Psychol. Belg.
1987, XXVII-2, 123-141.

TESTING EXPECTANCY MODELS: A CONJOINT MEASUREMENT APPROACH

Rudi VAN HOE

University of Ghent

The term expectancy theories refers to a category of social cognitive theories that use the same underlying Subjective Expected Utility (SEU) model for describing goal directed behavior. However, the mathematical SEU formulation has in many expectancy theories a dubious validity. Expectancy theorists lack an adequate measurement model to test their formal theory in a valid way. The conjoint measurement approach is proposed as a possible solution for this problem. In two experiments we tested an expectancy model respectively by means of the UNICON scaling algorithm for conjoint measurement and by means of an axiomatic approach. Evidence was found for the functional form of the SEU model, which implies that the subjective probability (expectancy) and the utility of an outcome of an action combine multiplicatively while the resultant products combine additively. In the discussion, it is argued that only the descriptive validity of a model has been proven. More research is needed to identify the cognitive processes that can be described by addition and multiplication.

The Basis of Expectancy Theories: the SEU Decisionmodel

The term *expectancy theories* refers to a category of social cognitive theories that — in spite of some minor differences — use the same underlying conceptual Subjective Expected Utility (SEU) model for describing and predicting goal-directed behavior.

According to the SEU decision theory, whether a person will perform an action X from the set of alternative actions depends on two properties of the possible outcomes of action X: the subjective probability of the occurrence of the outcomes should X be performed, and the attractiveness of each outcome to the subject. The first property is usually called the *subjective probability* or *expectancy* of an outcome (*P*); the second is the *utility* (*U*) of an outcome. The subjective probabilities and utilities of all possible outcomes of action X define an unique numeric property for this alternative, namely the *subjective expected utility* (*SEU*) according to the equation:

$$\mathrm{SEU}_x = \sum_{i=1}^{n} P_i U_i \tag{1}$$

I wish to thank Geert De Soete, Karel Poupeye and André Vandierendonck for their encouragement and valuable remarks in preparing this article.

where P_i stands for the subjective probability of outcome i of action X and where U_i denotes the utility of that outcome. According to the SEU model the alternative X will be chosen when it is the alternative with the highest SEU. This choice principle is the SEU maximization principle.

Although some assumptions of the SEU decision theory have been criticized, there has been an implicit acceptance of the SEU model in several areas (e.g., Ajzen & Fishbein, 1969, 1980; Atkinson, 1957; Atkinson & Birch, 1970; Heckhausen, 1977; Raynor, 1974; Vroom, 1964; Weiner, 1985). These theories have in common that the SEU model is used to describe behavior in terms of constant parameters, namely direction (which goal among the possible goals a person will select) and intensity (the vigor with which a subject exerts himself toward the chosen goal).

Testing Polynomial Expectancy Theories

A polynomial is a linear combination of products of real variables raised to nonnegative, integral powers. Simple polynomials are polynomials that are constructed only by addition (subtraction) and/or multiplication (division) of the variables in question.

The expectancy theories can be considered as *polynomial theories* (De Soete, 1984), because the relation between the variables are represented by means of a simple polynomial. In particular, it is hypothesized that the subjective probability and utility of a given outcome of action combine *multiplicatively*, and the resultant products combine *additively* to define the SEU of this action (cf. Eq.(1)).

This polynomial representation of the information integration process has in the majority of the cases only a *would-be* validity. A lot of expectancy theorists lack an adequate measurement theoretical basis to test the formal theory and to measure the basic variables of the model in a valid manner.

Accordingly, one of the objectives of this article is to examine: (1) how one can test a formal expectancy theory in a methodological adequate manner; and (2) how one can obtain valid measurement scales of the basic variables in the polynomial model.

The choice of the measurement model is crucial. This is illustrated by a study of Sheluga, Jacoby and Jaccard (1979). These researchers compared estimates of utilities derived from an additive conjoint measurement model as opposed to those obtained by a direct rating approach. They observed relatively low levels of correspondence between the estimates derived from the two measurement methods.

Measurement Models

The multivariate and functional measurement approach (Anderson, 1962, 1981) are frequently used to test SEU-based models. However, these measurement models can be criticized for some methodological shortcomings. The most severe is that both approaches implicitly assume that one can measure all variables on at least an interval level. However, in most psychological research this assumption is not met (Krantz & Tversky, 1971a, 1971b). For example, Anderson interprets the absence of significant interaction effects as evidence for an additive model and for the solidity of the assumption that the response scale is an interval scale. However, if there are significant interaction effects, one is not sure whether this is due to the judgmental process itself, or to the nonlinearity of the response function (De Soete, 1980).

It is safer to take only the ordinal characteristics into account by using conjoint measurement (Krantz, Luce, Suppes & Tversky, 1971; Luce & Tuckey, 1964). Indeed, the main advantage of conjoint measurement is that formal theories can be tested without first measuring the variables of the model. Conjoint measurement relies only on qualitative data and it produces — if the formal theory is correct — the measurements as a result.

