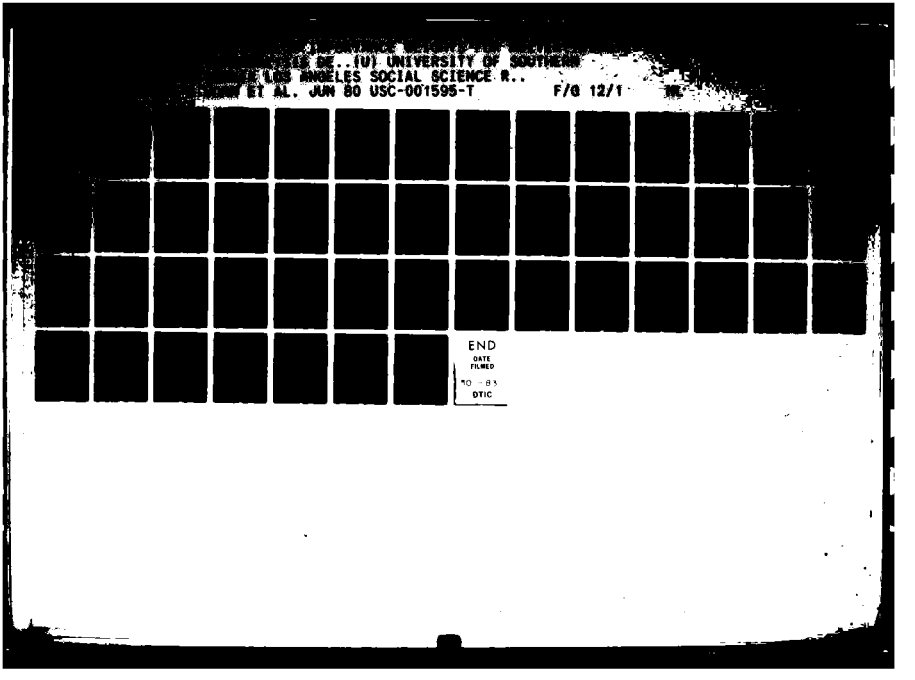
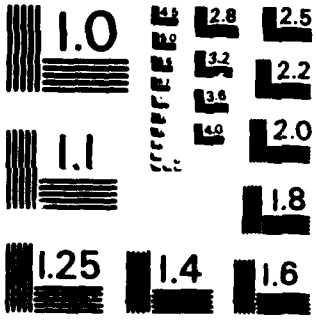

(U) UNIVERSITY OF SOUTHERN
LOS ANGELES SOCIAL SCIENCE R..
ET AL. JUN 80 USC-001595-T
F/G 12/1
END
DATE FILMED
10 - 83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

6/30--001595-T

social science research institute

UNIVERSITY OF SOUTHERN CALIFORNIA

Research Report 80-4

A COMPARISON OF IMPORTANCE WEIGHTS FOR MULTIATTRIBUTE UTILITY ANALYSIS DERIVED FROM HOLISTIC, INDIFFERENCE, DIRECT SUBJECTIVE AND RANK ORDER JUDGMENTS

Richard S. John, Ward Edwards
and Linda Collins

Social Science Research Institute
University of Southern California

Sponsored by:

Defense Advanced Research Projects Agency
Department of the Navy
Office of Naval Research

Monitored by:

Department of the Navy
Office of Naval Research
Pasadena Branch Office
Contract No. N00014-79-C-0038

Social Science Research Institute
University of Southern California
Los Angeles, California 90007
(213) 743-6955

INSTITUTE GOALS:

The goals of the Social Science Research Institute are threefold:

- To provide an environment in which scientists may pursue their own interests in some blend of basic and methodological research in the investigation of major social problems.
- To provide an environment in which graduate students may receive training in research theory, design and methodology through active participation with senior researchers in ongoing research projects.
- To disseminate information to relevant public and social agencies in order to provide decision makers with the tools and ideas necessary to the formulation of public social policy.

HISTORY:

The Social Science Research Institute, University of Southern California, was established in 1972, with a staff of six. In fiscal year 1978-79, it had a staff of over 90 full- and part-time researchers and support personnel. SSRI draws upon most University academic Departments and Schools to make up its research staff, e.g. Industrial and Systems Engineering, the Law School, Psychology, Public Administration, Safety and Systems Management, and others. Senior researchers have joint appointments and most actively combine research with teaching.

FUNDING:

SSRI Reports directly to the Executive Vice President of USC. It is provided with modest annual basic support for administration, operations, and program development. The major sources of funding support are federal, state, and local funding agencies and private foundations and organizations. The list of sponsors has recently expanded to include governments outside the United States. Total funding has increased from approximately $150,000 in 1972 to almost $3,000,000 in the fiscal year 1978-1979.

RESEARCH INTERESTS:

Each senior SSRI scientist is encouraged to pursue his or her own research interests, subject to availability of funding. These interests are diverse; a recent count identified 27. Four major interests persist among groups of SSRI researchers: crime control and criminal justice, methods of dispute resolution and alternatives to the courts, use of administration records for demographic and other research purposes, and exploitation of applications of decision analysis to public decision making and program evaluation. But many SSRI projects do not fall into these categories. Most projects combine the skills of several scientists, often from different disciplines. As SSRI research personnel change, its interests will change also.

A COMPARISON OF IMPORTANCE WEIGHTS FOR MULTIATTRIBUTE UTILITY ANALYSIS DERIVED FROM HOLISTIC, INDIFFERENCE, DIRECT SUBJECTIVE AND RANK ORDER JUDGMENTS

Richard S. John

Ward Edwards

Linda Collins

Social Science Research Institute

University of Southern California

Research Report 80-4

Sponsored by:

Defense Advanced Research Projects Agency

SUMMARY

Research done in the 1960's and early 1970's suggested that although statistical weights and subjective weights show some correspondence in regression-like situations, subjective weights tend to be too flat by comparison; statistical weights usually show that some attributes are quite important, while others are hardly important at all. More recent discussions of this literature, however, have pointed out a number of methodological problems with much of the early research, and have reached a more optimistic conclusion with respect to subjective weights. Several experiments support the more recent interpretation.

The present study compared weight estimation procedures for additive, riskless four-attribute value functions with linear single-attribute values. Self-explicated (subjective) weights were assessed from direct subjective and rank order estimates of attribute importance; observer-derived weights were determined both from indifference judgments (axiomatic approach) and from holistic evaluations (statistical approach) of alternatives. Assessed weights were compared to a "true" weight vector used to generate feedback during pre-assessment learning trials (constructed with zero inter-attribute correlations). Although self-explicated weights tended to be flatter than observer-derived weights, resulting composites correlated equally well with "true" composites. Only slight differences were found in ordinal correspondence between "true" and assessed weights.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

This research was supported by the Advanced Research Projects Agency of the Department of Defense, and was monitored by the Office of Naval Research under Contract No. N00014-C-0038. The authors would like to thank Detlof von Winterfeldt and Bill Stillwell, who made many valuable contributions to the research summarized in this report.

DISCLAIMER

The views and conclusion contained in this document are those or the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or of the United States Government.

