Huang, 1970; Fischer, 1976; Ullrich & Painter, 1974; Wallsten & Budescu, 1981). No study has used the methodology to study wholistic rather than hypothetical behaviors. Therefore, a step-by-step analysis of one respondent's data is presented first in order to demonstrate the methodology. Summary results are then presented for an entire group of respondents. Particular emphasis is given to verifying the existence of individual differences at the judgment, valuation, and integration stages of the decision-making process, as it is the existence of these differences that necessitates a within-persons methodology.

## Method

### Respondents

One hundred clients of a Boulder, Colorado family-planning clinic agreed to participate in the study. Both men and women were included in the sample; each respondent received \$5 for his or her participation.

### Materials and Procedure

Because this was a demonstration study, expository simplicity demanded that some constraints be placed upon the full range of individual differences that might exist. Accordingly, beliefs about birth control methods were restricted to beliefs associated with three consequences of contraceptive use: convenience, effectiveness/reversibility, and side effects. These particular consequences were selected because they seem to reflect the contraceptive concerns most often expressed in the birth control information pamphlets and booklets dis-

tributed by family-planning clinics and women's health care groups (e.g., *Basics of Birth Control*, 1977; *Birth Control Handbook*, 1973; Boston Women's Health Book Collective, 1976).

*Convenience* refers to such considerations as the number of visits to a physician that the use of a particular contraceptive method might require, whether the method requires preparation immediately before intercourse, and whether the method is "messy." *Effectiveness* is defined in terms of the percentage of women per year who become pregnant while using a contraceptive method. *Reversibility* refers to the possibility of being able to bear children after discontinuing use of a contraceptive method. Effectiveness and reversibility were combined into a single consequence because they are not independent; it is not possible, for example, for a contraceptive method to be both ineffective and irreversible. *Side effects* refers to the possible health risks that may arise from use of a contraceptive method.

A possible fourth consequence, *cost*, was not included because pilot studies showed that respondents either do not consider the cost of a contraceptive method to be important, or believe that contraceptive methods do not vary much in overall cost. Without such variance, the cost consequence cannot discriminate between contraceptive methods and therefore cannot affect predictive power, regardless of its importance.

Three levels of each contraceptive consequence were chosen so that the most popular birth control methods currently available might reasonably be characterized by one level of each consequence. Because the verbal descriptions of these levels are quite lengthy, the levels are often referred to in the text and tables in an abbreviated form. The consequence levels and their abbreviations are shown in Table 1.

Each respondent was given a list of eight real birth control methods: (1) tubal ligation, (2) IUD, (3) vasectomy, (4) pill, (5) condom, (6) diaphragm, (7) vaginal foam/cream/jelly, and (8) rhythm. Each method listed was accompanied by a short description that made no mention of its convenience, effectiveness, reversibility, or side effects. For example, the diaphragm was described

as "a thin rubber dome that is inserted into the vagina to cover the cervix and prevent sperm from entering the uterus."

As the first task (the *judgment* task), each respondent was asked to select the level of each consequence that he or she believed to be the most accurate description of each of the eight real birth control methods. Thus, a respondent might believe that the diaphragm is "somewhat inconvenient," that "between 15 and 35 women out of 100 will have accidental pregnancies in a year and that it is reversible ([one] can nearly always stop using this method and then have children)," and that its use causes "rarely any side effects at all."

For the second task (the *hypothetical contraceptive ranking* task), the three levels of each consequence were factorially combined in a  $3 \times 3 \times 3$  design resulting in 27 consequence-level combinations. Each of these combinations describes a hypothetical contraceptive alternative and was printed on a separate card. Each respondent was instructed to rank-order the 27 cards from "most preferred" to "least preferred" contraceptive as if he or she "didn't want to have a child just yet but did want a family in the future."

As the third task (the *real birth control ranking* task), each respondent was asked to rank-order the list of eight real birth control methods from most preferred to least preferred, as if he or she "didn't want to have a child just yet but did want a family in the future."

It may seem odd that a respondent asked to imagine that he or she "didn't want to have a child just yet but did want a family in the future" would be offered tubal ligation and vasectomy as choice alternatives. Tubal ligation and vasectomy were included because they *are* among the birth control methods available, and it is possible that a person would prefer an irreversible method to an ineffective one. This might happen, for instance, if a person were ambivalent about having children and would rather not have any children at all than risk having one now by using a less effective birth control method.

#### Analysis

The decision-making model was tested individ-

Table 1  
 Levels of the Three Contraceptive Consequences

| Consequence and Levels   | Abbreviation |
|--|--------------|
| Convenience  |              |
| (C1) very convenient   | very         |
| (C2) somewhat inconvenient   | somewhat     |
| (C3) inconvenient  | not          |
| Effectiveness/reversibility  |              |
| (ER1) less than 1 woman out of 100 will have an accidental pregnancy in a year but it is not reversible (once started, the use of the method cannot be discontinued)       | 99% IR       |
| (ER2) less than 5 women out of 100 will have accidental pregnancies in a year and it is reversible (can nearly always stop using this method and then have children)       | 95% R        |
| (ER3) between 15 and 35 women out of 100 will have accidental pregnancies in a year and it is reversible (can nearly always stop using this method and then have children) | 65-85% R     |
| Side effects   |              |
| (SE1) rarely any side effects at all   | rare         |
| (SE2) sometimes minor side effects—a nuisance, but not dangerous   | minor        |
| (SE3) sometimes minor side effects and possible major (dangerous to health) side effects   | major        |

ually for each respondent. Validation of the model consists of a sequence of three analytic steps: (1) testing one or more proposed integration rules, (2) derivation of the values assigned to the contraceptive consequences, and (3) determination and validation of the consequences believed associated with each of the eight real birth control methods. If one of the analytic steps in the sequence cannot be performed, analysis proceeds no further. If all three steps can be performed, then a correlation coefficient can be computed to determine the degree to which a person's beliefs, values, and integration rule predict his or her preferences across the set of real birth control methods and thereby validate the decision-making model. A detailed explanation of each analytic step is provided below. For concreteness, the dataset of one respondent is used to illustrate the analytic procedures.

The illustrative respondent's preference rank ordering of the 27 hypothetical contraceptive alternatives is shown in Table 2. The respondent's most preferred alternative, ranked 27, is one that is very convenient, is 95% effective and reversible, and rarely causes side effects. This contraceptive is obviously ideal for a person wishing to delay child-bearing. The remainder of the rank ordering exhibits a very marked degree of regularity. Alternatives that are 95% effective and reversible are always preferred to those that are 65% to 85% effective and reversible, which in turn are always preferred to those that are 99% effective and irreversible.

If the level of effectiveness/reversibility is held constant, an alternative that is very convenient is preferred to an alternative that is somewhat inconvenient as long as the first alternative is associated

Table 2  
Ranks and Overall Evaluations of the 27 Hypothetical  
Contraceptive Alternatives, Ordered From Most to Least Preferred,  
for the Illustrative Respondent

| Hypothetical Contraceptive Alternative |                                 |                 | Rank<br>Order | Overall<br>Evaluation | Rescaled<br>Overall<br>Evaluation |
|--|---------------------------------|-----------------|---------------|-----------------------|-----------------------------------|
| Convenience                            | Effectiveness/<br>Reversibility | Side<br>Effects |               |                       |                                   |
| very                                   | 95% R                           | rare            | 27            | 2.241                 | 1.00                              |
| very                                   | 95% R                           | minor           | 26            | 2.148                 | .94                               |
| somewhat                               | 95% R                           | rare            | 25            | 2.093                 | .90                               |
| very                                   | 95% R                           | major           | 24            | 2.055                 | .87                               |
| somewhat                               | 95% R                           | minor           | 23            | 2.000                 | .84                               |
| not                                    | 95% R                           | rare            | 22            | 1.945                 | .80                               |
| somewhat                               | 95% R                           | major           | 21            | 1.907                 | .77                               |
| not                                    | 95% R                           | major           | 20 (19)       | (1.759)               | (.67)                             |
| not                                    | 95% R                           | minor           | 19 (20)       | (1.852)               | (.74)                             |
| very                                   | 65-85% R                        | rare            | 18            | 1.741                 | .66                               |
| very                                   | 65-85% R                        | minor           | 17            | 1.648                 | .60                               |
| somewhat                               | 65-85% R                        | rare            | 16            | 1.593                 | .56                               |
| very                                   | 65-85% R                        | major           | 15            | 1.555                 | .54                               |
| somewhat                               | 65-85% R                        | minor           | 14            | 1.500                 | .50                               |
| not                                    | 65-85% R                        | rare            | 13            | 1.445                 | .46                               |
| somewhat                               | 65-85% R                        | major           | 12            | 1.407                 | .44                               |
| not                                    | 65-85% R                        | minor           | 11            | 1.352                 | .40                               |
| not                                    | 65-85% R                        | major           | 10            | 1.259                 | .34                               |
| very                                   | 99% IR                          | rare            | 9             | 1.241                 | .33                               |
| very                                   | 99% IR                          | minor           | 8             | 1.148                 | .26                               |
| somewhat                               | 99% IR                          | rare            | 7             | 1.093                 | .23                               |
| very                                   | 99% IR                          | major           | 6             | 1.055                 | .20                               |
| somewhat                               | 99% IR                          | minor           | 5             | 1.000                 | .16                               |
| not                                    | 99% IR                          | rare            | 4             | .945                  | .13                               |
| somewhat                               | 99% IR                          | major           | 3             | .907                  | .10                               |
| not                                    | 99% IR                          | minor           | 2             | .852                  | .06                               |
| not                                    | 99% IR                          | major           | 1             | .759                  | .00                               |

Note. Numbers in parentheses indicate ranks and evaluations after error correction.

only with rare or minor side effects. However, a somewhat inconvenient alternative that rarely causes side effects is preferred to a very convenient alternative that may cause major side effects. This distaste for alternatives with major side effects is evident throughout the rank ordering; the respondent always prefers to switch to an alternative with a less desirable level of convenience, rather than to one associated with major side effects. The regularity of the rank ordering suggests that it may satisfy the axioms of a simple polynomial integration rule.