In the following experiments, the empirical validity of two expectancy models was tested. As a consequence, we illustrate the relevance of two different — but complementary — conjoint measurement approaches as tools for testing expectancy theories. The first is the *axiomatic* approach: by relying on the ordinal characteristics of the data, one searches for necessary and sufficient conditions (axioms) that the data must satisfy in order to be numerically representable according to the postulated polynomial model. Axioms are sometimes difficult to test empirically. In these cases, one can choose for the *algorithmic* approach. This approach utilizes algorithms for numerically fitting conjoint measurements models to a set of data. These algorithms allow to find the numerical representation with the greatest goodness-of-fit, according to some composition rule specified by the researcher (Roskam, 1982). If one finds a numerical representation that fits the data, the axioms of the polynomial model in question are tested simultaneously.

Experiment 1

In experiment 1 we tested the well known expectancy theory of motivation that considers goal directed behavior as a multiplicative combination of motives, expectancies and values:

$$\text{Response Tendency} = \text{Motive} \times \text{Expectancy} \times \text{Value}. \quad (2)$$

This idea, in particular the role of a motive as multiplier is not new. It is central in some of the already mentioned theories (cf. Atkinson, 1957; Heckhausen, 1977), but it was Hull (1943) who gave it a first explicit quantitative formulation:

$$\text{Reaction Potential} = \text{Drive} \times \text{Habit} \quad (3)$$

(Anderson, 1974).

Some experimental studies — using functional measurement — tested the polynomial model of Eq. (2) or a simplification of this model:

$$\text{Response Tendency} = \text{Expectancy} \times \text{Value}. \quad (4)$$

Klitzner & Anderson (1977) found evidence for Eq. (2). Empirical support for Eq. (4) was found in different research domains: Shanteau & Nagy (1976) in a study of physical attractiveness, Arnold (1976) in a test of Vroom's motivation model, Klitzner (1977) in a study of snake fear and Sjöberg (1982) in the context of the behavioral intentions model.

The cited theories and functional measurement studies suggest that the multiplicative model is a valid description of the motivational process. However, there are four classes of simple polynomials in three variables (Krantz & Tversky, 1971a; Krantz et al., 1971). The other alternatives are: the additive (linear) ($M + E + V$), the distributive ($M \times (E + V)$) and the dual-distributive ($M + E \times V$ or $M \times E + V$) model.

Our objective is to find the best fitting polynomial model and to obtain valid measurement scales of the variables by using algorithmic conjoint measurement. This approach is unique in that we can obtain — as a result of the test of the model — valid measurements of a subject variable such as motives. This is in contrast with, for example, Atkinson's (1957) procedure where the measurement of the achievement motive precedes the test of the model.

Method

Subjects. Nineteen first year students, nine males and ten females, all enrolled in the Faculty of Psychological and Educational Sciences at the University of Ghent, participated in the experiment in fulfillment of course requirements.

Design. The design consisted of (per stimulus set) a subject variable (M, 10 or 9 subjects), 4 levels of the expectancy variable P and 4 levels of the utility variable U. It concerned a within-subjects design: each stimulus from the product set $P \times U$ was presented to the subjects.

Stimuli. Four stimulus sets were developed. Each stimulus set consisted of descriptions of hypothetical goals, which were characterized by the attributes expectancy and utility. The expectancy levels for all stimulus sets were the same: almost uncertain, 25% chance, 50% chance and almost certain.

Stimulus set I contained descriptions of the goal "snake". The utility levels were: boa, cobra, rattlesnake and viper. The subjects had to indicate on a 15 point scale how frightening they would consider situations such as:

> You are with a boa in a room
> It is almost uncertain that the boa will come up to you.

It was assumed that the individual differences in snake fear would act as a multiplicative motivational factor (see also Klitzner, 1977). Stimulus set II consisted of descriptions of the goal "job". The utility levels were: salary 100,000 Belgian fr., salary 40,000 Belgian fr., unemployment pay and free food and drinks. The subjects had to evaluate stimuli as:

> If one performs his job successfully, one earns monthly a salary of 100,000 fr
> The probability that one can perform this job successfully is very small

in terms of how likely they would apply for the described job. Stimulus set III were descriptions of the goal "success on an exam". The utilities were: moving up to next year, felicitations of the professor, two points extra, getting a good idea what exams are all about. The subjects had to scale exam situations as:

> The chance on success on the exam is 50%
> If one succeeds, one will be moved up to next year

in terms of how likely they will join the exam. Stimulus set IV consisted of descriptions of the goal "car". The utilities were: sports car, station car, delivery car, city car. Stimuli such as:

> The car on sale is a sports car
> It is almost certain (financially) that you can buy the car

were judged in terms of how likely they will enter the car shop and ask more information about the car in question.

Procedure. Ten subjects received the stimulus sets I and II (5 subjects in the order: I and II, the other five in the order: II and I), the other nine subjects received the set III and IV (5 subjects: III and IV, the other: IV and III). Within each stimulus set, each stimulus was presented three times.

The order of the stimuli was randomized per subject. Thus, each subject received 48 stimuli per set, in total 96 stimuli (for the two sets). Each stimulus set was presented to the subjects with an "introduction". As an example, we give the introduction on set III (translated from Dutch into English):

> "The objective of this study is to form an idea of the way people judge several situations. The experiment consists of two parts. Each part contains several descriptions of a specific situation. Your task is to judge these descriptions. You have to imagine the following situation: it concerns an exam situation where one can decide if one joins the exam or not. You are not obliged to join the exam, the decision to participate is left to you. We ask you to scale each description of this situation in terms of how likely you will join the exam:
> -7 -6 -5 -4 -3 -2 -1 0 $+1$ $+2$ $+3$ $+4$ $+5$ $+6$ $+7$"

As already mentioned, each stimulus was presented three times. On the basis of the median of the three ratings of each stimulus we could construct an ordering of stimuli for each subject.