INTRODUCTION

Judgments about the relative desirability of acts or objects are inherently subjective. They depend on subjective likelihoods of the consequences of choosing an act or object, on subjective values for these consequences, and on subjective trade-offs among different consequences. Multi-attribute utility analysis (MAUA) models such subjective value judgments by eliciting value relevant attributes of the objects or acts, by assessing single-attribute utilities and weights, and by aggregating these inputs into an overall value index. Proponents of MAUA argue that the choices dictated by MAUA will, on the average, yield more favorable consequences than choices based on other types of evaluations, e.g., intuition. However, since both inputs and outputs of MAUA are subjective numbers, and since the consequences of any choice are subjectively experienced, researchers have faced substantial difficulties in validating that claim.

In this paper we will explore a validation paradigm based on the thesis that in many cases value is simply a surrogate for probability. This paradigm allows us to validate the MAUM claim, and to test competing MAUA procedures, by applying evaluation methods in situations in which probabilistic relationships between choices and their consequences can be ascertained. One need only compare the resultant evaluations of choices (derived from various MAUA procedures and intuition) to the (known) distribution of consequences associated with each alternative. In the following we will discuss the conceptual basis and an operationalization of this paradigm in more detail. Subsequently, we will describe an experiment which validated four MAUA weighting procedures within this paradigm.

In many evaluation problems, the relationship between value and probability is obvious: A "good" applicant for graduate school is likely to

succeed in the graduate program; a "good" credit applicant is unlikely to default; a "good" scientific manuscript is likely to be accepted for publication in a prestigious journal. However, in every one of the examples above, the defining characteristics of the alternatives are probabilistically related to future consequences that are determined once the choice is made. In most cases, degree of deservedness (worth) is dependent upon the alternative's likelihood of resulting in each possible consequence (outcome state) and the desirability of each consequence.

In a credit granting decision, for example, the outcome states might be discrete (such as default vs. no default) or continuous (such as the dollar amount of profit made on the loan). In the discrete (dichotomous) case, worth is often considered monotonic to the likelihood of the "good" outcome, e.g., no default, while in the continuous case, worth is normally thought to vary monotonically along a bipolar continuum from "bad" (e.g., substantial dollar loss) to "good" (e.g., large dollar profit). Thus, an alternative possesses no worth or "deservedness" in and of itself; rather, worth is induced upon the alternative as a function of the probabilistic relationship between alternative characteristics and future consequences. This theoretical position is widely held in modern psychology: beliefs (probabilistic relationships) determine affects (worth evaluations), which in turn determine behavior (choices). In other words, what we think influences what we feel, and what we feel influences what we do.

Most day to day choices are made from evaluations based on a casual learning of the relevant probabilistic relationship between alternatives and consequences. Indeed, there may be little or no thought given to the beliefs and affects that influence choice. Important decisions, such

as those listed above usually require accurate evaluations, which in turn are best obtained by a precise knowledge of the probabilistic relationship between alternative characteristics and outcome states. In such cases, prior decisions and their resulting outcomes may be scutinized. If the decision is important enough, and if a sufficient number of past decisions and consequences have been documented and stored, professional learners (such as applied statisticians, management scientists, and industrial psychologists) may be employed to use complex retrospective techniques for uncovering useful probabilistic relationships between alternative characteristics and outcome states.

For many important decision problems (e.g., choosing a school desegregation plan) there is very little or no documented prior experience. Even when many past observations have been collected and stored, the probabilistic relationship may prove too complex for traditional post hoc analyses. Yet, although (normative) belief structures can not be explicated, affect will usually persist. That is, even in the absence of explicit relationships between alternative characteristics and consequences, various properties of the alternatives will be viewed as more or less desirable or worthy than other properties. Unlike the probabilistic relationships, which may be discovered by analyzing the environment, affect structures can only be explicated by studying the decision maker(s).

We used an operationalization of this validation strategy (c.f., Pearl, Note 1), tested by John and Edwards (Note 2) and similar to that utilized by Schmidt (1978), to compare weight estimation procedures for additive, riskless four-attribute value functions with linear single attribute values.

Estimated weights were compared to the "true" weights in the "artificial environment of choice-reward", i.e., in the linear model used to generate outcome feedback.

Of central interest is the performance of client explicated methods (such as rank weights, subjective [ratio] estimation, and constant sum) relative to so called observed derived methods (such as pricing-out, trading-off to the most important dimension, regression weights, and ANOVA weights derived from an orthogonal design). (For reviews of the client explicated vs. observer derived distinction, see Fischer, 1975, 1979; Huber, 1974a, 1974b; Johnson and Huber, 1977.) All client explicated approaches assume an additive model form and depend upon direct subjective estimates of all parameters, including weights. Subjective estimation techniques determine scale values of attributes on a dimension of "importance in determining the overall construct of evaluation". These scale values are called weights.

In contrast, observer derived approaches typically rely on (holistic) judgments that relate directly to the relative standing of some subset of choice alternatives on the construct of evaluation. Proposed aggregation rules are accepted only if the holistic judgments do not indicate violations of axioms or rejection of statistical hypotheses necessary for the model representation. Each holistic judgment can be thought of as representing one equation with some number of unknowns, depending upon the complexity of the accepted model form. In general, axiomatic procedures require a number of holistic judgments (equations) equal to the number of unknowns, and the parameter values (including weights) can be thought of as simply the solution to a set of simultaneous equations. Often, independent sets of holistic judgments (equations) are obtained, and the solution parameters

from each are compared. This is called sensitivity analysis. On the other hand, statistical procedures usually require a much larger number of holistic judgments (equations) than unknowns. Here, each judgment (equation) contains an error term, and parameter values are usually the critical point (minimum) of a loss function (such as least squares) defined over the errors. The sensitivity of statistical models is often gauged by the errors of estimate of the parameters.

Over twenty years after Paul Hoffman's (1960) seminal work on the correspondence of subjective (self-explicated) and statistical (observer derived) weights, there is little consensus as to whether weights should be "constructed" via direct assessments of importance. A very influential review by Slovic and Lichtenstein (1971) set the tone for much of the research for the past ten years, and their conclusions have been echoed by researchers across a diverse literature: management science (Zeleny, 1976, p. 14); attitude theory (Fishbein and Ajzen, 1972, p. 501; 1975, p. 159); verbal reporting on mental processes (Nisbett and Wilson, 1977, p. 254). Early results suggested that although statistical weights and subjective weights show some correspondence in regression-like situations, subjective weights tend to be too flat by comparison; statistical weights usually show that some attributes are quite important, while others are hardly important at all. More recent discussions of this literature, however, have pointed out a number of methodological problems with much of the early research, and have reached a more optimistic conclusion with respect to subjective weights (Schmitt and Levine, 1977; John and Edwards, Note 3). Several experiments support the more recent interpretation (Brehmer and Qvarnstrom, 1976; Schmitt, 1978; John and Edwards, Note 2).

Method

Overview and Independent Variables

Forty-six college studnets were taught a four attribute MAU model of diamond worth using the paradigm of multiple cue probability learning and outcome feedback; after training, subjects assessed MAU weight parameters via a variety of elicitation techniques. Although all subjects saw diamond profiles and outcome feedback with similar multivariate distributions (equal attribute variances and means, zero intercorrelations among attributes, and weight parameters in the ratio of 8:4:2:1), three task variables thought to affect learning were manipulated. Monetary payoff was manipulated by telling half of the subjects that they could earn up to $10.00 in cash, the exact amount depending upon their performance during the experiment. The other half were given no monetary incentive. Task uncertainty was set at one of two levels; half of the subjects received small random error in the diamond worth feedback (1% of total variance), while the other half received larger random error (18% of total variance). Exposure to the MAU model was manipulated by varying the total number of learning trials. Half of the subjects were trained for 120 trials, while the other half completed only 60 trials. Immediately after model training every subject made several independent assessments of attribute importance. These elicitation techniques will be described in detail subsequently.