The preference pattern also suggests the respondent's relative valuation of the three contraceptive consequences. The consequences and consequence levels that are relinquished first in the rank ordering are those that are valued less. Thus, the rank ordering reveals that the consequence of effectiveness/reversibility is much more important to this respondent than either convenience or side effects, and that convenience is somewhat more important than side effects. These qualitative observations are verified by the formal axiomatic conjoint measurement analysis.

**Step 1: Test of the Integration Rule**

Axiomatic conjoint measurement can diagnose three classes of simple polynomial integration rules in three variables: additive rules ( $A + P + U$ ), distributive rules  $[(A + P)U]$ , and dual-distributive rules ( $AP + U$ ) (Krantz et al., 1971; Krantz & Tversky, 1971). It can also diagnose variants of these rules, such as an additive rule in three variables in which one of the variables is irrelevant or a distributive rule in which the three variables are permuted. The additive integration rule in three variables is tested first because it is the simplest of the polynomial integration rules. Other integration rules are considered only if the additive rule is rejected.

An additive integration rule implies that the overall evaluation of each hypothetical contraceptive alternative can be computed by summing the values associated with the three consequence levels describing that alternative. For an integration rule for a  $3 \times 3 \times 3$  factorial design to be additive, a preference rank ordering must satisfy the axioms of (1) simple independence, (2) joint independence, and (3) double cancellation. *Simple independence* means that the ordering of the levels of each consequence must be invariant over every combination of the levels of the other two consequences. *Joint independence* means that the ordering of every combination of the levels of any two consequences must be invariant over the levels of the third consequence. Joint independence implies simple independence, but the converse is not true. However, when simple independence is satisfied, violations of joint independence are not common.

*Double cancellation* has no simple intuitive explanation and is best illustrated by an example. Suppose that the level of any one of the three consequences is held constant and that the cells of the resulting  $3 \times 3$  factorial combination of the other two consequences (X and Y) are labeled as shown in Figure 1. If the rank of cell b is greater than or equal to the rank of cell f and if the rank of cell d is greater than or equal to the rank of cell h, then the rank of cell a must be greater than or equal to the rank of cell i if double cancellation is to be satisfied. For a rank ordering to be considered con-

sistent with an additive composition rule, double cancellation must be satisfied for all possible combinations of two of the three consequences at every level of the third consequence. Double cancellation is rarely violated if the rank ordering satisfies simple and joint independence.

Examination of the illustrative respondent's preference rank ordering of the 27 hypothetical contraceptive alternatives helps to clarify the nature of these three axioms. This examination is simplified if the ranks of the alternatives are arranged into three different row  $\times$  column  $\times$  plane matrices, as shown in Table 3. In each matrix, the column consequence is the consequence being tested for simple independence. Table 3a shows that the preference ordering of the levels of convenience (C) is  $C1 > C2 > C3$  at each of the nine possible combinations of the levels of effectiveness/reversibility and side effects; convenience is therefore independent of effectiveness/reversibility and side effects.

Table 3b shows that the preference ordering of the levels of effectiveness/reversibility (ER) is  $ER2 > ER3 > ER1$  at each of the nine possible combinations of the levels of side effects and convenience; effectiveness/reversibility is therefore independent of side effects and convenience.

Figure 1  
 The Double Cancellation Axiom

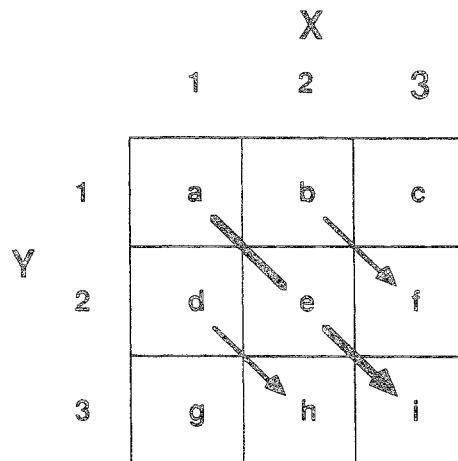


Table 3  
Three Arrangements of the Preference Rank Ordering  
of the Hypothetical Contraceptive Alternatives  
for the Illustrative Respondent

| (a) ER × C × SE |     |    |    |     |    |    |     |    |    |
|-----------------|-----|----|----|-----|----|----|-----|----|----|
|                 | SE1 |    |    | SE2 |    |    | SE3 |    |    |
|                 | C1  | C2 | C3 | C1  | C2 | C3 | C1  | C2 | C3 |
| ER1             | 9   | 7  | 4  | 8   | 5  | 2  | 6   | 3  | 1  |
| ER2             | 27  | 25 | 22 | 26  | 23 | 19 | 24  | 21 | 20 |
| ER3             | 18  | 16 | 13 | 17  | 14 | 11 | 15  | 12 | 10 |

| (b) SE × ER × C |     |     |     |     |     |     |     |     |     |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|                 | C1  |     |     | C2  |     |     | C3  |     |     |
|                 | ER1 | ER2 | ER3 | ER1 | ER2 | ER3 | ER1 | ER2 | ER3 |
| SE1             | 9   | 27  | 18  | 7   | 25  | 16  | 4   | 22  | 13  |
| SE2             | 8   | 26  | 17  | 5   | 23  | 14  | 2   | 19  | 11  |
| SE3             | 6   | 24  | 15  | 3   | 21  | 12  | 1   | 20  | 10  |

| (c) C × SE × ER |     |     |     |     |     |     |     |     |     |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|                 | ER1 |     |     | ER2 |     |     | ER3 |     |     |
|                 | SE1 | SE2 | SE3 | SE1 | SE2 | SE3 | SE1 | SE2 | SE3 |
| C1              | 9   | 8   | 6   | 27  | 26  | 24  | 18  | 17  | 15  |
| C2              | 7   | 5   | 3   | 25  | 23  | 21  | 16  | 14  | 12  |
| C3              | 4   | 2   | 1   | 22  | 19  | 20  | 13  | 11  | 10  |

In Table 3c, however, the preference ordering of the levels of side effects (SE) is SE1 > SE2 > SE3 at only eight of the nine possible combinations of the levels of convenience and effectiveness/reversibility. At combination C3 × ER2, the ordering is SE1 > SE3 > SE2. The reversal of SE2 and SE3 violates simple independence for side effects.

In each matrix in Table 3, the consequences being tested for joint independence and double cancellation are the row and column consequences. In Table 3a the preference ordering of the nine combinations of effectiveness/reversibility and convenience is the same at each of the three levels or planes of the side effects consequence. This is easy to see if the combinations in each plane are renumbered from 9 (most preferred) to 1 (least preferred). Reading across each row, the nine combinations are ordered 3, 2, 1, 9, 8, 7, 6, 5, 4 in each plane; therefore effectiveness/reversibility and convenience are jointly independent of side effects.

If the nine combinations of side effects and effectiveness/reversibility in each plane of Table 3b are renumbered in a similar fashion and read across each row, the combinations are ordered 3, 9, 6, 2, 8, 5, 1, 7, 4 in only the first and second planes of convenience; the ordering in the third plane reverses the alternatives with the (renumbered) ranks of 7 and 8. As a result of this reversal, side effects and effectiveness/reversibility are not jointly independent of convenience.

In Table 3c the nine combinations of convenience and side effects have the same (renumbered) ordering of 9, 8, 6, 7, 5, 3, 4, 2, 1 in only the first and third planes of effectiveness/reversibility; the ordering in the second plane reverses the alternatives with the (renumbered) ranks of 1 and 2. Therefore, convenience and side effects are not jointly independent of effectiveness/reversibility.

Because the antecedent conditions do not hold, there are no tests of double cancellation for the



combination of effectiveness/reversibility and convenience shown in Table 3a, or for the combination of side effects and effectiveness/reversibility shown in Table 3b. A "no test" situation does not constitute a violation of additivity. In Table 3c double cancellation is satisfied for the combination of convenience and side effects at all three levels of effectiveness/reversibility.

In practice, computer programs such as CONJOINT (Holt & Wallsten, 1974), CMSICAL (Nygren, 1986), or PCJM2 (Ullrich & Cummins, 1973; Ullrich, Cummins, & Walkenbach, 1978) are used to test the axioms for any proposed simple polynomial integration rule. PCJM2 was used to test the three additivity axioms for the illustrative respondent's preference rank ordering of the hypothetical contraceptive alternatives. The program yielded the same conclusions regarding the satisfaction of the additivity axioms as did direct examination of the three matrices in Table 3.