Results

Two subjects were removed from the analysis because they gave each stimulus a rating -7. It concerned two subjects of set III and IV. So the design of these sets contain 7 levels of the subject variable motive.

The ordinal data (cf. median values) were analysed with Roskam's algorithm for conjoint measurement UNICON (Roskam, 1974). With UNICON one can (1) obtain valid measurement scales of the expectancies, utilities and motives on the basis of the ordinal properties of the data and (2) compare the goodness-of-fit of several simple polynomials.

In general, a multiplicative model ($M \times E \times V$) is hypothesized. We computed also the goodness-of-fit of alternative polynomial models: the additive ($M + E + V$), the distributive ($(E + V) \times M$) and two dual-distributive models ($E \times M + V$ and $V \times M + E$).

The stress-value is in the algorithm the measure of the goodness-of-fit of the models. We choose .02 as critical stress-value. Table 1 contains for each stimulus set the stress-values of the tested polynomial models. The data show that the multiplicative model has the lowest stress-value within each stimulus set. In other words, the multiplicative model has clearly the best goodness-of-fit.

Tab. 1. — Stress-values of the polynomial models

polynomial model	Stimulus set I	II	III	IV
$M \times E \times M$	0.0955	0.0185	0.0178	0.0107
$M + E + M$	0.3977	0.5418	0.5410	0.4770
$M \times (E + V)$	0.3998	0.5029	0.5410	0.4816
$M \times E + V$	0.4465	0.5153	0.4889	0.4844
$M \times V + E$	0.3973	0.4027	0.3958	0.4280

Per set, we received also the values of the independent variables M, E and V according to the different conjoint measurement models.

Conclusion

The data clearly support the multiplicative model of Eq. (2). However, we tested only a restricted interpretation of the general expectancy model. A more fundamental test of the model would be a test of the functional implication of Eq. (1), namely that expectancy and utility combine multiplicatively and the resultant products combine additively. Therefore, in the next experiment we tested the bilinear version of Eq. (1).

EXPERIMENT 2

In experiment 2 we tested the bilinear expectancy model:

$$\text{Response Tendency} = P_1 \times U_1 + P_2 \times U_2. \tag{5}$$

A bilinear model is defined as a class in which the variables can be partitioned into two disjoint subsets such that if the variables are held constant in either one subset the function is linear in the other. The testable axioms for this model were developed by Coombs and Lehner (1981a, 1981b).

The bilinear expectancy model was already tested within the information integration theory. Shanteau (1974) found evidence for the assumption that the subjective probability and utility of a behavior outcome combine multiplicatively. On the other hand, he found violations of the assumption that the expectancy × utility products combine additively. These results were replicated by Lynch and Cohen (1978) within the context of helping behavior and by Lynch (1979) within the context of hypothetical bets. These studies indicate in general a *subadditive* effect.

An exception is the experimental study by Klitzner and Anderson (1977). They found that their data could be predicted by the following motivational model:

$$\text{Response Tendency} = M_1 \times P_1 \times U_1 + M_2 \times P_2 \times U_2. \qquad (6)$$

Studies within the context of multiple regression (Ajzen & Fishbein, 1969, 1980) and path analysis (Bentler & Speckart, 1979, 1981; Fredericks & Dossett, 1983), on the contrary, support the additive expectancy model.

In the present experiment, the expectancy variables P_1 and P_2 were interpreted respectively as the subjective probability of a successful outcome and the subjective probability of an unsuccessful outcome. The utilities U_1 and U_2 were respectively the attractiveness of the successful and unsuccessful outcome. This interpretation of the bilinear model is inspired by Coombs and Lehner's risk model (1981a, 1981b), where a difference is made between the "good" and "bad" outcomes and by Atkinson's achievement motivation model (1957), where the number of possible outcomes is limited to two, namely "success" and "failure". So, expression (5) can be restated as:

$$R(A) = \varnothing_1(P_1)\varnothing_2(U_1) + \varnothing_3(P_2)\varnothing_4(U_2) \qquad (7)$$

where A = achievement situation $\{(P_1, U_1), P_2, U_2)\}$ and $\varnothing_1 \neq \varnothing_3$.

Method

Subjects. Twenty six first year students, nine males and seventeen females, all enrolled in the Faculty of Psychological and Educational Sciences at the University of Ghent, participated in the experiment for course requirements.

Stimuli. The stimuli are presented in a matrix (cf. Table 2). The matrix contains two levels of the probability of success P_1 (.1 and .4) and two levels of the utility of a successful outcome U_1 (5 and 10). Their cartesian product generates the four columns of the matrix. The matrix contains also two levels of the probability of failure P_2 (.1 and .4) and two levels of the utility of an unsuccessful outcome U_2 (-5 and -10). Their cartesian product generates the rows of the matrix. Thus, the bilinear expectancy model was tested by means of a $2 \times 2 \times 2 \times 2$ factorial design presented in a 4×4 matrix.