Subjects

Forty-six students (26 males and 20 females) were selected from a much larger pool of volunteers from an Introductory Psychology course. The criteria for selection were scores of at least 600 (males) or 550 (females) on the mathematical aptitude section of the SAT, and a requirement that all subjects be whites whose native tongue was English -- the latter requirement because the experimental stimuli, hypothetical diamonds, relate to cultural mores.

Subjects were run either individually or in groups of up to six. Each session lasted 1 to 2 hours. Subjects received some payment (see below) and experimental credit in fulfillment of a course requirement.

Training Procedure

Each subject sat in front of a computer terminal with a CRT display. A written set of instructions said that subjects were to learn, via computer assisted instruction, a manner in which diamonds are appraised. Diamonds are evaluated on the basis of cut, color, clarity, and carat weight. The instructions explained these dimensions in considerable detail, and asserted (incorrectly) that any diamond can be described as a profile of four numbers, each between 0.0 and 10.0, representing the diamond's rating on the four attributes. Value increases with rating on each dimension.

During the experiment, the CRT would display a profile of four labelled numbers. The prompt "PRICE?" then appeared, and the subject entered a dollar estimate on the keyboard. Then the CRT displayed the "true" price of that diamond, the signed difference between "true" price and the subject's estimate,

and a standardized error score calculated as follows:

$$\text{Error Score} = \frac{E(MSE)_{equal} - (Error)^2}{E(MSE)_{equal} - E(MSE)_{beta}} \qquad (1)$$

$E(MSE)_{equal}$ is the expected mean squared error using equal weights, $(Error)^2$ is the squared deviation of the subject's estimate from the feedback, and $E(MSE)_{beta}$ is the expected mean squared error using the optimal beta weights. The instructions explained that a score of 1 is excellent, a score of 0 is very poor, and that scores above 1 or below 0 are possible but very infrequent. Subjects also recorded on paper any errors they detected in the feedback about the difference between estimated and true value; these are terminal errors. The few such instances were later corrected by editing.

Stimulus Generation

The ratings came from uniform distributions over the 0 to 10 range on each dimension. Consequently the expected value for each attribute was 5.0, and its standard deviation was 2.9. The expected intercorrelation between any pair of attributes was 0. The same set of 120 diamond profiles were presented in the same order to all subjects who saw 120 profiles; those who saw only 60 profiles saw the first 60 of those. Sample statistics by 30-trial blocks for all stimuli are acceptably close to their population values.

Outcome feedback was calculated from the following model:

$$\text{True Price} = 320(C_1) + 160(C_2) + 80(C_3) + 40(C_4) + k(N(0,1)) \qquad (2)$$

In Equation 2, C_i is the rating on the ith dimension, k is a constant that determines the precision of the model, and N(0,1) is standardized normal random error. The values $k = 100$ and $k = 500$ were used for different groups of subjects. The expected value of the true price is \$3000; its standard deviation is 1069 if $k = 100$ and 1176 if $k = 500$. Consequently, the

expected squared multiple correlation between the true price and the four attributes was .99 for k = 100 and .82 for k = 500. Four different assignments of attribute labels to weights were devised, and one was chosen randomly for each subject.

Post-Learning Weighting Judgments

Upon completion of the learning trials, the subject went individually to another room, and received a seven page self-administered booklet for weight assessments. The experiment asked the subject to read the instructions at the top of each page, and to ask any questions before starting work on that page. The order of assessment procedures was identical for all subjects. No subject could change previous responses after turning a page.

Bootstrapping. Raw regression weights were obtained by standard least squares regression analysis of each subject's responses over the last 30 learning trials.

Ranking. The subject simply rank ordered the four attributes from most to least important in determining price.

Most important dimension. The subject identified a most important dimension, and assigned a percentage that represented its importance in determining price. The instructions said that the ratio of the assigned percentage to 100 minus that percentage represented the ratio of the importance of the most important attribute to the total combined importance of the other three attributes.

Constant sum. The subject distributed 100 points across the four attributes, according to importance. The instructions said only that more important attributes should receive higher percentages.

Ratio estimation. First the subject once more ranked the attributes in order of decreasing importance. The least important dimension was assigned a weight of 10, and the subject provided weights for the other three dimensions using that weight as an anchor. The instructions said that the ratio of any given pair of weights should reflect the number of times more important one attribute is than the one with which it is being compared. (This is the response mode Edwards [1977] proposed in his SMART procedure.)

Pricing out. The subject was told to imagine that he or she possesses $3000 in cash and a diamond that scores (0, 0, 0, 0) -- worst possible scores on all four dimensions. For each dimension, the subject states how much he or she would be willing to pay in order to exchange that diamond for one that scores 10 on that dimension and 0 on the other three. (For details, see Keeney & Raiffa, 1976, p.125)

Trading off to the most important dimension. The subject once more identifies the most important dimension. For convenience of exposition, suppose that is the first one listed. Then the subject must specify a value of x such that diamonds (x, 0, 0, 0) and (0, 10, 0, 0) are equivalent in price. This judgment must be made four times, once for each dimension. Of course, when the most important attribute is set to 10, the two diamonds will be identical; this judgment was used to make sure the subject understood the instructions. (Again, for details, see Keeney & Raiffa, 1976, p. 121.)

Holistic Orthogonal Parameter Estimation (HOPE). HOPE simply required the subject to appraise 17 diamonds holistically. The set of 17 diamonds is carefully chosen so that parameters can be recovered from the judgments. (The HOPE procedure, developed by Barron and Person [1979], is closely akin to standard fractional replication ANOVA designs.)