Note that all of the violations of simple and joint independence are due to the ordering of a single pair of alternatives. These alternatives differ only in their levels of the side effects consequence. The more preferred alternative (ranked 20 in Table 2) may have major side effects, whereas the less preferred alternative (ranked 19) may have minor side effects. This preference probably represents a mistake by the respondent because (1) it involves alternatives with adjacent ranks, (2) it makes no sense to prefer major to minor side effects, and (3) it is inconsistent with all her other preferences.

Such errors are the bane of axiomatic conjoint measurement. Although attempts are being made to develop an error theory that will specify the number of axiom violations that should cause rejection of a proposed integration rule (e.g., Arbuckle & Larimer, 1976; Falmagne, 1978; McClelland, 1977; Nygren, 1985, 1986), no adequate error theory is yet available. Because data are typically fallible, however, an evaluation of the validity of an integration rule should take the possibility of error into account (Krantz & Tversky, 1971). Requiring perfect consistency with the axioms would be unrealistic; it is analogous to rejecting the model in a correlational study if the correlation coefficient is less than 1.00.

One strategy for coping with error that has seen prior use (e.g., Coombs, 1979; Coombs & Bowen, 1971; Coombs et al., 1975; Ullrich & Painter, 1974) is to allow a small number of nonsystematic errors before rejecting the proposed integration rule. In the present study, up to three errors were allowed, where an error is defined as either the reversal of a pair of alternatives with adjacent ranks, or an alternative that is clearly misplaced in the ordered card deck. Allowing only three errors for a rank ordering of 27 alternatives is a stringent criterion, considering the possible number of errors that might be made (Nygren, 1985). In the illustrative respondent's data, the pair of alternatives with ranks 19 and 20 constitutes the only error. Because reversing the ranks of these alternatives eliminates all axiom violations, the rank ordering was judged consistent with an additive integration rule according to the above strategy for dealing with error.

Had the ordering been judged inconsistent with an additive integration rule, it would have been tested for consistency with an additive-variant rule, or a distributive or dual-distributive rule. Space considerations do not permit presentation of the procedures for testing these rules. Nickerson and McClelland (1984) defined a number of additive-variant rules and described error correction procedures for these rules. Krantz and Tversky (1971) discussed axiom testing for distributive and dual-distributive rules and provided diagnostic flowcharts describing the sequence in which integration rules are proposed and tested.

### Step 2: Derivation of Values

Once the integration rule has been determined, values may be derived. Given a preference rank ordering that perfectly satisfies the axioms of the proposed integration rule, or that has been corrected to satisfy those axioms as in the example above, various scaling programs may be used to produce a set of values, one value for each level of each consequence. These values are then combined according to the validated integration rule to yield an overall evaluation of each hypothetical alternative.

Scaling programs for additive and additive-var-

iant integration rules include ADDALS (de Leeuw, Young, & Takane, 1976), OPSCOLES (Emery, 1978), NMRG (Johnson, 1975), MONANOVA (Kruskal, 1965), ORDMET (Delbeke, 1979; Lehner & Noma, 1980; McClelland & Coombs, 1975), UNICON (Roskam, 1974), and POLYCON (Young, 1972). DIST (Emery, 1977a) and DULST (Emery, 1977b) scale data for distributive and dual-distributive integration rules, respectively. These programs employ iterative search algorithms or least-squares regression techniques, or solve systems of linear inequalities, to estimate values for consequence levels so that the overall evaluations of the hypothetical alternatives form an interval scale that is monotonic with the original preference rank ordering of those alternatives.

MONANOVA (Kruskal, 1965) was used to produce the values shown in Table 4 for the illustrative respondent's preference rank ordering of the hypothetical contraceptive alternatives. The range of values across the levels of a consequence reflects the relative importance of that consequence in determining the respondent's overall evaluations of the hypothetical alternatives. Thus effectiveness/reversibility (range = 1.000) is more than three times as important as convenience (range = .296), which in turn is about twice as important as side effects (range = .186). A relative importance percentage can be computed for each consequence by dividing

the value range of that consequence by the sum of the value ranges of all three consequences, then multiplying by 100. The relative importance percentage of each contraceptive consequence is shown in the last column of Table 4. The quantitative assessment of the importance to this respondent of the three contraceptive consequences thus verifies the qualitative assessment made previously.

It should be emphasized that the importance attached to each consequence is relative to the set of consequences and consequence levels being studied, and might change if different consequences and/or consequence levels were to be used. It might also change depending upon the purpose for which a contraceptive method is being chosen. For example, a respondent may place different values on effectiveness/reversibility depending upon whether she wishes to postpone or to cease childbearing.

Because the integration rule is an additive one, the overall evaluation of each hypothetical contraceptive alternative equals the sum of the three values corresponding to the three consequence levels describing that alternative. Thus, the overall evaluation of the alternative that is very convenient, 95% effective and reversible, and rarely causes side effects (ranked 27) is  $.648 + 1.000 + .593 = 2.241$ . The overall evaluation of the alternative that is very convenient, 95% effective and reversible,

Table 4  
 Values and Relative Importance Percentages  
 of the Three Contraceptive Consequences  
 for the Illustrative Respondent

| Consequence and Levels      | Value | Relative Importance |
|-----------------------------|-------|---------------------|
| Convenience                 |       |                     |
| (C1) very                   | .648  | 20.0%               |
| (C2) somewhat               | .500  |                     |
| (C3) not                    | .352  |                     |
| Effectiveness/reversibility |       |                     |
| (ER1) 99% IR                | 1.000 | 67.5%               |
| (ER2) 95% R                 | 1.000 |                     |
| (ER3) 65-85% R              | .500  |                     |
| Side effects                |       |                     |
| (SE1) rare                  | .593  | 12.5%               |
| (SE2) minor                 | .500  |                     |
| (SE3) major                 | .407  |                     |

and risks minor side effects (ranked 26) is  $.648 + 1.000 + .500 = 2.148$ , and so on. The complete set of overall evaluations is shown in column 5 of Table 2. For intuitive appeal, the set of overall evaluations has been rescaled and rounded so that the overall evaluations of the most and least preferred alternatives equal 1.0 and 0.0, respectively. Rescaling is permissible because the set of overall evaluations is unique only up to a positive linear transformation. The rescaled overall evaluations are shown in the last column of Table 2.

**Step 3:  
 Determination and Validation of Beliefs**

The determination and validation of a person's beliefs about the consequences of the eight real birth control methods is a two-part procedure. First, a correspondence between the real birth control methods and the hypothetical contraceptive alternatives is established by using the information obtained in the judgment task to match each real birth control method to one of the hypothetical contraceptive alternatives. In the judgment task, the illustrative respondent chose the level of each consequence that she believes to describe most accurately each of the eight real birth control methods. For example, she believes the diaphragm to be a somewhat inconvenient, 95% effective and reversible birth control method with possible minor side effects; this description matches the hypothetical alternative with rank 23. Consequence levels, ranks, and overall evaluations of the hypothetical alternatives matching the eight real birth control methods are shown in Table 5. Note that tubal ligation and vasectomy, and the pill and the IUD, have tied ranks and evaluations because the respondent believes that they are described by the same level of each of the three consequences.

Second, the validity of the beliefs about the real birth control methods is ascertained by computing the correlation between the ranking of the real birth control methods and the overall evaluations of the hypothetical contraceptive alternatives matched to those methods. The overall evaluations of the hypothetical alternatives are by definition a function only of the integration rule and values used by the

Table 5  
 Belief Descriptions and Rankings of the Eight Real Birth Control Methods for the Illustrative Respondent

| Birth Control Method | Convenience | Effectiveness/<br>Reversibility | Side Effects | Hypothetical Alternatives |              | Rescaled Overall Evaluation |              | Real Birth Control |  |
|----------------------|-------------|---------------------------------|--------------|---------------------------|--------------|-----------------------------|--------------|--------------------|--|
|                      |             |                                 |              | Ranking Task              | Ranking Task | Evaluation                  | Ranking Task | Ranking Task       |  |
| Pill                 | very        | 95% R                           | major        | 24 (7.5)                  |              | .87 (7.5)                   |              | 8                  |  |
| Intrauterine device  | very        | 95% R                           | major        | 24 (7.5)                  |              | .87 (7.5)                   |              | 6                  |  |
| Diaphragm            | somewhat    | 95% R                           | minor        | 23 (6)                    |              | .84 (6)                     |              | 7                  |  |
| Condom               | somewhat    | 65-85% R                        | rare         | 16 (5)                    |              | .56 (5)                     |              | 5                  |  |
| Rhythm               | not         | 65-85% R                        | rare         | 13 (4)                    |              | .46 (4)                     |              | 4                  |  |
| Foam/cream/jelly     | not         | 65-85% R                        | minor        | 11 (3)                    |              | .40 (3)                     |              | 3                  |  |
| Tubal ligation       | very        | 99% IR                          | minor        | 8 (1.5)                   |              | .26 (1.5)                   |              | 2                  |  |
| Vasectomy            | very        | 99% IR                          | minor        | 8 (1.5)                   |              | .26 (1.5)                   |              | 1                  |  |

Note. The numbers in parentheses in columns 5 and 6 are the ranks and evaluations re-ranked from 8 (most preferred) to 1 (least preferred) for ease in comparing the values in these columns with those in column 7.

respondent, because the "beliefs" or consequence levels are specified in the hypothetical descriptions. In contrast, the ranking of the real birth control methods is a function of the respondent's integration rule and values *and* of her beliefs about the consequences of each birth control method. If it is assumed that the respondent used the same integration rule and values in rank-ordering both the real methods and the hypothetical alternatives, this correlation constitutes a test of the accuracy of the consequences and the consequence levels used in this study for describing her beliefs about the real birth control methods.