Tab. 2. — The 4 × 4 basic matrix (stimulus set)

$P_2 \times U_2$ \ $P_1 \times U_1$	.1 5	.1 10	.4 5	.4 10
.1 −5	.1 5 .1 −5	.1 10 .1 −5	.4 5 .1 −5	.4 10 .1 −5
.1 −10	.1 5 .1 −10	.1 10 .1 −10	.4 5 .1 −10	.4 10 .1 −10
.4 −5	.1 5 .4 −5	.1 10 .4 −5	.4 5 .4 −5	.4 10 .4 −5
.4 −10	.1 5 .4 −10	.1 10 .4 −10	.4 5 .4 −10	.4 10 .4 −10

Procedure. The matrix contains sixteen cells. These stimuli were presented as sixteen different descriptions of an exam situation. For example, the cell: .1 10
.4 −5
was interpreted as the description of the following exam situation:

> You have one chance out of ten that you pass on the exam.
> If you pass the exam, you receive ten points extra.
>
> You have four chances out of ten that you fail the exam.
> If you fail the exam, you lose five points.

In order to obtain a number of replications per subject, the stimuli were arranged in a balanced incomplete blockdesign. The block design consisted of 24 blocks of six stimuli (cf. Cochran & Cox, 1957, plan 11.28, p. 478). The subjects were asked to rankorder stimuli within blocks. In this manner each stimulus was judged nine times and replicated three times. Per subject, the ordering of the presentations and the ordering of the stimuli within each block was randomized.

The task for the subjects was, per block, to rank order the six exam situations according to attractiveness (cf. SEU):

> "This research has as general goal to form an idea of how people judge situations. More specifically, we want to explore in what way first year students perceive exam situations. Your task will be to judge some descriptions of exam situations.
>
> On the next pages, you will find several descriptions of an exam situation.

The description of an exam situation has always the same form, for example:

You have seven chances out of ten that you pass the exam.
If you pass, you receive ten points extra.

You have three chances out of ten that you fail the exam.
If you fail, you lose five points.

Each page contains six such descriptions of an exam situation. Special to the exam situation is that you may choose whether you will or not take the exam. Namely, your task is to rank order, per page, the six exam situations according to your preference. I.e., the exam situation you like most receives number six; the exam situation that you prefer next receives number five; ... The exam situation you find the least attractive of all, receives the score one. Two or more exams of the same page cannot get the same rank number.

Take for each page your time, and thanks for the cooperation."

Results

Two subjects were removed from the analysis because they did not finish their task. For the remaining subjects the *stochastic dominant ordering* was identified. As each pair of stimuli is replicated three times, a stimulus x is said to dominate stochastically y, whenever x dominates y in more than one of the three comparisons of (x, y) (De Soete, 1980).

Independence axiom. The first step in the conjoint analysis of the ordinal data is testing *mutual independence* of $P_1 \times U_1$ (the "good" behavior outcome) and $P_2 \times U_2$ (the "bad" outcome). One can say that $P_1 \times U_1$ is independent of $P_2 \times U_2$, when the rank order of $P_1 \times U_1$ is the same at all levels of $P_2 \times U_2$. In the same way, $P_2 \times U_2$ is independent of $P_1 \times U_1$, if the rank order of $P_2 \times U_2$ is the same for all levels of $P_1 \times U_1$.

The ordinal data of 17 of the 24 subjects satisfied perfectly the independence of $P_1 \times U_1$ with respect to $P_2 \times U_2$. Fifteen of the 24 subjects satisfied the independence of $P_2 \times U_2$ with respect to $P_1 \times U_1$.

Direction of independence: consensus independence. On the ground of the bilinear expectancy model stated in Eq. (7), one can predict the direction of the independence of $P_1 \times U_1$ with respect to $P_2 \times U_2$: the larger the probability of success (P_1) the larger the attractiveness of the exam situation should be, all other things being held constant; and the more points one gets with a successful outcome (U_1) the larger the attractiveness should be, all other variables being held constant. Coombs and Lehner (1981b) talk about the *consensus independence* (*CI*) of $P_1 \times U_1$ with respect to $P_2 \times U_2$.

As a measure of the consensus independence of $P_1 \times U_1$ with respect to $P_2 \times U_2$ the percentage of *implicit* pairwise preference orderings was computed that satisfied the independence axiom. Per row of a matrix, there are five implicit pairwise comparisons, so 480 in total. In the present experiment, the percentage was equal to 98.4%.

In the same way, one can speak of the consensus independence of $P_2 \times U_2$ with respect to $P_1 \times U_1$: the smaller the probability of failure (P_2) the larger the attractiveness should be, all other things held constant; and the less points one can lose when failing (U_2) the larger the attractiveness should be, all other variables held constant. The percentage pairwise permutations that satisfied the independence of $P_2 \times U_2$ with respect to $P_1 \times U_1$ was 99.1%.

Further test of the model proceeds by distinguishing two test patterns: I and II. The analysis according to the test patterns I and II consists of a systematic decomposition of the 4×4 matrix into subsets of 2×2 matrices which permits a detailed examination of the bilinear model.

Test pattern I: identification of independence violations. Test pattern I contains the qualitative tests for the *associated* variables, namely the tests between P_1 and U_1 and between P_2 and U_2. The four levels of the row factor of the stimulus matrix generates the 2×2 matrix as illustrated in Figure 1. The same 2×2 matrix can be found in all four rows of the basic matrix, each row corresponds with a different fixed level of $P_2 \times U_2$. The direction of the consensus independence is in the 2×2 matrix illustrated with single — vertical and horizontal — arrows.