Results

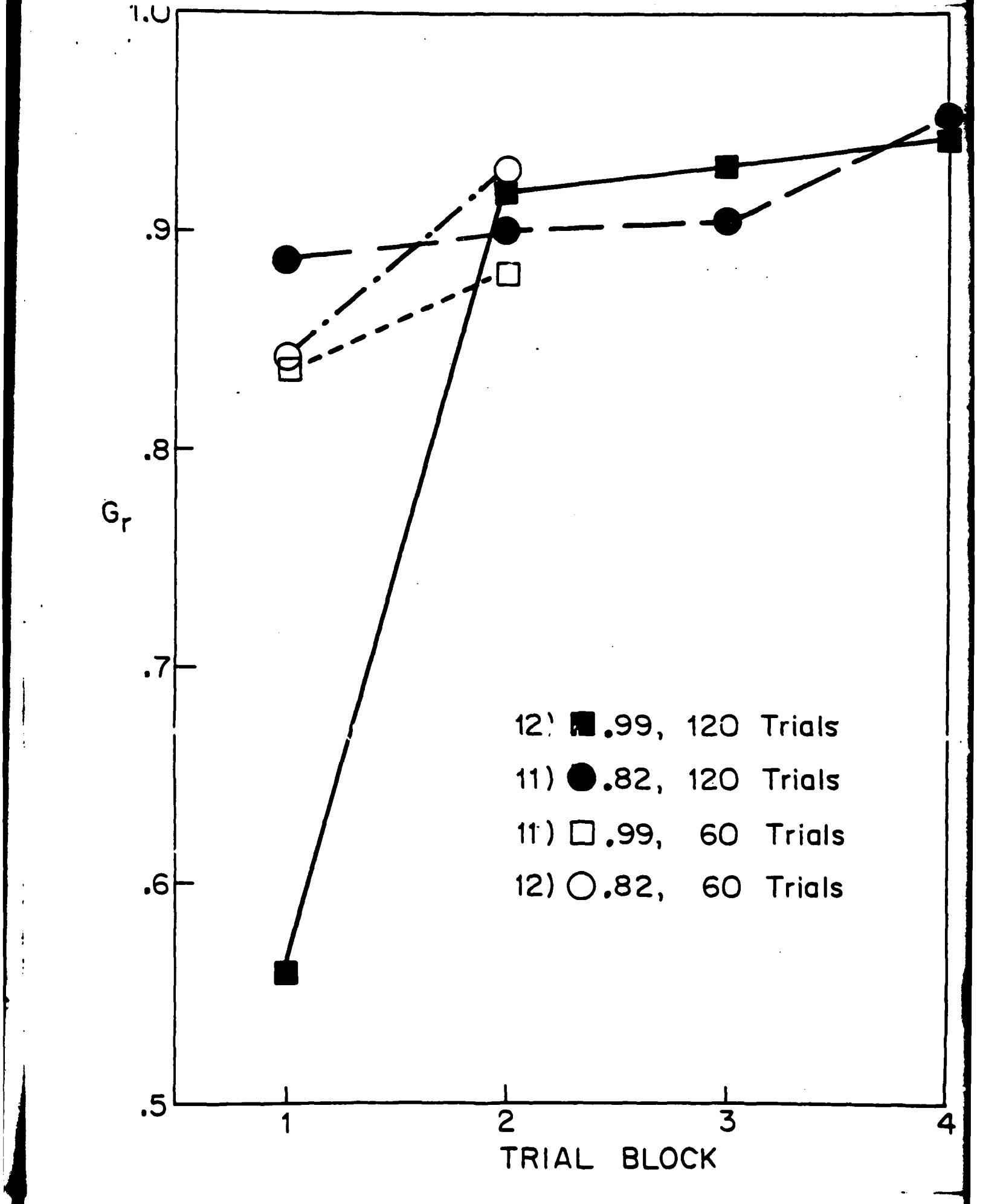
MAU Model Learning

The lens model index of matching (G) is the correlation between composites derived from consistent application of the weights used to generate outcome feedback and the weights derived statistically from subjects' holistic diamond appraisals. Thus, G (often called "knowledge", appropriately enough), is a measure of the extent to which the subject's combination rule (weight vector) corresponds to that of the "true" model in creating composites. For the specific MAU model we taught our subjects, the correlation between composites from different sets of weights is directly related to the parameters of the bivariate distribution describing the weights. When all attribute correlations are zero and the attribute variances are equal, the correlation between composites from subject's weights and from true weights is given by the following formula (Gulliksen, 1950, p. 319):[1]

$$R_{X_s X_t} = \frac{r_{st}\,(\sigma_s/\bar{s})\,(\sigma_t/\bar{t}) + 1}{\sqrt{1 + (\sigma_s/\bar{s})^2}\sqrt{1 + (\sigma_s/\bar{t})^2}} \qquad (2)$$

(where s and t are the subject's and "true" weighting schemes, respectively, and X_s and X_t are the composite evaluations resulting from them). Equation 2 was used to calcualte matching scores for every subject for each block of 30 learning trials.[2]

Figure 1 shows average G scores as a function of number of trials and k. (Payoff or its absence make no difference in the data.) Figure 1 shows a significant increase in matching from the first trial block to the second ($F(1,38)=15.39$, $p < .05$). This result also holds for the 60 trial subjects


1.0
.9
.8
.7
.6
.5
G_r
1
2
3
4
TRIAL BLOCK
12) .99, 120 Trials
11) .82, 120 Trials
11) .99, 60 Trials
12) .82, 60 Trials

considered separately ($F(1,19)=4.82$, $p < .05$), and across all four trial blocks for the 120 trial subjects ($F(3,57)=10.79$, $p < .05$). There is no significant increase in performance across the last trial blocks for the 120 trial subjects ($F(2,38=2.29$, $p > .05$). Subject's learning about the weights is v rtually complete by about trial #30. For both subjects who received payoffs and thosewho did not, the combination of little task uncertainty and an expectancy of many learning trials produced very poor performance in the first 30 trials.

Weight Assessments

Subjects' knowledge of the weights after the first trial block is not mediated by monetary payoffs, task uncertainty, or the number of learning trials completed (see Figure 1); thus, we have collapsed weight assessments across all three task manipulations. Whether or not the subject assigned the largest weight to the most important attribute and whether or not he/she assigned weights in the correct rank ordering are good indications of weight correspondence. The number of subjects who correctly indicated the most important dimension (ties not counted) and the number who indicated the correct rank ordering (including at most 1 tie) are shown in Table 1 for each of the seven assessment techniques.

Subjects most often correctly identified the most important dimension using the ratio technique and most often indicated the correct rank ordering using the bootstrapping method. However, there were no significant differences on either of these measures ($X^2(6) = 8.18$, $p > .05$ and $X^2(6) = 2.28$, $p > .05$, respectively). A more sensitive measure of correspondence is the number of inverted attribute pair orders (a linear transformation of

TABLE 1

Weight Orders

Assessment Technique	# of Ss Correctly Identifying Most Important Dimension	# of Ss with <1 Inversions with True Weights	Mean # of Inversions with True Weights
Bootstrapping	35	16	1.06
Ranking	35	7	1.37
Constant Sum	32	11	1.44
Ratio	36	9	1.42
Pricing-Out	34	12	1.41
Trading-Off	33	8	1.73
HOPE	30	7	1.68

Kendall's T). The mean number of such inversions for each technique is also shown in Table 1. The fewest inversions resulted from the bootstrapping weights while trading-off to the most important dimension and HOPE produced the most inversions. The mean number of inversions was significantly different across assessment procedures ($F(6,228)=3.09$, $p < .05$). Well over 90% of the subjects yielded 3 or fewer inversions for all of the obtained attribute orderings. Furthermore, all of the cumulative distributions of inversions are significantly different from that expected if subjects were simply providing random orderings (by the Kolmogorov goodness of fit test, $p < .05$).