The Goodman-Kruskal (1954) gamma correlation for rank order data with ties was used to assess the relationship between the ranking of the real birth control methods and the overall evaluations of the hypothetical contraceptive alternatives matched to those methods. The gamma for the illustrative respondent's data was .92, indicating a slightly-less-than-perfect capturing of her beliefs about real birth control methods. Note that the respondent described the pill and the IUD with the same consequence levels, but assigned those methods ranks of 8 and 6, respectively, in the real birth control ranking task, causing the slight discrepancy.

Two possible reasons (other than random error) exist for this discrepancy. The respondent may have considered some consequence(s) in addition to those of convenience, effectiveness/reversibility, and side effects when ranking the real birth control methods. The additional consequence(s) distinguishes the pill from the IUD, with the pill having a more desirable consequence level. Alternatively, the respondent may have been able to make finer distinctions in consequence levels than were permitted in this study. For example, she may believe that use of the IUD may cause more serious side effects than does use of the pill, even though she described both as possible causes of major side effects.

The correlation between the overall evaluations of the hypothetical contraceptive alternatives and the ranking of the real birth control methods was described above as being a test of the accuracy of the consequences and consequence levels used. This correlation actually reflects *both* the accuracy of

the consequences and consequence levels in describing the respondent's beliefs about the real birth control methods *and* the degree to which the respondent's beliefs, values, and integration rule predict her preferences across the real birth control methods. This is because testing the decision-making model with the extended axiomatic conjoint measurement methodology is a funnel-like procedure involving a sequence of analytic steps. If one of the steps in the sequence cannot be performed, analysis proceeds no further. The validation of beliefs is the last step in the sequence; hence, if the beliefs can be validated, then the decision-making model can be validated as well.

A gamma correlation coefficient of 1.00 would indicate that the decision-making model has perfectly captured the respondent's integration rule, values, and beliefs, and accurately described her decision-making process. The term "perfectly" is used in the context of the set of consequences and consequence levels included in this study. It would be possible for a respondent to hold slightly different beliefs than those allowed in this study and still obtain a gamma of 1.00. For example, a respondent may really believe that the condom is 60% effective and reversible, but may have chosen 65% to 85% effective and reversible because a consequence level describing her true belief was not available. A perfect gamma could still occur in this situation as long as the true beliefs are close enough to the allowed beliefs so that the ranking of the real birth control methods is not affected.

If it is assumed that the same integration rule and values were used in both preference rank ordering tasks, then failure of the decision-making model must be attributed to inaccuracy in the description of the beliefs, as discussed above. There is no way to test whether the respondent did indeed use the same integration rule and values in both tasks. Note, however, that the slightly-less-than-perfect correlation occurs because she matched the pill and the IUD to the same hypothetical alternative so that they ought to have adjacent ranks in the real birth control method rank ordering, regardless of the integration rule used. Instead, the diaphragm is ranked between them. This suggests that the slight failure of the model must be due to a failure

of the consequences and consequence levels used in this study to reflect accurately the respondent's beliefs about real birth control methods, rather than to the use of a different integration rule or different values in the two tasks.

**Results**

The analysis outlined above was performed individually for each respondent in the study. A summary of the integration rules, values, and beliefs for all 100 respondents follows.

**Integration Rules**

The preference rank ordering of the 27 hypothetical contraceptive alternatives was consistent with an additive or an additive-variant integration rule for 66 of the 100 respondents. Each of these rules is defined below, and their frequency distribution is shown in Table 6.

The rank orderings of 32 respondents satisfied the axioms of simple independence, joint independence, and double cancellation for all three contraceptive consequences with the correction of three or fewer errors, as defined above. The integration rules of these respondents were classified as "additive over three consequences."

Two respondents produced rank orderings that satisfied the three additivity axioms for only two of the three contraceptive consequences, with random ordering of the levels of the third and least important consequence. These respondents were assumed to be indifferent to that consequence, which is labeled "inessential." (See Coombs et al., 1975, p. 285, and Krantz et al., 1971, p. 256, for formal definitions of essential and inessential variables.) The integration rules of these respondents were classified as "additive over two consequences."

An additional nine respondents were indifferent to two of the three contraceptive consequences. The rank orderings of these respondents show the set of hypothetical alternatives divided into three subsets of nine alternatives. The alternatives in each subset are characterized by one level of the important consequence, with random ordering of the alternatives within each subset. The integration rules of these respondents were classified as "one-consequence" rules. (One-consequence rules might also be considered "additive over one consequence;" however, these rules are not additive in the strict sense, as two or more consequences are required for additivity.)

For 21 respondents, the rank orderings of the 18 most preferred alternatives satisfied the three additivity axioms, but the nine least preferred alter-

Table 6  
 Classification of Integration Rules

| Integration Rule  | <i>n</i> |
|---|----------|
| Additive over three consequences: C × ER × SE             | 32       |
| Additive over two consequences                            | 2        |
| C × SE (ER inessential)                                   | 2        |
| Additive over one consequence                             | 9        |
| C only (ER and SE inessential)                            | 2        |
| ER only (C and SE inessential)                            | 5        |
| SE only (C and ER inessential)                            | 2        |
| Conjunctive/additive over three consequences: C × ER × SE | 21       |
| (C3 unacceptable)   | 1        |
| (ER1 unacceptable)  | 9        |
| (ER3 unacceptable)  | 1        |
| (SE3 unacceptable)  | 10       |
| Conjunctive/additive over two consequences                | 2        |
| C × ER (SE inessential, ER1 unacceptable)                 | 1        |
| ER × SE (C inessential, SE3 unacceptable)                 | 1        |
| Not classified  | 34       |

natives—all nine described by the least desirable level of one consequence—were randomly ordered. It appears that these respondents first set aside these alternatives without ordering them (for example, the nine alternatives with the greatest risk of major side effects), then ordered the remaining alternatives consistent with an additive integration rule. The integration rules of these respondents were classified as “conjunctive/additive over three consequences.” A conjunctive integration rule classifies an alternative as acceptable only if the levels of all the consequences describing that alternative are acceptable (Anderson et al., 1981, pp. 54–55). A conjunctive/additive integration rule additionally orders the acceptable alternatives according to an additive integration rule, while the unacceptable alternatives remain randomly ordered (Nickerson & McClelland, 1984).

The integration rules of two additional respondents were classified as “conjunctive/additive over two consequences.” The preference rank orderings of the remaining 34 respondents did not satisfy the axioms of any of the proposed integration rules. These respondents were eliminated from subsequent analysis because the derivation of values requires the validation of an integration rule.