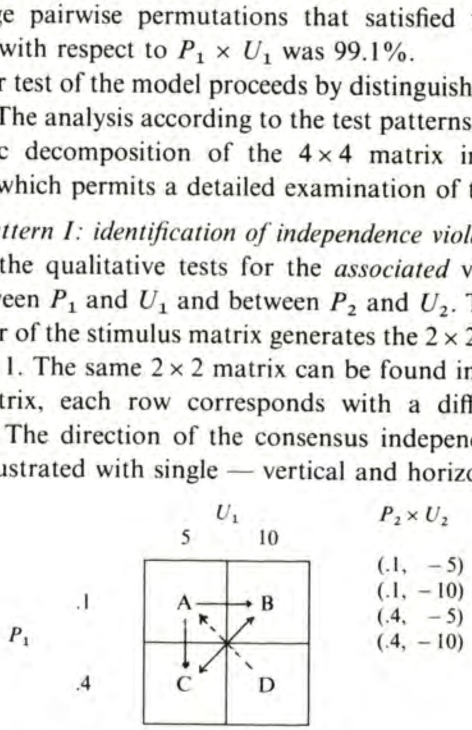

Fig. 1. — The 2×2 submatrices: the rows of the basic matrix.

Test pattern I is not so much a set of axioms, but rather a tool to identify the violations of the independence axiom and to evaluate the direction of the tradeoff.

Four different types of violations of the independence axiom are possible: (1) violations of the independence of P_1 with respect to U_1 (the variable P_1 is independent of U_1 when the rank order of P_1 is the same for all levels of U_1, with $P_2 \times U_2$ fixed on four different levels), (2) violations of the independence of U_1 with respect to P_1 (the factor U_1 is independent of P_1 when the rank order of U_1 is the same for all levels of P_1, with fixed $P_2 \times U_2$ levels), (3) inconsistency of the tradeoff between P_1 and U_1 and (4) violations of the transitivity.

The rank order of the four cells in a 2×2 matrix, which corresponds with a row of the 4×4 matrix, can be decomposed into six pair comparisons: two comparisons test the consensus independence of P_1 with respect to U_1 (cf. vertical arrows in Figure 1), two of them test the consensus independence of U_1 with respect to P_1 (cf. horizontal arrows in Figure 1), one comparison tests the direction of the tradeoff between P_1 and U_1 (cf. diagonal arrow between C and B in Figure 1) and one of them tests the consensus transitivity (cf. dashed arrow between A and D). In the present experiment, we found two violations of the independence of P_1 with respect to U_1, four violations of the independence of U_1 with respect to P_1, four inconsistent tradeoffs and no violations of transitivity were found.

The tests for the violations of the independence of $P_1 \times U_1$ with respect to $P_2 \times U_2$ must also be applicable for the violations of the independence of $P_2 \times U_2$ with respect to $P_1 \times U_1$. In this case, these tests are performed for the four 2×2 submatrices as presented in Figure 2. These submatrices correspond with the columns of the parent matrix. The vertical and horizontal arrow in Figure 2 represent the direction of the consensus independence. The tradeoff is the comparison between A and D. In the experiment no transitivity violations (cf. dashed arrow between B and C) were found. We found three violations of the independence of U_2 with respect to P_2 and one violation of the independence of P_2 with respect to U_2. Seven subjects had an inconsistent tradeoff.

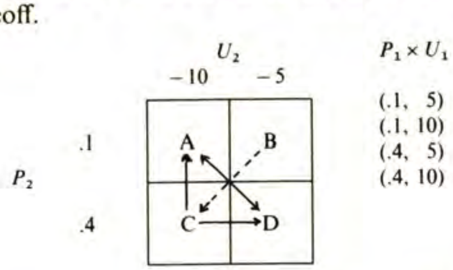

Fig. 2. — The 2×2 submatrices: columns of the basic matrix.

Test pattern I: direction of the tradeoff. The second use of test pattern I lies in the direction of the tradeoff. The tradeoff is a comparison of the effects of the probability of success (probability of failure) and the utility of success (utility of failure). For example, in Figure 1, the change from A to B is a change in U_1 from 5 to 10 and has increased attractiveness. The change from A to C is a change in P_1 form .1 to .4 and is also an increase in the preference judgment. The rank order between B and C indicates which of these increments had the greater effect. For example, if C were judged more attractive than B, this implies that the increment in P_1 had a greater effect in the attractiveness of A than the increment in U_1.

If cell C is consistently judged more attractivce than B, this signifies in terms of the model:

$$C \gtrsim B \langle = \rangle \frac{\text{Ø}_1(.4)}{\text{Ø}_1(.1)} \geq \frac{\text{Ø}_2(10)}{\text{Ø}_2(5)}. \tag{8}$$

The rank order between B and C is an indication whether the tradeoff of the effects of the expectancies is larger or smaller than the tradeoff of the utilities. Per subject we have four such comparisons between B and C; in total 96 tradeoff comparisons. In 96% of the cases C was preferred above B. This implies that the expectancies had a larger effect on the judgment of the exam situations than the utilities.

The same analysis was performed for the direction of the tradeoff between P_2 and U_2 (the diagonal arrow in Figure 2). The inequality is as follows:

$$A \gtrsim D \langle = \rangle \frac{\text{Ø}_3(.1)}{\text{Ø}_3(.4)} \geq \frac{\text{Ø}_4(-5)}{\text{Ø}_4(-10)}. \tag{9}$$

In 91.6% of the tradeoff comparisons A was preferred above D. So, also the P_2 variable had — in comparison with the U_2 variable — a larger effect on the judgment of the exam situations.