In addition to assigning weights in the correct rank ordering, we would like subjects to spread the weights appropriately. One good indication of the weight spread is the ratio of the weight assigned to the most important dimension to the sum of the weights assigned to the remaining three dimensions. Since a log transformation of this ratio is essentially linear with the normalized weight assigned to the most important dimension, we have elected simply to use the normalized weights For four dimensions, specification of the weight on the most important dimension severely restricts the variance the range of the weight vector. Of course, what constitutes an appropriate weight on the most important dimension depends upon whether the subject correctly identified the most important dimension or not. If he/she did, then the optimal weight is 53.3; if some other attribute receives a higher weight than the "true" most important dimension, flatter weights are better than more extreme ones, i.e., the closer to 25, the better. Table 2 displays mean maximum weights for each assessment technique conditional upon those subjects who: correctly identified the most important dimension (see Table 1, column a, for sample sizes); correctly identified the most important

TABLE 2

Mean Weights on the Most Important Dimension (MID)

Assessment Technique	S Correctly Identified MID: Correct for EACH Technique (a)	S Correctly Identified MID: Correct for ALL Techniques N=16 (b)	S Correctly Identified MID: % with Weight <53 (c)	S Incorrectly Identified MID: Weight on Ss MID (d)
Bootstrapping	52.3	55.2	57%	32.6
Direct Assess of MID	43.6	43.9	88%	39.0
Constant Sum	41.9	43.9	91%	36.3
Ratio	41.6	42.1	100%	41.8
Pricing-Out	42.2	42.2	94%	44.5
Trading-Off	46.3	51.7	79%	38.8
HOPE	48.3	51.6	80%	39.9

dimension for all seven assessment techniques (N = 16,column b); incorrectly identified the most important dimension (sample sizes are 46 minus the sample sizes for (a), column d). In addition, the percentage of those subjects correctly identifying the most important attribute (column a) who gave weights less than the optimal value (53.3) is shown in column (c).

In general, all of the weighting techniques, with the exception of bootstrapping, underestimated weights to the correctly identified most important dimension. HOPE and trading-off to the most important dimension tended to provide more extreme weights on the correctly identified most important dimension then did the remaining four assessment techniques. A repeated measures analysis of variance, not including the 3 task manipulations, was run over the 16 subjects who correctly identified the most important dimension on all seven assessments. The means in column (b) were found to be significantly different from one another ($(F(6,90)=4.24, p<.05)$. A comparision of columns (a) and (b) suggests that mean weights on the most important dimension are larger for those subjects who correctly identified the most important dimension for all assessments than for those who did so for only a subset of them. Comparing column (a) with column (d) suggests the pleasant finding that subjects who did not know the most important dimension assigned flatter weights.

The results of this analysis suggest that bootstrapping weights are best in terms of producing both the correct rank ordering among the attributes and the correct weight magnitudes. HOPE and trading-off to the most important dimension are better than average in terms of magnitude or spread, but are the poorest at generating the correct rank ordering. Of course, these two effects will tend to cancel each other. All of the other techniques

produce highly similar orderings and spreads. Thus, it is not surprising that correlations between composites (calculated from Equation 2) from the subjects' weights and true model weights (assuming equal expected variances, all zero intercorrelations, and the expected OLS regression weights) show little differentiation. Average correlations range from .88 for trading-off and pricing-out to .92 for bootstrapping weights. These slight differences were not significant ($F(5,190)=1.88$, $p > .05$). Neither is it surprising that correlations between composites from the subject's bootstrapping weights and various other subjective weights demonstrate no differences ($F(4,152)=1.18$, $p > .05$). Mean correlations with bootstrapping range from .89 for pricing-out to .92 for HOPE. This overall level of performance is quite good, considering that equal weights produce a composite correlation of only .81 and extreme weights (using the most important dimension only) yield a composite correlation of .87 with the true weights.

Rank Weighting

Four sets of rank weights were generated from each subject's rank ordering of the attributes; two are designed so that the weight on the most important dimension matches that directly assessed by the subject. Rank-sum weights are a linear transformation of the ranks and rank-reciprocal weights are proportional to the reciprocals of the ranks. Decision-rule rank weights are determined by comparing the subject's directly assessed weight on the most important dimension to the weight on the most important dimension for rank-sum, rank-reciprocal, and equal weights. That rank weighting procedure producing the least discrepant weight on the most important dimension yields the decision-rule rank weights. Rank-exponent weights, proposed by Stillwell and Edwards (Note 5), are determined from Equation 3:

$$W_i = (K + 1 - R_i)^z \Big/ \sum_{j=1}^{K} R_j^z. \qquad (3)$$

(W_i is the normalized weight on the ith dimension, R_j is the subjects' ranking of the jth dimension, and k is the number of dimensions.) By substituting the elicited value of the weight on the most important dimension for W_1 in Equation 3, z is easily determined by iterative numerical methods to any degree of accuracy desired.

For the "true" MAU model we used, all four rank weighting schemes can potentially perform quite well. A subject who yields the correct rank ordering of the attributes (zero inversions) would obtain a correlation with the true weight composites of .97 for rank sum weights (40,30,20,10) and .99 for rank reciprocal weights (48,24,16,12). As we saw in Table 1, the direct ranking procedure was quite good in providing nearly correct rank orderings of attributes; thus, it is not surprising that the average rank sum and reciprocal correlations were .92 and .95, respectively. Had a subject not only provided the correct rank ordering, but also the correct directly assessed weight on the most important dimension (53.3), the decision rule rank weights would have been the same as the rank reciprocal weights; rank exponent weights under these conditions (53.3,30.0,13.3,3.3) yield a correlation very close to 1.0. Since the directly assessed weights to the most important dimension were underestimated (Table 2), it is also not surprising that decision rule rank weights and rank exponent weights performed no better than the rank weights not utilizing the directly assessed weight to the most important dimension. Correlations between composites weighted with true weights and with decision rule rank and rank exponent weights were .93 and .92, respectively.

We have shown that the directly assessed weight of the most important dimension is, like that for most of the other techniques, underestimated. However, one issue concerning rank-exponent and decision-rule rank weights is the degree to which directly assessing the weight on the most important dimension is even possible. One critical question concerns the degree to which the direct assessment will correspond to assessments involving all dimensions. An ordinal analsyis is presented in Table 3 showing the frequencies of subjects estimating the weight to the most important dimension (direct assessment) less than, greater than, and within 5% of the weight estimate provided by the other six elicitation procedures. Recall that the constant sum technique immediately followed the direct assessment of the weight of the most important dimension and that the response modes both required an estimate in terms of a percentage of 100. Somewhat surprisingly, 13 subjects changed their estimates, with 11 choosing to assign fewer points in the constant sum method. Thus, subjects reassessed already too flat weights as even flatter when asked to provide weights to the other three dimensions.

Discussion

All of the weight assessment techniques we studied yielded weights corresponding to the "true" weights to about the same degree. No significant differences in the correlation among composites were evidenced in our comparision of holistic procedures (bootstrapping and HOPE), indifference procedures (trading-off to the most important dimension and pricing-out), direct subjective estimates (method of constant sum and magnitude estimates with ratio instructions), and arithmetic transformations of rank orders (rank-sum, rank-reciprocal, rank-exponent, and decision-rule rank

TABLE 3

Direct Assessment of Weight on the Most Important Dimension (MID)

# of Subjects	Direct Assessment Greater	Direct Assessment Less	Equal ± 2.5	Changed MID	Tied MID
Assessment Technique:					
Bootstrapping	7	19	2	17	
Constant Sum	11	2	28	3	2
Ratio	10	13	21	1	1
Pricing Out	16	13	9	6	2
Trading Off	13	14	13	5	1
HOPE	4	20	5	16	1

techniques). All of these procedures substantially outperformed equal weighting and somewhat outperformed extreme weighting. Subjects exhibited knowledge of the "true" weighting scheme beyond simply knowing that all attributes are related to overall price (i.e., equal weighting) or that one attribute is highly related to overall price (i.e., extreme weighting). These results replicate those reported by John and Edwards (Note 2).