### Values

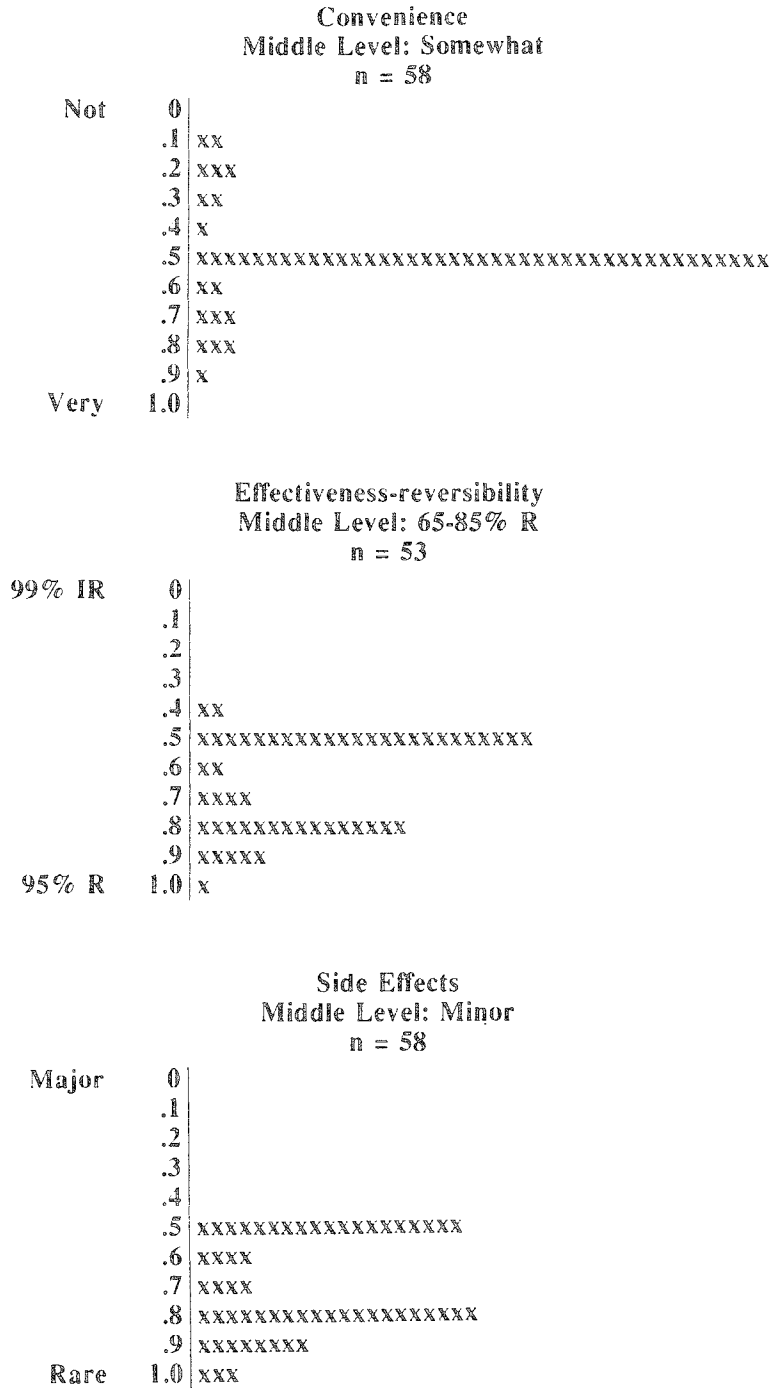
Values may vary across respondents in several different ways. First, respondents may vary in their preference orderings of the levels of the contraceptive consequences. In this study, all 66 respondents for whom integration rules could be determined favored very convenient over somewhat inconvenient over inconvenient contraceptive alternatives, and rare over minor over major side effects (given that the convenience and side effects consequences are essential). Of the respondents for whom effectiveness/reversibility was essential, 88% preferred alternatives that are 95% effective and reversible over those that are 65% to 85% effective and reversible over those that are 99% effective and irreversible, with the remaining respondents having a different preference ordering of these levels. This lack of variation in the preference orderings of the consequence levels is probably due to

the particular consequences included in the study. It is easy to imagine other contraceptive consequences where great individual variation might be expected, for example, the consequence *responsibility*—which partner is responsible for obtaining and using the contraceptive method—with possible levels “male,” “female,” and “both.”

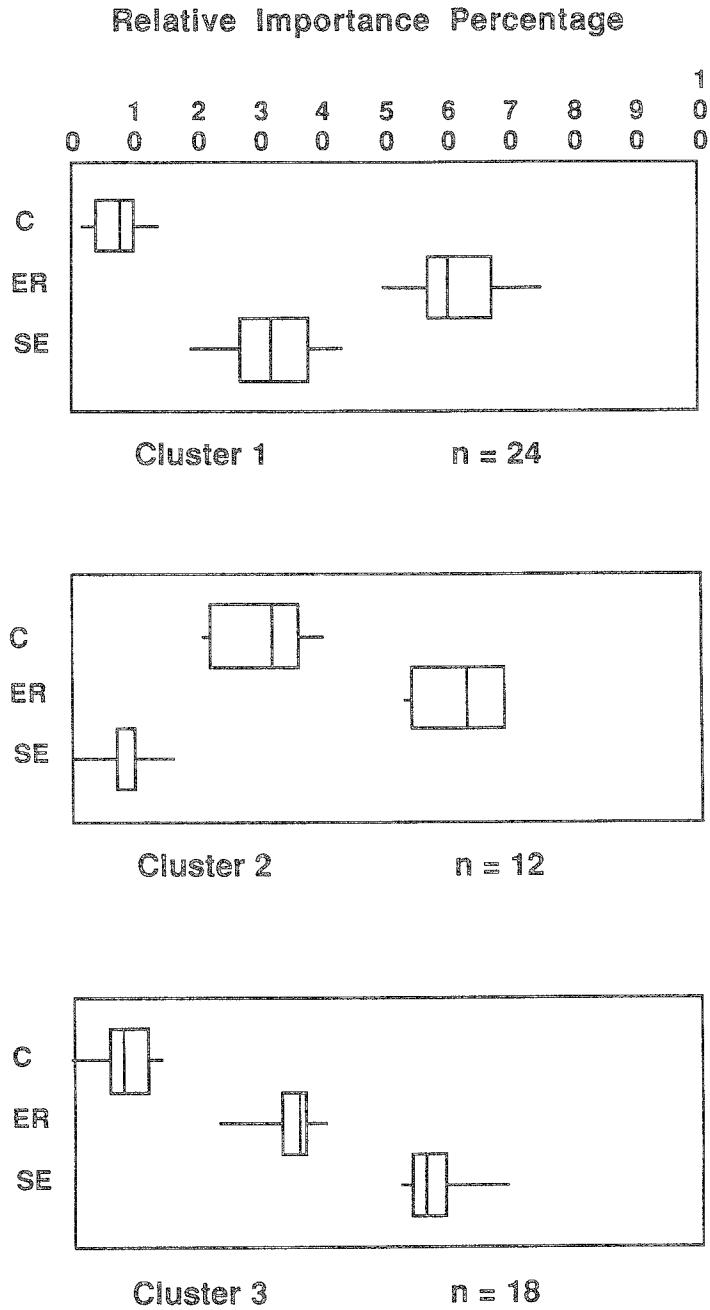
Second, respondents may vary in the values they assign to the levels of the consequences. Comparison of these values across respondents is simplified if they are rescaled so that the most and least preferred levels of each consequence equal 1.0 and 0.0, respectively, in order to establish a common scale across all respondents. Individual differences in the valuation of the levels of a consequence then appear as spread in the distribution of the values of the “middle” level of the consequence—the middle level being the one that is neither most nor least preferred. The middle-level value distributions for convenience, side effects, and effectiveness/reversibility are shown in Figure 2. (Respondents indifferent to a consequence are omitted from the distribution for that consequence. Only respondents who preferred the 95% effective and reversible alternatives over the 65% to 85% effective and reversible alternatives over the 99% effective and irreversible alternatives are included in the distribution for effectiveness/reversibility.) Examination of the three distributions indicates great individual variation in the valuation of the middle levels of the three consequences. Furthermore, this variation is not due to a few idiosyncratic respondents; the distributions for effectiveness/reversibility and side effects are bimodal, and nearly one-third of the respondents are not at the single mode of the distribution for convenience.

Third, respondents may also differ in their valuations of the contraceptive consequences themselves. A cluster analysis (BMDP2M; Dixon & Brown, 1977) of the relative importance percentages of the three contraceptive consequences showed that values varied widely for the 66 respondents for whom integration rules could be determined. The cluster analysis produced three primary and six smaller clusters of respondents. The three primary clusters are depicted by the “box-and-whisker” plots (Tukey, 1977, pp. 27–56) in Figure 3. In these plots

Figure 2  
 Frequency Distributions of the Middle-Level Values  
 of the Three Contraceptive Consequences



**Figure 3**  
 Box-and-Whisker Plots of the Relative Importance Percentages  
 of the Three Contraceptive Consequences  
 for the Three Largest Clusters of Respondents





the endpoints of the "whiskers" represent the minimum and the maximum relative importance percentages of a consequence in a given cluster, the vertical line represents the median, and the "box" encloses the interquartile range.

Respondents in the largest cluster ( $n = 24$ ) are moderately concerned about effectiveness/reversibility (median relative importance percentage of 60%), somewhat concerned about side effects (32%), and only slightly concerned about convenience (8%). Respondents in a second cluster ( $n = 12$ ) also show moderate concern about effectiveness/reversibility (64%), but are somewhat concerned about convenience (33%) and only slightly concerned about the risk of side effects (8%). Respondents in the third cluster of respondents ( $n = 18$ ) are moderately concerned about side effects (56%), somewhat concerned about effectiveness/reversibility (36%), and only slightly concerned about convenience (8%). Three of the six smaller clusters contain only respondents for whom only one consequence is essential, and the other three each contain a single idiosyncratic respondent. It is important to note that the relative importance percentages of the three contraceptive consequences are homogeneous within clusters but heterogeneous across clusters, evidence of the existence of large individual differences in values.

### Beliefs

The median gamma correlation coefficient for the 66 respondents for whom integration rules can be determined was .65. Standard tests of significance are not appropriate for within-persons analyses; therefore a gamma of .50 was arbitrarily selected to be "acceptable." The rationale for this selection is that this value assures that at least 75% of all possible pair comparisons (between the ranking of the real birth control methods and the overall evaluations of the hypothetical contraceptive alternatives matched to those methods) had the same ordering or are tied. Of the 66 respondents, 40 (61%) had acceptable gamma correlation coefficients, with a median of .75. If it is assumed that respondents used the same integration rule and val-

ues in the real birth control ranking task as they did in the hypothetical contraceptive ranking task, then the results of the gamma analyses indicate that the consequences and consequence levels employed in this study accurately describe beliefs about real birth control methods for a majority of the respondents for whom integration rules can be determined.