Test pattern II: testing the unassociated variables. Test pattern II contains the qualitative tests for the *unassociated* variables. There are four cross partitions of unassociated variables: $P_1 \times P_2$, $P_1 \times U_2$, $U_1 \times P_2$, $U_1 \times U_2$. We shall illustrate the test by means of the partition $U_1 \times U_2$. The extension to the other partitions is straightforward. The two levels of the column factor U_1 and the two levels of the row factor U_2 generate the 2×2 matrix as illustrated in Figure 3. The direction of the consensus independence is illustrated with a vertical and horizontal arrow. The same 2×2 matrix can be found four times in the parent matrix. Thus, the 4×4 matrix contains four 2×2 submatrices, each submatrix corresponding with a different fixed level of $P_1 \times P_2$ (cf. Figure 3).

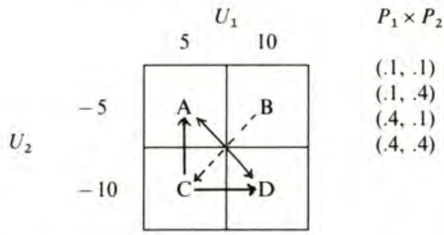

Fig. 3. — The 2×2 submatrices of the cross partition $U_1 \times U_2$.

As in test pattern I, one can identify the violations of the consensus independence in the 2×2 submatrices of each partition. For the $U_1 \times U_2$ partition three violations of the CI of U_1, with respect to U_2 and three violations of the CI of U_2 with respect to U_1 were identified. For the $P_1 \times P_2$ partition two violations of the CI of P_1 with respect to P_2 and one violation of the CI of P_2 with respect to P_1 were found. For the $P_1 \times U_2$ partition there were two violations of the CI of P_1 with respect to U_2 and three violations of the CI of U_2 with respect to P_1. Last but not least, for the $U_1 \times P_2$ partition two violations of the CI of P_2 with respect to U_1 and four violations of the CI of U_1 with respect to P_2. For all partitions no violations of consensus transitivity were found.

In test pattern II emphasis was put on the analysis of the tradeoffs between the unassociated variables. If the direction of the tradeoff between U_1 and U_2 is consistently $D \geqslant A$, this implies in terms of the model (for a fixed level (.1,.1) of $P_1 \times P_2$):

$$D \gtrsim A \langle = \rangle \frac{\text{Ø}_1(.1)}{\text{Ø}_3(.1)} \geq \frac{\text{Ø}_4(-5) - \text{Ø}_4(-10)}{\text{Ø}_2(10) - \text{Ø}_2(5)}. \tag{10}$$

For a fixed level (.4,.1) of $P_1 \times P_2$:

$$D \gtrsim A \langle = \rangle \frac{\text{Ø}_1(.4)}{\text{Ø}_3(.1)} \geq \frac{\text{Ø}_4(-5) - \text{Ø}_4(-10)}{\text{Ø}_2(10) - \text{Ø}_2(5)}. \tag{11}$$

For a fixed level (.1,.4) of $P_1 \times P_2$:

$$D \gtrsim A \langle = \rangle \frac{\text{Ø}_1(.1)}{\text{Ø}_3(.4)} \geq \frac{\text{Ø}_4(-5) - \text{Ø}_4(-10)}{\text{Ø}_2(10) - \text{Ø}_2(5)}. \tag{12}$$

For a fixed level (.4,.4) of $P_1 \times P_2$:

$$D \gtrsim A \langle = \rangle \frac{\text{Ø}_1(.4)}{\text{Ø}_3(.4)} \geq \frac{\text{Ø}_4(-5) - \text{Ø}_4(-10)}{\text{Ø}_2(10) - \text{Ø}_2(5)}. \tag{13}$$

The left hand side of each inequality is a value of the *decision variable*, the right hand side is called the *threshold criterion* (Coombs & Lehner, 1981b, p. 578). The threshold criterion is a constant over the four inequalities. By virtue of the fact that $\varnothing_1$, and $\varnothing_3$ are positive and increasing in the arguments, there is a necessary partial order of the values of the decision variable as illustrated in Figure 4a. The bilinear model requires that the following set of *decision implications* is associated with this partial order (cf. Figure 4b): if a subject says $D \geqslant A$ at one point in the partial order of Figure 4a he must say the same for any point above that in the partial order and if he says $A \geqslant D$ one point in the partial order he must say the same below that point in the partial order.