None of the more complicated weighting procedures performed any better than the simple technique of directly assessing the rank ordering and arithmetically transforming the ranks into weights. Although this might suggest that subjects' weight assessments contain no more useful information beyond that embodied in their rank ordering of the attributes, we must be cautious. The true weight ratios chosen for this experiment (8:4:2:1) along with the attribute structure (4 attributes, zero intercorrelations, and equal variances) provide an ideal setting for rank weights. That is, a correct rank ordering produces a minimum correlation among composites of .97 for the rank transformations suggested by Stillwell and Edwards (Note 5). In short, after ranks are known, there is little room for improvement. Of course, in the absence of analytical work, we have no way of assessing the generalizability of this example.

That rank weights outperformed equal weights is an important replication of a somewhat surprising finding by John and Edwards (Note 2). Although rank weighting procedures for MAUA have been extensively studied for at least fifteen years (e.g., Eckenrode, 1965; Permut, 1973), earlier results had suggested no differences between rank and equal weights (e.g., Beckwith & Lehmann, 1973; Eils & John, 1980; Einhorn & McCoach, 1977; Lehmann, 1971) or inferior performance by rank weights (e.g., Newman, 1977).

In addition to the main findings cited above, four other specific results are noteworthy. First, we found that model weights were learned in somewhat fewer than 60 trials, probably between 20 and 40. The rate at which subjects learned the model was altered by the combination of task uncertainty and number of learning trials. Specifically, subjects who expected to see many trials (120) and whose outcome feedback was relatively certain (1% variance unaccounted for by the diamond profile) learned weights at a slower pace than did other subjects. The final levels of weight knowledge observed were not mediated by any of the task variables. For the number of learning trials variable, a smaller value (less than 40 or so) would be needed to produce any potent manipulation for the four attribute task situation we studied. Monetary payoffs did not effect final levels of weight knowledge, probably for one or both of two reasons: (1) Most subjects did not seem to care about such a "small" amount of money($10.00); and (2) Many subjects commented that they found the "diamond appraisal" task quite interesting and stimulating. Both of these casual observations fit our stereotype of USC undergraduates. The lack of any main effect for task uncertainty is an important finding. Subjects were able to learn and accurately report weights in a task environment in which 18% of the variance was not accounted for by the five attributes. In real world settings, much of the variance in overall alternative value is often not accounted for by the specific sets of attributes chosen to represent the MAU structure. Furthermore, weights are often learned in highly uncertain real world environments in which all factors that ultimately determine an alternative's overall worth are not always know. Thus, our positive results in the 18% unaccounted for variance condition are suggestive that subjective weights can be obtained in complex, real world-like settings.

Our second specific result concerns the relative ability of the different weighting schemes to reproduce the correct rank ordering of attribute weights. The best orderings were clearly produced by bootstrapping weights. Trading-off to the most important dimension and HOPE yielded the greatest number of inversions. This is a puzzling result. HOPE and bootstrapping are virtually identical in terms of the subjects' task requirement (simple holistic evaluations), yet their performance was quite disparate. Also, trading-off to the most important dimension and pricing-out are very similar indifference procedures, yet trading-off yielded poorer orders. Curiously, bootstrapping was the first order obtained, and pricing-out and HOPE were the last for all subjects. Although we did not expect it to be the case, subjects may have become bored with the "somewhat repetitive" eliciation procedures, or they may simply have forgotten the weight ratios learned previously. (This explanation is most plausible for explaining the poorer rank orders from HOPE, the last elicitation. After making holistic evaluations and receiving outcome feedback from a computer, the paper and pencil method with no feedback may have seemed substantially less glamorous.) Although bootstrapping did enjoy the informational advantage of contiguous feedback, it is also true that bootstrapping weights are based on 30 holistic responses, the first 29 of which are made before the subject had completed all of the learning trials.

The third specific result has been reported in the literature on subjective weights many times: judged weights were too flat. Although all of our procedures produced weights flatter than the "true" weights, HOPE and bootstrapping weights were considerably more extreme than the others. Since we conditionalized on only those subjects who correctly identified the most

important dimension, we conclude that HOPE and bootstrapping yielded more nearly optimal weight spreads than did the other techniques. Thus, we seem to have replicated previous findings that subjective (non-holistic) assessment procedures produce too flat weights in comparison to holistic ones.

The final specific result concerns the four methods we tested for combining ordinal assessments of attribute importance with arithmetic transformations of the ranks to arrive at a weight vector. Recall that rank-sum and rank-reciprocal weights were based on the rank order assessment alone, whereas rank-exponent and decision rule rank weights combine rank order information with a direct assessment of the weight on the most important dimension. Our results showed no advantage to the methods that utilize the weight to the most important dimenison assessment. Since most direct assessments of importance on the most important dimension were about equal to the rank-sum weight on the most important dimension, both decision rule rank and rank-exponent weights were quite close to rank-sum weights for most subjects.

Footnotes

1. Gulliksen (1950, p. 316) assumed that the attributes were in z-score form (mean zero, variance one), and McClelland (Note 4) proved a similar theorem by assuming that the variances were all equal to one. Both of these assumptions are overly strong in terms of obtaining Equation 2. That the attribute variances are equal is a sufficient condition.

2. The Fisher z transformation for the Pearson correlation coefficient was not applied in the present report because all correlations were calculated using population parameters (equal attribute variances, zero intercorrelations, and "true" weights in the exact ratio of 8:4:2:1). Since our matching scores are theoretical population values, there is no reason to correct for biases in the sampling distribution of r. Had we actually applied the weights to a given sample of diamond profiles and calculated the correlation between the composites, then the z transformation would have been appropriate.

Reference Notes

1. Pearl, J. On demonstrating effectiveness of decision-analysis (UCLA Paper ENG-1272). Los Angeles: University of California, School of Engineering and Applied Science, September, 1972.

2. John, R.S. and Edwards, W. Subjective versus statistical importance weights: A criterion validation (SSRI Tech. Report 78-7). Los Angeles: University of Southern California, Social Science Research Institute, December, 1978.

3. John, R.S. and Edwards, W. Importance weight assessment for additive, riskless preference functions: A review (SSRI Tech. Report 78-5). Los Angeles: University of Southern California, Social Science Research Institute, December, 1978.

4. McClelland, G. Equal versus differential weighting for multiattribute decisions: There are no free lunches (Center Report #207). Boulder, Colo.: University of Colorado, Institute of Behavioral Science, August, 1978 (Revised Version).

5. Stillwell, W.G. and Edwards, W. Rank weighting in multiattribute utility and decision making: Avoiding the pitfalls of equal weights (SSRI Tech. Report 79-2). Los Angeles: University of Southern California, Social Science Research Institute, September, 1979.

References

Barron, H.F. and Person, H.B. Assessment of multiplicative utility functions via holistic judgments. Organizational Behavior and Human Performance, 1979, 24, 147-166.

Beckwith, N.E. and Lehmann, D.R. The importance of differential weights in multiple attribute models of consumer attitude. Journal of Marketing Research, 1973, 10, 141-145.

Brehmer, B. and Qvarnstrom, G. Information integration and subjective weights in multiple-cue judgments. Organizational Behavior and Human Performance, 1976, 17, 118-126.

Eckenrode, R.T. Weighting multiple criteria. Management Science, 1965, 12, 180-192.

Edwards, W. Use of multiattribute utility measurement for social decision making. In D.E. Bell, R.L. Keeney, and H. Raiffa (eds.), Conflicting objectives in decisions. New York: Wiley, 1977.