Table 7 presents the beliefs about the eight real birth control methods for the 40 respondents with acceptable gamma coefficients. It shows that there are indeed individual differences in respondents' beliefs about the consequences associated with the methods. Specifically, respondents disagreed in their judgments of the convenience of the IUD, pill, condom, diaphragm, foam/cream/jelly and rhythm, but generally agreed on the convenience of tubal ligation and vasectomy. There was nearly unanimous agreement on the effectiveness/reversibility of six of the eight real birth control methods; disagreement occurred only for the condom and the diaphragm. Respondents generally agreed in their judgments of the side effects of the IUD, pill, condom, and rhythm, but disagreed on the side effects of tubal ligation, vasectomy, diaphragm, and foam/cream/jelly.

It should be noted that the percentages reported in Table 7 would not vary substantially if based upon the beliefs of all 100 respondents, rather than on the beliefs of just the 40 respondents with acceptable gamma correlations. The important point is that for these latter respondents, the consequences and consequence levels used in this study accurately describe beliefs about real birth control methods. They do not do so for the remaining 26 respondents for whom integration rules can be determined. The validity of the beliefs of respondents for whom integration rules cannot be determined cannot be tested.

### Discussion

#### Individual Differences

The results of this study clearly show that there are substantial individual differences in contraceptive beliefs, values, and integration rules even in

Table 7  
 Percentages of the 40 Respondents With Acceptable Gammas  
 for Whom a Consequence Was Essential, Who Described Each Real  
 Birth Control Method in Terms of a Particular Level of Each Consequence

| Birth Control Method | Convenience |          |     | Effectiveness/<br>Reversibility |       |          | Side Effects |       |       |
|----------------------|-------------|----------|-----|---------------------------------|-------|----------|--------------|-------|-------|
|                      | Very        | Somewhat | Not | 99% IR                          | 95% R | 65-85% R | Rare         | Minor | Major |
| Tubal ligation       | 79          | 9        | 12  | 100                             | 0     | 0        | 41           | 29    | 30    |
| Intrauterine device  | 49          | 45       | 6   | 0                               | 92    | 8        | 0            | 12    | 88    |
| Vasectomy            | 79          | 12       | 9   | 95                              | 5     | 0        | 50           | 32    | 18    |
| Pill                 | 39          | 49       | 12  | 0                               | 100   | 0        | 0            | 3     | 97    |
| Condom               | 12          | 42       | 46  | 0                               | 23    | 77       | 100          | 0     | 0     |
| Diaphragm            | 3           | 70       | 27  | 0                               | 67    | 33       | 71           | 29    | 0     |
| Foam/cream/jelly     | 0           | 48       | 52  | 0                               | 5     | 95       | 59           | 38    | 3     |
| Rhythm               | 21          | 24       | 55  | 0                               | 5     | 95       | 94           | 6     | 0     |

a small, relatively homogeneous sample of respondents. Although the respondents in this study either were contraceptive users or intended to become users after their visits to the family-planning clinic, no effort was made to obtain a sample of respondents representative (in terms of sociodemographic characteristics, etc.) of contraceptive users in general; hence, the substantive results (e.g., the precise cluster definitions and the proportion of respondents in each cluster) may not be generalizable. However, there is no reason to expect that the degree of heterogeneity found in this study would not be typical of a larger, more representative sample.

Individual differences are important for their methodological and theoretical implications. For example, many if not most studies employing a behavioral model that incorporates an integration rule assume that that rule is an additive one and do not test it. It is also assumed that the integration rule is additive over all variables in the model. As the present study shows, these assumptions are not warranted. The integration rules of only one-third of the respondents were additive over all three consequences. The integration rules were additive variants for another third, and for the remaining third, no integration rules could be determined.

The assumption of integration rule additivity can have considerable impact on the results of a study. For example, the "part-worth" preference model

used by consumer researchers usually assumes an additive integration rule. This model is often tested with conjoint analysis (e.g., Green & Srinivasan, 1978; Green & Wind, 1973), which is not axiomatic conjoint measurement but a numerical scaling technique. Conjoint analysis does not test axioms but instead searches (usually through an iterative scaling algorithm such as MONANOVA; Kruskal, 1965) for an appropriate monotonic transformation of the rank ordering that best satisfies the assumed additive rule; then it evaluates the correspondence between the original and the transformed data (i.e., the overall evaluations) with a goodness-of-fit measure. It is generally believed that this measure will indicate whether the assumption of an additive integration rule is valid. This belief is incorrect. Nickerson and McClelland (1984) have demonstrated that it is possible to obtain a value for the goodness-of-fit measure that indicates that the rank ordering is *perfectly* consistent with an additive integration rule, even though that rule is in fact an interactive or a conjunctive/additive one.

It is also generally believed that even if the integration rule is not additive, conjoint analysis will nevertheless achieve "good prediction" if there are monotonic preference functions over increasing levels of the consequences, as is often the case in consumer research studies (Green & Srinivasan, 1978). However, the criterion being used for good

prediction is the value of the goodness-of-fit measure. It is possible for conjoint analysis to produce a goodness-of-fit measure that indicates very good or even perfect prediction, even though the scaling is extremely distorted (Nickerson & McClelland, 1984). This phenomenon occurs because scaling algorithms tend to tie the overall evaluations of alternatives in order to attenuate violations of additivity. Tying maintains the monotonicity necessary to maximize a goodness-of-fit measure, while simultaneously distorting the values calculated for the levels of the consequences (Srinivasan & Shocker, 1973). The tendency of scaling algorithms to produce these degenerate solutions suggests that their use should always be preceded by axiom testing procedures (Nickerson & McClelland, 1984; Nygren, 1985).

The existence of individual differences has practical as well as methodological and theoretical implications. For example, the diversity in beliefs suggests that inaccurate information—nonveridical beliefs about the consequences of available birth control methods—may be a significant determinant of contraceptive choice for some persons. In this study, 30% of the respondents believed that use of tubal ligation risks major side effects even though such effects are extremely uncommon (Saidi & Zainie, 1980). The influence of incorrect information is to make decisions suboptimal because the person will not achieve the desired goal by using the contraceptive method selected. Family planning counselors and contraceptive advocates may therefore find that providing the target population with accurate information about particular contraceptive methods can be as effective in modifying fertility behavior as attempts to change values.

#### Usefulness and Validity of the Model and the Methodology

Because there *are* individual differences in contraceptive beliefs, values, and integration rules, it is important to use a model and a methodology that can accommodate this diversity. The McClelland (1980) decision-making model, which allows for and even expects individual differences, and the

extended axiomatic conjoint measurement methodology, which depends on within-persons analytic techniques, were proposed as a means whereby this diversity might be studied. In this study the model accurately described the contraceptive decision-making process for 40 of the 100 respondents (40%). This percentage, although not extremely high, seems promising in light of the facts that (1) this is the first application of the model and its methodology to fertility-related decision making, (2) some respondents apparently found the tasks rather difficult, (3) beliefs about the consequences of real birth control methods were restricted to one of only three levels of only three contraceptive consequences, and (4) quite stringent criteria were adopted for integration rule classification and beliefs validation.

A major strength of the decision-making model and the extended axiomatic conjoint measurement methodology is the ability to pinpoint, for each respondent for whom the model cannot be validated, the stage in the decision process where the model first fails. The sequential method of analysis determines the integration rule, derives values, and evaluates the accuracy of the beliefs in describing the real birth control methods. If the entire analysis can be completed for a respondent, the decision-making model is judged to describe accurately the contraceptive decision-making process of that respondent. If the analysis reveals a failure of the model at a particular stage of the decision process, the respondent is eliminated from further analysis at that stage. Tabulation of these failures across respondents then gives an assessment of the overall success of the model in describing the contraceptive decision making of the respondents.

For example, in the present study 34 of the 100 respondents were eliminated because their integration rules could not be classified within the limits of the adopted error criterion (and therefore their values could not be derived), and 26 of the remaining respondents were eliminated because their gamma correlation coefficients showed that the consequences and consequence levels used in this study did not accurately describe their beliefs about real birth control methods. In a study using the typical cross-persons methodology to investigate

behavioral choice, it is not at all clear where the problem lies if a low correlation is obtained between the beliefs and values and the behavioral choice. It is not possible, for instance, to determine whether the low correlation results from respondents' use of some integration rule other than the assumed additive one, or from problems with the measurement of beliefs and values.

The inability to determine integration rules for one-third of the respondents in this study does not imply inadequacy of the decision-making model or the methodology per se; rather it is an empirical result, of interest in and of itself. Close examination of the preference rank orderings of these respondents shows that a few of these orderings exhibit substantial regularity but would require the correction of more than the three allowable errors in order to satisfy the axioms of any of the proposed integration rules. The rank orderings of the majority, however, generated huge numbers of axiom violations. Several of these rank orderings appeared to be random or very nearly so. The others fell into one of two patterns. In the first pattern, the several most preferred alternatives were ordered as if consistent with an additive integration rule, but the ordering of the rest of the alternatives was random or very nearly so. The second pattern resembled the first, except that the nine least preferred alternatives—all nine described by the least desirable level of one consequence—were set aside as in a conjunctive/additive rule, and only the middle alternatives—those neither most nor least preferred—were randomly ordered.