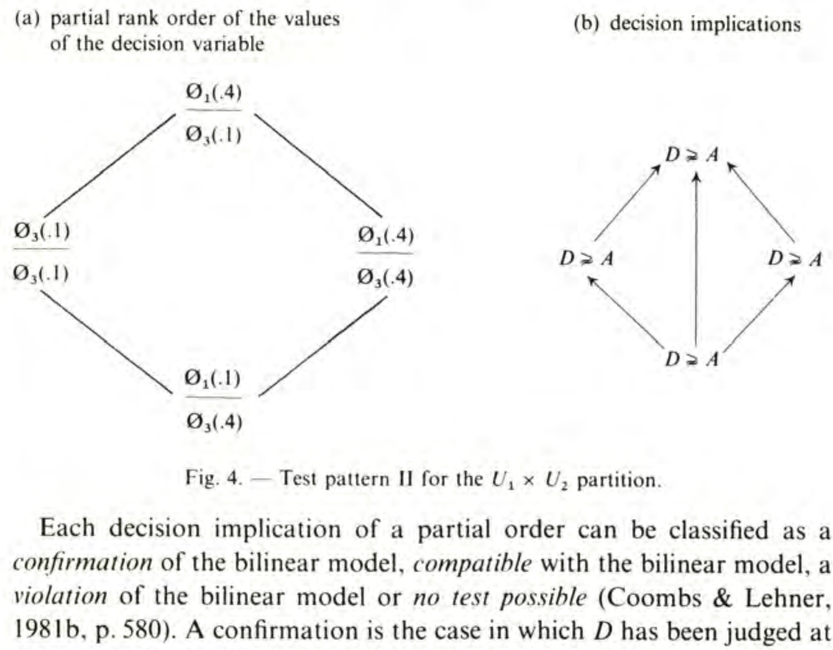

Fig. 4. — Test pattern II for the $U_1 \times U_2$ partition.

Each decision implication of a partial order can be classified as a *confirmation* of the bilinear model, *compatible* with the bilinear model, a *violation* of the bilinear model or *no test possible* (Coombs & Lehner, 1981b, p. 580). A confirmation is the case in which D has been judged at least as attractive as A at one level in the partial order and the same event occurs for a ratio (value of the decision variable) larger than that and equivalenty the case in which A has been judged more attractive than D at one level in the partial order and the same event occurs for a ratio below that in the partial order. A violation occurs when D has been judged at least as attractive as A at one point in the partial order and A has been judged more attractive than D for a higher ratio. The case in which A has been judged more attractive than D at some point in the partial order and D has been judged at least as attractive as A for a higher level of the decision variable is compatible with the bilinear model, but is in fact a weak confirmation. For each individual rank order that does not satisfy the consensus independence axiom no test is possible.

Per partial order five decision implications are associated with one subject. So for the 24 subjects there are 120 implications. Each implication was classified as a confirmation, compatible, violation or no test possible. Table 3 represents the results of this classification: most implications confirmed or were at least compatible with the bilinear model.

Tab. 3. — Results of test pattern II for all partitions

partition	Confirmation	Compatible	Violation	No test
$U_1 \times U_2$	55%	30%	3%	12%
$P_1 \times P_2$	65%	25%	2.5%	7.5%
$P_1 \times U_2$	73%	12.5%	2%	12.5%
$U_1 \times P_2$	80%	5%	4%	11%

Conclusion

As a test of the bilinear model we performed several qualitative tests. Most individual rankorders satisfied the independence axiom and the consensus independence property. The decision implications of the partial orders of the unassociated variables confirmed or were compatible with the bilinear model. In view of these results, one can say that the bilinear expectancy model is a valid description of the data of experiment 2. Furthermore, test pattern I revealed that the expectancy variable had a larger effect on the judgment of the exam situations than the utility variable.

One has to be careful with this conclusion. Firstly, the analysis was restricted to an ordinal analysis, namely we could not use statistical techniques to identify whether the violations of the model are substantive or negligible. Coombs and Lehner (1981b, p. 586) remark that the statistical problem is one of the most severe of the conjoint analysis of the bilinear model. Second, the analysis was performed on the rank order of one stimulus set. A more general analysis would be when several stimulus sets are used. In that case not only a qualitative test within but also between the several matrices should be possible.

Discussion

Taking the results of the two experiments together, one can say that the functional form of the SEU model has been confirmed. I.e., the subjective probability and utility of a given outcome of action combine multiplicatively, while the resultant products combine additively to define the SEU of a goal directed action. We discussed and illustrated the usefulness of conjoint measurement analysis as a valid approach for the test and measurement problem of the expectancy theorists. In experiment 1, in particular, the conjoint measurement approach is used to obtain valid measurements of the subject variable motive.

Note, that only the descriptive validity of this model has been proven, namely it is as if subjects combine the subjective probabilities and utilities multiplicatively and the resultant products additively. The important empirical question is what subjects really do when it seems that they multiply expectancies and utilities. Do they really multiply numbers or do they use some analog form of the multiplication? The same question counts for the additive composition rule.

Lopes and Ekberg (Lopes, 1976; Lopes & Ekberg, 1980) suggest that serial fractionation is the cognitive process that can be described by multiplication. Serial fractionation is a variant of the anchoring and adjustment heuristics (Tversky & Kahneman, 1974), in the sense that the adjustment always occurs downward in proportion to the information being integrated. Thus serial fractionation is an explanation of how subjects produce multiplicative data without really multiplying numbers. For the subjective valuation of a gamble, for example, subjects should use the amount to be won (U) as anchor and than adjusting downward in proportion to the probability of success (P) (Lopes & Ekberg, 1976). More research on this question is certainly needed.

Another point of discussion is the fact that expectancy models are examples of *static* models. An expectancy theory implicitly assumes that behavior can be conceptualized as a sequence of goal directed episodes. Each episode is described in terms of direction and intensity of the goal directed activity. A more realistic conception would be the definition of behavior as a continuum of actions characterized by the change of one action into another. So, expectancy theorists should strive for the development of *dynamic* models of behavior. A first attempt is already made by Atkinson and Birch (1970, 1974) and by Coombs and Smith (1973).