Eils, L.C. and John, R.S. A criterion validation of multiattribute utility analysis and of group communication strategy. Organizational behavior and Human Performance, 1980, 25, 268-288.

Einhorn, H.J. and McCoach, W. A simple multiattribute utility procedure for evaluation. Behavioral Science, 1977, 22, 270-284.

Fischer, G.W. Experimental applications of multiattribute utility models. In D. Wendt (ed.), Utility, probability, and human decision making. Dordrecht-Holland: D. Reidel Publishing Co., 1975, 47-85.

Fischer, G.W. Utility models for multiple objective decisions: Do they accurately represent human preferences? Decision Sciences, 1979, 10, 451-479.

Fishbein, M. and Ajzen, I. Attitudes and opinions. Annual Review of Psychology, 1972, 23, 487-544.

Fishbein, M. and Ajzen, I. Belief, attitude, intention, and behavior: an introduction to theory and research. Reading, Mass.: Addison-Wesley, 1975.

Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.

Hoffman, P.J. The paramorphic representation of clinical judgment. Psychological Bulletin, 1960, 47, 116-131.

Huber, G.P. Multiattribute utility models: A review of field and field-like research. Management Science, 1974, 20, 1393-1402(a).

Huber, G.P. Methods for quantifying subjective probabilities and multiattribute utilities. Decision Sciences, 1974, 5, 430-458(b).

Johnson, E.M. and Huber, G.P. The technology of utility assessmnet. IEEE Transactions on Systems, Man, and Cybernetics, 1977, SMC-7, 311-325.

Keeney, R.L. and Raiffa, H. Decisions with multiple objectives. New York: Wiley, 1976.

Lehmann, D.R. Television show preference: Application of a choice model. Journal of Marketing Research, 1971, 8, 47-55.

Newman, J.R. Differential weighting in multiattribute utility measurement: when it should not and when it does make a difference. Organizational Behavior and Human Performance, 1977, 20, 312-324.

Nisbett, R.E. and Wilson, T.D. Telling more than we can know: verbal reports on mental processes. Psychological Review, 1977, 34, 231-259.

Permut, S.E. Cue utilization patterns in student-faculty evaluation. Journal of Psychology, 1973, 83, 41-48.

Schmitt, N. Comparison of subjective and objective weighting strategies in changing task situations. Organizational Behavior and Human Performance, 1978, 21, 171-188.

Schmitt, N. and Levine, R.L. Statistical and subjective weights: ome problems and proposals. Organizational Behavior and Human Performance, 1977, 20, 15-30.

Slovic, P. and Lichtenstein, S. Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organizational Behavior and Human Performance, 1971, 6, 649-744.

Zeleny, M. The attribute-dynamic attitude model (ADAM). Management Science, 1976, 23, 12-26.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: USC-001595-T, SSRI-RR-80-4 ADA132 757	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): A comparison of importance weights for multi-attribute utility analysis derived from holistic, indifference, direct subjective and rank order judgments	5. TYPE OF REPORT & PERIOD COVERED: Technical, June, 1980
	6. PERFORMING ORG. REPORT NUMBER: 80-4
7. AUTHOR(s): Richard S. John, Ward Edwards, and Linda Collins	8. CONTRACT OR GRANT NUMBER(s): N00014-79-C-0038
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Social Science Research Institute, University of Southern California, Los Angeles, CA 90007	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research - Dept. of Navy, 800 Quincy Street, Arlington, VA 22217	12. REPORT DATE: June, 1980 13. NUMBER OF PAGES: 30
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research, Research Branch Office, 1030 East Green Street, Pasadena, California 91101	15. SECURITY CLASS. (of this report): unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
additive, riskless, holistic, true weight, value, probability

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Research done in the 1960's and early 1970's suggested that although statistical weights and subjective weights show some correspondence in regression-like situations, subjective weights tend to be too flat by comparison; statistical weights usually show that some attributes are quite important, while others are hardly important at all. More recent discussions of this literature, however, have pointed out a number of methodological problems with much of the early research, and have reached a more optimistic conclusion with respect to subjective weights. Several experiments support the more recent interpretation.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

The present study compared weight estimation procedures for additive, riskless, four-attribute value functions with linear single-attribute values. Self-explicated (subjective) weights were assessed from direct subjective and rank order estimates of attribute importance; observer-derived weights were determined both from indifference judgments (axiomatic approach) and from holistic evaluations (statistical approach) of alternatives. Assessed weights were compared to a "true" weight vector used to generate feedback during pre-assessment learning trials (constructed with zero inter-attribute correlations). Although self-explicated weights tended to be flatter than observer-derived weights, resulting composites correlated equally well with "true" composites. Only slight differences were found in ordinal correspondence between "true" and assessed weights.

CONTRACT DISTRIBUTION LIST
(Unclassified Technical Reports)

Director Advanced Research Projects Agency Attention: Program Management Office 1400 Wilson Boulevard Arlington, Virginia 22209	2 copies
Office of Naval Research Attention: Code 455 800 North Quincy Street Arlington, Virginia 22217	3 copies
Defense Documentation Center Attention: DDC-TC Cameron Station Alexandria, Virginia 22314	12 copies
DCASMA Baltimore Office Attention: Mr. K. Gerasim 300 East Joppa Road Towson, Maryland 21204	1 copy
Director Naval Research Laboratory Attention: Code 2627 Washington, D.C. 20375	6 copies

Revised August 1978

SUPPLEMENTAL DISTRIBUTION LIST
(Unclassified Technical Reports)

Department of Defense

Director of Net Assessment
Office of the Secretary of Defense
Attention: MAJ Robert G. Gough, USAF
The Pentagon, Room 3A930
Washington, DC 20301

Assistant Director (Net Technical Assessment)
Office of the Deputy Director of Defense
Research and Engineering (Test and
Evaluation)
The Pentagon, Room 3C125
Washington, DC 20301

Director, Defense Advanced Research
Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, Cybernetics Technology Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Director, ARPA Regional Office (Europe)
Headquarters, U.S. European Command
APO New York 09128

Director, ARPA Regional Office (Pacific)
Staff CINCPAC, Box 13
Camp H. M. Smith, Hawaii 96861

Dr. Don Hirta
Defense Systems Management School
Building 202
Ft. Belvoir, VA 22060

Chairman, Department of Curriculum
Development
National War College
Ft. McNair, 4th and P Streets, SW
Washington, DC 20319

Defense Intelligence School
Attention: Professor Douglas E. Hunter
Washington, DC 20374

Vice Director for Production
Management Office (Special Actions)
Defense Intelligence Agency
Room 1E963, The Pentagon
Washington, DC 20301

Command and Control Technical Center
Defense Communications Agency
Attention: Mr. John D. Hwang
Washington, DC 20301

Department of the Navy

Office of the Chief of Naval Operations
(OP-951)
Washington, DC 20450

Office of Naval Research
Assistant Chief for Technology (Code 200)
800 N. Quincy Street
Arlington, VA 22217

Office of Naval Research (Code 230)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Naval Analysis Programs (Code 431)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Operations Research Programs (Code 434)
800 North Quincy Street
Arlington, VA 22217

Office of Naval Research
Information Systems Program (Code 437)
800 North Quincy Street
Arlington, VA 22217

Director, ONR Branch Office
Attention: Dr. Charles Davis
536 South Clark Street
Chicago, IL 60605