There are several possible explanations for these preference rank orderings:

1. Comments by a few respondents indicate that they found the task too long or too difficult, primarily because of the long verbal descriptions used to characterize the hypothetical contraceptive alternatives. Subsequent research has shown that it is possible to ameliorate this problem by using brief code words instead of long verbal descriptions to characterize the alternatives.
2. It is possible that some respondents did not take the task seriously and therefore responded

almost randomly. It is well known that "real people," such as the family-planning clinic clients serving as respondents in the present study, often perform less well than the more usual college-student respondents.

3. The decision-making model assumes that a fertility-related behavior is the result of deliberative choice. However, there are certainly some respondents for whom the use of a contraceptive method does not involve a decision. For example, a person may simply accept the contraceptive method prescribed by her physician or may request the contraceptive method used by a friend or a sister without giving the matter much, if any, thought. The decision-making model cannot be expected to describe the decision-making process of a person if no decision is being made. The ranking tasks would not make much sense to these respondents.
4. Although the decision-making model assumes a causal direction from beliefs and values to choices, the possibility exists that a respondent might select a contraceptive method and then justify that choice by adjusting his or her beliefs and values to match it. It is highly unlikely that such a respondent could maintain the consistency of beliefs and values required by the axiomatic conjoint measurement methodology.
5. It may be that respondents know what they like, and perhaps also what they dislike, but have difficulty selecting from among the mediocre alternatives in between, all of which are characterized by a less-than-desirable level of at least one consequence. This may be particularly true for respondents who have not given the matter of selecting a birth control method much thought.
6. The effectiveness/reversibility consequence actually consists of two combined consequences, effectiveness and reversibility. It is well known that correlated consequences increase inconsistency and random errors in decision making (Einhorn, Kleinmuntz, & Kleinmuntz, 1979). For many persons, the effectiveness and reversibility consequences are un-

doubtedly negatively correlated; the most preferred level of effectiveness is combined with the least preferred level of reversibility, and vice-versa. It is possible that this correlation caused difficulty for some respondents, even though the consequence was presented as a single consequence rather than as two separate consequences.

Any one of these situations can explain the lack of regularity in the preference rank orderings of the 34 respondents for whom integration rules cannot be determined. It must be emphasized, however, that these explanations are speculative only. No definitive explanation for the large numbers of axiom violations in the rank orderings can be obtained from examination of the rank orderings themselves. Einhorn et al. (1979) have suggested that questions regarding persons' cognitive processes during decision making may be answerable if the formal decision-making tasks are supplemented with process-tracing procedures such as concurrent or retrospective verbal protocols (Ericsson & Simon, 1980; Newell & Simon, 1972). It is recommended that future applications of the extended axiomatic conjoint measurement methodology include these procedures.

Another strength of the funnel-like analytic procedure used in the present study is that it suggests where modifications need to be made in the study design. Thus, the inability of the consequences and consequence levels to capture 26 respondents' beliefs about real birth control methods indicates that other consequences and/or consequence levels need to be used in addition to or in place of those used in the present study. For example, some persons may consider important the means by which a contraceptive method is obtained, that is, whether it can be purchased at a drugstore, whether it requires a visit to a physician or a clinic, or whether it requires surgery.

Although it seems likely that a respondent considers no more than about four consequences when selecting a contraceptive method because of cognitive limitations, a study should include all of the consequences actually considered by a respondent and the consequence levels believed possible for

each consequence. The factorial combination of more than a few consequences would create a very large number of behavioral alternatives, resulting in a task that would be too burdensome for most respondents. However, axiomatic conjoint measurement permits the use of fractional factorial designs as a means of reducing the number of behavioral alternatives to be ranked. Alternatively, the task materials might be tailored for each individual respondent. That is, each respondent would first select, from a large set of possible consequences, the few that he or she considers most important in selecting a birth control method, and perhaps also indicate possible levels of each consequence. The materials for the judgment task and the hypothetical contraceptive ranking task would then be based upon only those consequences and consequence levels selected by the respondent.

#### Comparison With Expectancy-Value Models

The fertility decision-making model is obviously very similar to the class of behavior models known as "expectancy-value" or "expectancy-valence" models. These models have been used extensively by psychologists to study a wide variety of behavioral phenomena (Feather, 1982; Lawler, 1971). Expectancy-value models all theorize that the strength or "force" of a person's tendency to act in a certain way is an additive function of the strength of his or her expectancies or beliefs that the act or behavior will be followed by given consequences or outcomes multiplied by the desirabilities, values, or valences of those consequences or outcomes:

$$F_i = \sum_{j=1}^n (E_{ij}V_j) \quad (1)$$

where  $F_i$  is the force on the person to perform act  $i$ ,

$E_{ij}$  is the strength of the expectancy or belief that act  $i$  will be followed by outcome  $j$ ,

$V_j$  is the valence or value of outcome  $j$ , and

$n$  is the number of outcomes.

Expectancy-value theories assume that persons select from among alternative courses of action the one with the greatest force (Vroom, 1964).

The fertility decision-making model differs from expectancy-value models in that (1) it tests rather than assumes an integration rule, (2) it does not restrict the integration rule to being an additive one, and (3) it does not multiply beliefs (expectancies) by values (valences). These are not trivial differences. The results of the present study show that the assumption that all persons use the same (additive) integration rule is not warranted. When some persons employ the assumed integration rule but others do not, tests of the model may not accurately reflect its validity. Schmidt (1973) has criticized the multiplication operation in expectancy-value models because it presumes that the measures of expectancies and values are ratio scales; these measures (usually five- or seven-point Likert rating scales) are interval scales at best. The result is that tests of the model are both logically and empirically ambiguous. The methodology developed in this paper avoids these two problems.

The most important difference between the fertility decision-making model and the expectancy-value models, however, is a methodological rather than a theoretical one. Both the decision-making model and the expectancy-value models are models of the behavior of an *individual*. Both theorize that a person is most likely to perform the behavior towards which he or she holds the most favorable beliefs and values; prediction as to whether a person will perform a given behavior must depend therefore on a comparison of his or her beliefs and values regarding that behavior with his or her beliefs and values regarding alternative competing behaviors. Tests of these models quite clearly require a within-persons method of analysis.

The extension of axiomatic conjoint measurement used here to study contraceptive choice is a within-persons methodology. The methodology used in nearly all expectancy-value studies, however, has been an across-persons one (Mitchell, 1974). For example, an expectancy-value model known as the "Fishbein model" (Fishbein, 1980; Fishbein & Ajzen, 1975) has often been used to study contraceptive choice (e.g., Cohen, Severy, & Ahtola, 1978; Davidson & Jaccard, 1975; Fishbein & Jaccard, 1973; Jaccard & Davidson, 1972; Werner & Middlestadt, 1979). Typically the Fishbein model

is tested by obtaining rating scale measures of beliefs and values regarding a single contraceptive behavior, performing the indicated algebraic operations on the ratings, and then correlating across respondents the value of  $F$  (known in the Fishbein model as behavioral intention or BI) predicted by the model with a separate rating scale measure of BI. For example, Davidson and Jaccard (1975) used the Fishbein model to predict respondents' intentions to use birth control pills.

Note that the assumption underlying this regression-based across-persons methodology is that the persons most likely to perform a behavior are those whose beliefs and values are most favorable toward that behavior; prediction as to whether a person will perform a given behavior therefore depends not upon a comparison of his or her beliefs and values regarding that behavior with a comparison of his or her beliefs regarding possible alternative behaviors, but upon a comparison of his or her beliefs and values regarding that behavior with the beliefs and values of other persons regarding that behavior. Thus, although the theoretical model clearly dictates that comparisons should be made across multiple behavioral alternatives for one person, the methodology makes comparisons across multiple persons for one behavioral alternative.

There is a discrepancy, then, between expectancy-value models and the regression-based methodology almost invariably used to test them; the methodology does not in fact provide a valid test of those models. Furthermore, the methodology is clearly an inappropriate one for the explanation and prediction of the behavior of *persons*. Why, then, has this methodology been used?

Undoubtedly, the primary reason why an across-persons methodology has been used to test within-persons models is that few researchers have recognized that a model/methodology discrepancy exists (Mitchell, 1974). However, the methodology is also used by researchers who are cognizant of the discrepancy, probably because within-persons analyses are viewed as "impractical" (Mitchell, 1974). Adapting the regression-based across-persons methodology to a within-persons methodology requires that rating scale measures of beliefs and values be collected for each level of each con-

sequence for each behavioral alternative under consideration. Thus, 144 responses would be required from each respondent in order to collect the same information obtained from the judgment task and the hypothetical contraceptive ranking task in the present study. Collection of these responses would be tedious for the respondents and their analysis cumbersome for the researcher. The extension of axiomatic conjoint measurement proposed here as a solution to a methodological dilemma encountered in the study of fertility-related decision making should also provide a practical, workable alternative to a regression-based within-persons methodology for testing expectancy-value models.