REFERENCES

Ajzen, I., & Fishbein, M. (1969). The prediction of behavioral intentions in a choice situation. *Journal of Experimental Social Psychology, 5*, 400-416.

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Hillsdale, NJ: Prentice-Hall.

Anderson, N. H. (1962). Applications of an additive model to impression formation. *Science, 138*, 817-818.

Anderson, N. H. (1974). Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 236-305). New York: Academic Press.

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Arnold, H. J. (1976). *Cognitive processing in expectancy theories of motivation*. Unpublished doctoral dissertation, Yale University.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359-372.

Atkinson, J. W., & Birch, D. (1970). *The dynamics of action*. New York: Wiley.

Atkinson, J. W., & Birch, D. (1974). The dynamics of achievement motivation. In J. W. Atkinson & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 271-325). Washington: Winston.

Bentler, P. M., & Speckart, G. (1979). Models of attitude-behavior relations. *Psychological Review, 86*, 452-464.

Bentler, P. M., & Speckart, G. (1981). Attitudes "cause" behaviors: A structural equation analysis. *Journal of Personality and Social Psychology, 40*, 226-238.

Cochran, W.G., & Cox, G.M. (1957). *Experimental designs* (2nd ed.). New York: Wiley.

Coombs, C. H., & Smith, J. E. K. (1973). On the detection of structure in attitudes and developmental processes. *Psychological Review, 80*, 337-351.

Coombs, C. H., & Lehner, P. E. (1981a). Evaluation of two alternative models for a theory of risk: I. Are moments of distributions useful in assessing risk? *Journal of Experimental Psychology: Human Perception and Performance, 7*, 1120-1123.

Coombs, C. H., & Lehner, P. E. (1981b). The conjoint analysis of the bilinear model, illustrated with a theory of risk. In I. Borg (Ed.), *Multidimensional data representations, when and why* (pp. 563-589). Ann Arbor: Mathesis Press.

De Soete, G. (1980). The integration of motivational information: A conjoint measurement analysis. *Acta Psychologica, 46*, 257-269.

De Soete, G. (1984). Conjunct meten: situering en recente ontwikkelingen. *Nederlands Tijdschrift voor de Psychologie, 39*, 88-99.

Fredericks, A. J., & Dossett, D. L. (1983). Attitude-behavior relations: A comparison of the Fishbein-Ajzen and the Bentler-Speckart models. *Journal of Personality and Social Psychology, 45*, 501-512.

Heckhausen, H. (1977). Achievement motivation and its constructs: A cognitive model. *Motivation and Emotion, 1*, 283-329.

Hull, C. L. (1943). *Principles of behavior*. New York: Appleton.

Klitzner, M. D. (1977). *Small animal fear. An integration-theoretical analysis.* Unpublished doctoral dissertation, University of California, San Diego.

Klitzner, M. D., & Anderson, N. H. (1977). Motivation × Expectancy × Value: A functional measurement approach. *Motivation and Emotion, 1*, 347-365.

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. (1971a). Conjoint measurement analysis of composition rules in psychology. *Psychological Review, 78*, 151-169.

Krantz, D. H., & Tversky, A. (1971b). Letters to the editor: An exchange on functional and conjoint measurement. *Psychological Review, 78*, 457-458.

Lopes, L. L. (1976). Model-based decision and inference in stud poker. *Journal of Experimental Psychology: General, 105*, 217-239.

Lopes, L. L., & Ekberg, P. S. (1980). Test of an ordering hypothesis in risky decision making. *Acta Psychologica, 47*, 161-167.

Luce, R. D., & Tuckey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental judgment. *Journal of Mathematical Psychology, 1*, 1-17.

Lynch, J. G. (1979). Why additive utility models fail as descriptors of choice behavior. *Journal of Experimental Social Psychology, 15*, 397-417.

Lynch, J. G., & Cohen, J. L. (1978). The use of subjective utility theory as an aid to understanding variables that influence helping behavior. *Journal of Personality and Social Psychology, 36*, 1138-1151.

Raynor, J. O. (1974). Future orientation in the study of achievement motivation. In J. W. Atkinson & J. O. Raynor (Eds.), *Motivation and achievement* (pp. 121-154). Washington: Winston & Sons.

Roskam, E. E. Ch. I. (1974). *Unidimensional conjoint measurement for multifaceted designs.* Internal Report 74-MA-09. University of Nijmegen.

Roskam, E. E. Ch. I. (1982). *Conjoint measurement.* Lecture notes, European Institute for Advanced Study in Management, Brussels.

Shanteau, J. C. (1974). Component processes in risky decision making. *Journal of Experimental Psychology, 103*, 680-691.

Shanteau, J. C., & Nagy, G. (1976). Decisions made about other people: A human judgment analysis of dating choice. In J. Caroll & J. Payne (Eds.), *Cognition and social judgment* (pp. 123-141). New York: Erlbaum.

Sheluga, D., Jaccard, J., & Jacoby, J. (1979). Process-descriptive methods and the estimation of preference: An integrative approach. *Journal of Consumer Research, 6*, 166-176.

Sjöberg, L. (1982). Beliefs and values as attitude components. In L. Sjöberg (Ed.), *Social attitudes and psychological measurement* (pp. 199-216). Hillsdale NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science, 185*, 1124-1131.

Vroom, V. H. (1964). *Work and motivation.* New York: Wiley.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548-573.

Dienst voor Psychologie
H. Dunantlaan 2
9000 Gent

Received December 1986