Director, ONR Branch Office
Attention: Dr. J. Lester
495 Summer Street
Boston, MA 02210

Director, ONR Branch Office
Attention: Dr. E. Gloye
1030 East Green Street
Pasadena, CA 91106

Director, ONR Branch Office
Attention: Mr. R. Lawson
1030 East Green Street
Pasadena, CA 91106

Office of Naval Research
Scientific Liaison Group
Attention: Dr. M. Bertin
American Embassy - Room A-407
APO San Francisco 96503

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps (Code RD-1)
Washington, DC 20380

Headquarters, Naval Material Command
(Code 0331)
Attention: Dr. Heber G. Moore
Washington, DC 20360

Dean of Research Administration
Naval Postgraduate School
Attention: Patrick C. Parker
Monterey, CA 939-0

Superintendent
Naval Postgraduate School
Attention: R. J. Roland, (Code 52R1)
C^3 Curriculum
Monterey, CA 93940

Naval Personnel Research and Development
Center (Code 305)
Attention: LCDR O'Bar
San Diego, CA 92152

Navy Personnel Research and Development
Center
Manned Systems Design (Code 311)
Attention: Dr. Fred Muckler
San Diego, CA 92152

Naval Training Equipment Center
Human Factors Department (Code N215)
Orlando, FL 32813

Naval Training Equipment Center
Training Analysis and Evaluation Group
(Code N-00T)
Attention: Dr. Alfred F. Smode
Orlando, FL 32813

Director, Center for Advanced Research
Naval War College
Attention: Professor C. Lewis
Newport, RI 02840

Naval Research Laboratory
Communications Sciences Division (Code 54
Attention: Dr. John Shore
Washington, DC 20375

Dean of the Academic Departments
U.S. Naval Academy
Annapolis, MD 21402

Chief, Intelligence Division
Marine Corps Development Center
Quantico, VA 22134

Department of the Army

Deputy Under Secretary of the Army
(Operations Research)
The Pentagon, Room 2E611
Washington, DC 20310

Director, Army Library
Army Studies (ASDIRS)
The Pentagon, Room 1A534
Washington, DC 20310

U.S. Army Research Institute
Organizations and Systems Research Laboratory
Attention: Dr. Edgar M. Johnson
5001 Eisenhower Avenue
Alexandria, VA 22333

Director, Organizations and Systems
Research Laboratory
U.S. Army Institute for the Behavioral
and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Technical Director, U.S. Army Concepts
Analysis Agency
8120 Woodmont Avenue
Bethesda, MD 20014

Director, Strategic Studies Institute
U.S. Army Combat Developments Command
Carlisle Barracks, PA 17013

Commandant, Army Logistics Management Center
Attention: DRXMC-LS-SCAD (ORSA)
Ft. Lee, VA 23801

Department of Engineering
United States Military Academy
Attention: COL Ar F. Grum
West Point, NY 10996

Marine Corps Representative
U.S. Army War College
Carlisle Barracks, PA 17013

Chief, Studies and Analysis Office
Headquarters, Army Training and Doctrine
Command
Ft. Monroe, VA 23351

Commander, U.S. Army Research Office
(Durham)
Box CM, Duke Station
Durham, NC 27706

Department of the Air Force

Assistant for Requirements Development
and Acquisition Programs
Office of the Deputy Chief of Staff for
Research and Development
The Pentagon, Room 4C331
Washington, DC 20330

Air Force Office of Scientific Research
Life Sciences Directorate
Building 410, Bolling AFB
Washington, DC 20332

Commandant, Air University
Maxwell AFB, AL 36112

Chief, Systems Effectiveness Branch
Human Engineering Division
Attention: Dr. Donald A. Topmiller
Wright-Patterson AFB, OH 45433

Deputy Chief of Staff, Plans, and
Operations
Directorate of Concepts (AR/XOCCC)
Attention: Major R. Linhard
The Pentagon, Room 4D 1047
Washington, DC 20330

Director, Advanced Systems Division
(AFHRL/AS)
Attention: Dr. Gordon Eckstrand
Wright-Patterson AFB, OH 45433

Commander, Rome Air Development Center
Attention: Mr. John Atkinson
Griffis AFB
Rome, NY 13440

IRD, Rome Air Development Center
Attention: Mr. Frederic A. Dion
Griffis AFB
Rome, NY 13440

HQS Tactical Air Command
Attention: LTCOL David Dianich
Langley AFB, VA 23665

Other Government Agencies

Chief, Strategic Evaluation Center
Central Intelligence Agency
Headquarters, Room 2G24
Washington, DC 20505

Director, Center for the Study of
Intelligence
Central Intelligence Agency
Attention: Mr. Dean Moor
Washington, DC 20505

Mr. Richard Heuer
Methods & Forecasting Division
Office of Regional and Political Analysis
Central Intelligence Agency
Washington, DC 20505

Office of Life Sciences
Headquarters, National Aeronautics and
Space Administration
Attention: Dr. Stanley Deutsch
600 Independence Avenue
Washington, DC 20546

Other Institutions

Department of Psychology
The Johns Hopkins University
Attention: Dr. Alphonse Chapanis
Charles and 34th Streets
Baltimore, MD 21218

Institute for Defense Analyses
Attention: Dr. Jesse Orlansky
400 Army Navy Drive
Arlington, VA 22202

Director, Social Science Research Institute
University of Southern California
Attention: Dr. Ward Edwards
Los Angeles, CA 90007

Perceptronics, Incorporated
Attention: Dr. Amos Freedy
6271 Varial Avenue
Woodland Hills, CA 91364

Stanford University
Attention: Dr. R. A. Howard
Stanford, CA 94305

Director, Applied Psychology Unit
Medical Research Council
Attention: Dr. A. D. Baddeley
15 Chaucer Road
Cambridge, CB 2EF
England

Department of Psychology
Brunel University
Attention: Dr. Lawrence D. Phillips
Uxbridge, Middlesex UB8 3PH
England

Decision Analysis Group
Stanford Research Institute
Attention: Dr. Miley W. Merkhofer
Menlo Park, CA 94025

Decision Research
1201 Oak Street
Eugene, OR 97401

Department of Psychology
University of Washington
Attention: Dr. Lee Roy Beach
Seattle, WA 98195

Department of Electrical and Computer
Engineering
University of Michigan
Attention: Professor Kan Cher
Ann Arbor, MI 94135

Department of Government and Politics
University of Maryland
Attention: Dr. Davis B. Bobrow
College Park, MD 20747

Department of Psychology
Hebrew University
Attention: Dr. Amos Tversky
Jerusalem, Israel

Dr. Andrew P. Sage
School of Engineering and Applied
Science
University of Virginia
Charlottesville, VA 22[illegible]

Professor Raymond Tanter
Political Science Department
The University of Michigan
Ann Arbor, MI 48109

Professor Howard Raiffa
Morgan 302
Harvard Business School
Harvard University
Cambridge, MA 02163

Department of Psychology
University of Oklahoma
Attention: Dr. Charles Gettys
455 West Lindsey
Dale Hall Tower
Norman, OK 73069

Institute of Behavioral Science #3
University of Colorado
Attention: Dr. Kenneth Hammond
Room 201
Boulder, Colorado 80309

Decisions and Designs, Incorporated
Suite 600, 8400 Westpark Drive
P.O. Box 907
McLean, Virginia 22101