### References

- Anderson, B. F., Deane, D. H., Hammond, K. R., McClelland, G. H., & Shanteau, J. C. (1981). *Concepts in judgment and decision research: Definitions, sources, interrelations, comments*. New York: Praeger.
- Arbuckle, J., & Larimer, J. (1976). The number of two-way tables satisfying certain additivity axioms. *Journal of Mathematical Psychology*, *13*, 89–100.
- Basics of birth control*. (1977). New York: Planned Parenthood Federation of America.
- Birth control handbook*. (1973). Montreal: Montreal Health Press.
- Boston Women's Health Book Collective. (1976). *Our bodies, ourselves*. New York: Simon & Schuster.
- Bourgeois-Pichat, J. (1967). Social and biological determinants of human fertility in nonindustrial societies. *Proceedings of the American Philosophical Society*, *111*, 160–163.
- Cohen, J. B., Severy, L. J., & Ahtola, O. T. (1978). An extended expectancy-value approach to contraceptive alternatives. *Journal of Population*, *1*, 22–41.
- Coombs, C. H., & Bowen, J. N. (1971). A test of VE-theories of risk and the effect of the central limit theorem. *Acta Psychologica*, *35*, 15–28.
- Coombs, C. H., Coombs, L. C., & McClelland, G. H. (1975). Preference scales for number and sex of children. *Population Studies*, *29*, 273–298.
- Coombs, C. H., & Huang, L. C. (1970). Polynomial psychophysics of risk. *Journal of Mathematical Psychology*, *7*, 317–338.
- Coombs, L. C. (1979). The measurement of commitment to work. *Journal of Population*, *2*, 203–223.
- Davidson, A. R., & Jaccard, J. J. (1975). Population psychology: A new look at an old problem. *Journal of Personality and Social Psychology*, *31*, 1073–1082.
- de Leeuw, J., Young, F. W., & Takane, Y. (1976). Additive structure in qualitative data: An alternating least squares method with optimal scaling features. *Psychometrika*, *41*, 471–503.
- Delbeke, L. L. A. (1979). Variables and methods of measurement. In J. A. Michon, E. G. J. Eijkman, & L. F. W. de Klerk (Eds.), *Handbook of psychonomics* (Vol. 1, pp. 23–97). Amsterdam: North-Holland.
- Dixon, W. J., & Brown, M. B. (Eds.). (1977). *BMDP biomedical computer programs, P series*. Berkeley CA: University of California Press.
- Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process tracing models of judgment. *Psychological Review*, *86*, 465–485.
- Emery, D. R. (1977a). DIST: A numerical conjoint measurement program designed to scale data to a distributive model in three dimensions. *Journal of Marketing Research*, *14*, 413–414.
- Emery, D. R. (1977b). DULST: A numerical conjoint measurement program designed to scale data to a dual-distributive model in three dimensions. *Journal of Marketing Research*, *14*, 558–559.
- Emery, D. R. (1978). *Optimal scaling via ordinary least squares* (Working Paper No. WP-31-78). Calgary, Alberta: University of Calgary, Faculty of Business.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.
- Falmagne, J. C. (1978). Probabilistic choice behavior theory: Axioms as constraints in optimization. In J. N. Castellan, Jr., & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 93–104). Hillsdale NJ: Erlbaum.
- Feather, N. T. (1982). *Expectations and actions: Expectancy-value models in psychology*. Hillsdale NJ: Erlbaum.
- Fischer, G. W. (1976). Multidimensional utility models for risky and riskless choice. *Organizational Behavior and Human Performance*, *17*, 127–146.
- Fishbein, M. (1980). A theory of reasoned action: Some applications and implications. In M. M. Page & H. E. Howe, Jr. (Eds.), *Nebraska Symposium on Motivation, 1979* (Vol. 27, pp. 65–116). Lincoln NE: University of Nebraska Press.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading MA: Addison-Wesley.
- Fishbein, M., & Jaccard, J. J. (1973). Theoretical and methodological considerations in the prediction of family planning intentions and behavior. *Representative Research in Social Psychology*, *4*, 37–51.
- Goodman, L. A., & Kruskal, W. A. (1954). Measures of association for cross-classifications. *Journal of the American Statistical Association*, *49*, 732–764.
- Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, *5*, 103–123.
- Green, P. E., & Wind, Y. (1973). *Multivariate decisions in marketing: A measurement approach*. Hinsdale IL: Dryden Press.

- Holt, J. O., & Wallsten, T. S. (1974). *A user's manual for CONJOINT: A computer program for evaluating certain conjoint-measurement axioms* (Technical Report 42). Chapel Hill NC: University of North Carolina, L. L. Thurstone Psychometric Laboratory.
- Jaccard, J. J., & Davidson, A. R. (1972). Toward an understanding of family planning behaviors: An initial investigation. *Journal of Applied Social Psychology*, 2, 228-235.
- Johnson, R. M. (1975). A simple method for pairwise monotone regression. *Psychometrika*, 40, 163-168.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). New York: Academic Press.
- Krantz, D. H., & Tversky, A. (1971). Conjoint-measurement analysis of composition rules in psychology. *Psychological Review*, 78, 151-169.
- Kruskal, J. B. (1965). Analysis of factorial experiments by estimating monotone transformations of the data. *Journal of the Royal Statistical Society, Series B*, 27B, 251-263.
- Lawler, E. E., III. (1971). *Pay and organizational effectiveness: A psychological view*. New York: McGraw-Hill.
- Lehner, P. E., & Noma, E. (1980). A new solution to the problem of finding all numerical solutions to ordered metric structures. *Psychometrika*, 45, 135-137.
- McClelland, G. H. (1977). A note on Arbuckle and Larimer, "The number of two-way tables satisfying certain additivity axioms". *Journal of Mathematical Psychology*, 15, 292-295.
- McClelland, G. H. (1980). A psychological and measurement theory approach to fertility decision-making. In T. K. Burch (Ed.), *Demographic behavior: Interdisciplinary perspectives on decision-making* (pp. 141-161). Boulder CO: Westview Press.
- McClelland, G. H., & Coombs, C. H. (1975). ORDMET: A general algorithm for constructing all numerical solutions to ordered metric structures. *Psychometrika*, 40, 269-290.
- Mitchell, T. R. (1974). Expectancy models of job satisfaction, occupational preference, and effort: A theoretical, methodological, and empirical appraisal. *Psychological Bulletin*, 81, 1053-1077.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs NJ: Prentice-Hall.
- Nickerson, C. A., & McClelland, G. H. (1984). Scaling distortion in numerical conjoint measurement. *Applied Psychological Measurement*, 8, 183-198.
- Nygren, T. E. (1985). An examination of conditional violations of axioms for additive conjoint measurement. *Applied Psychological Measurement*, 9, 249-264.
- Nygren, T. E. (1986). A two-stage algorithm for assessing violations of additivity via axiomatic and numerical conjoint analysis. *Psychometrika*, 51, 483-491.
- Roskam, E. E. (1974). *Unidimensional conjoint measurement (UNICON) for multi-faceted designs*. Nijmegen, The Netherlands: University of Nijmegen, Psychologisch Laboratorium.
- Saidi, M. H., & Zainie, C. M. (1980). *Female sterilization: A handbook for women*. New York: Garland.
- Schmidt, F. L. (1973). Implications of a measurement problem for expectancy theory research. *Organizational Behavior and Human Performance*, 10, 243-251.
- Srinivasan, V., & Shocker, A. D. (1973). Linear programming techniques for multidimensional analysis of preferences. *Psychometrika*, 38, 337-369.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading MA: Addison-Wesley.
- Ullrich, J. R., & Cummins, D. E. (1973). PCJM: A program for conjoint measurement analysis of polynomial composition rules. *Behavioral Science*, 18, 226-227.
- Ullrich, J. R., Cummins, D. E., & Walkenbach, J. (1978). PCJM2: A program for the axiomatic conjoint measurement analysis of polynomial composition rules. *Behavior Research Methods and Instrumentation*, 10, 89-90.
- Ullrich, J. R., & Painter, J. R. (1974). A conjoint-measurement analysis of human judgment. *Organizational Behavior and Human Performance*, 12, 50-61.
- Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.
- Wallsten, T. S. (1972). Conjoint-measurement framework for the study of probabilistic information processing. *Psychological Review*, 79, 245-260.
- Wallsten, T. S. (1976). Using conjoint-measurement models to investigate a theory about probabilistic information processing. *Journal of Mathematical Psychology*, 14, 144-185.
- Wallsten, T. S., & Budescu, D. V. (1981). Additivity and nonadditivity in judging MMPI profiles. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1096-1109.
- Werner, P. D., & Middlestadt, S. E. (1979). Factors in the use of oral contraceptives by young women. *Journal of Applied Social Psychology*, 9, 537-547.
- Young, F. W. (1972). A model for polynomial conjoint analysis algorithms. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling: Theory and applications in the behavioral sciences* (Vol. 1, pp. 69-104). New York: Seminar Press.

#### Acknowledgments

This research was supported in part by grants HD-10802 and HD-14403 from the National Institute of Child Health and Human Development. The authors gratefully acknowledge the technical assistance of Perle Bochert,



*Susan Kegeles, Michael O'Reilly, Robbie Toney, and Doreen Victor, and the many helpful editorial suggestions of Alice Healy and Lou McClelland. This paper is a revision of CRJP Technical Report No. 234 (July, 1981).*

**Author's Address**

Send requests for reprints or further information to Gary H. McClelland, Center for Research on Judgment and Policy, Campus Box 344, Muenzinger Psychology Building, University of Colorado, Boulder CO 80309, U.S.A.