MANAGEMENT SCIENCE
Vol. 15, No. 11, July, 1969
Printed in U.S.A.

DECISION MAKING IN A COMPLEX ENVIRONMENT: THE USE OF SIMILARITY JUDGEMENTS TO PREDICT PREFERENCES*[1]

DAVID KLAHR

University of Chicago

Judgments of the relative similarity of pairs of alternatives are used to construct a model of the decision space of a group of college admissions officers. This model is then used to predict the preferences of the officers. The accuracy of the predictions supports the hypothesis that preference judgments are made on the basis of the similarity of given alternatives to an "ideal" alternative. A nonmetric multidimensional scaling procedure is used to construct the space. This procedure yields a dimensional representation based upon very few assumptions about the nature of the similarity measures.

I. Introduction

Measurements of the degree of sameness or differentness among stimuli abound in psychological research.[2] Such measurements yield information about the similarity of one pair of stimuli relative to another pair. Coombs (1964, p. 441) points out that similarities data is important to theoretical psychology in two senses:

i. As a means of studying cognitive structure. This use of similarities data is relatively new, and has been greatly facilitated by recent advances in multidimensional scaling (Shepard, 1962a, b; Kruskal, 1964a, b; Torgerson, 1965; Lingoes, 1964).

ii. As an explanatory principle for such psychological phenomena as perceptual organization, association, transfer and retroactive inhibition (Atteneave, 1950).

This paper describes the use of similarities data in both of the above senses: to study the cognitive structure of a group of college admissions officers; and to use the structure obtained to explain the preferences of the officers for a set of hypothetical college applicants.

The primary result of this research is an empirical demonstration that supports the following intuitively plausible but thus far untested theoretical position.

> A set of alternatives can be represented as a set of points in a multidimensional space. In that same space there exists for each individual an ideal object (*i.e.*, one that, if it existed would always be preferred to all other objects). The individual's preference ordering of alternatives is simply the inverse of the ordering of distances in the space from the ideal object to all alternatives.

---

* Received September 1967 and revised January 1968.

[1] This is part of a paper presented at the Fourteenth International Meetings of the Institute for Management Science, in August, 1967, Mexico City. It is based upon my Ph.D. dissertation (Klahr, 1968). I am indebted to my committee, especially Herbert A. Simon and Kenneth R. MacCrimmon for criticism and support; and to my Chicago colleagues, R. Weil, S. Becker, and P. Goodman, for their comments on an earlier draft. The research was supported in part by a Ford Foundation Fellowship. The computer work was done at the Computation Centers at both Carnegie-Mellon University and the University of Chicago.

[2] Although measurements of similarity have been commonly used in psychological research (*e.g.*, Messick, 1956; Shepard, 1964), as Atteneave (1950) points out: "The question 'what makes things seem alike or different?' is so fundamental that very few psychologists have been naive enough to ask it." In recent years, the development of multidimensional scaling techniques has facilitated the investigation of this fundamental question.

This position has been developed by Coombs and his co-workers (Coombs, 1950, 1964; Bennett and Hays, 1960) under the rubric of the "unfolding technique." The empirical applications of Coombs' ideas have been rather limited, however, due to the difficulty of implementing the procedures.[3] The approach described in this paper, while based upon the theoretical position outlined above, follows a very different procedure. Coombs' technique takes as input the judges' preference orderings for alternatives and yields as output the spatial representation that is consistent with those inputs. Our approach takes as input a set of similarity judgments and an independent determination of an ideal object; it yields as output a spatial representation that is consistent with the similarity data and a prediction of the preference ordering.[4] The prediction is then tested against an independently obtained set of preference judgments. The results indicate that a spatial configuration based upon judgments of similarity can be used to predict preference.

A secondary result of the research is a contribution to the clinical judgment area: the determination of the evaluation function that can be used to characterize the college admissions officers. This study of a set of experienced decision makers behaving in a complex, realistic, and familiar environment is an example of the trend in investigations of human judgment to move out of the artificially contrived circumstances of the laboratory into the realistic complexity of everyday decision making.[5]

II. Definitions and Theory

The psychological space, E, of the decision maker consists of a set of multidimensional objects, S, and an ideal multidimensional object, O'. A decision consists of the identification of an object in E by the decision maker as being closest to, or most similar to, his subjectively perceived O'. The objects in S and the ideal object O' are composed of many attributes (or dimensions, factors, components, *etc.*).[6]

For most decisions no object in E exactly matches O'. If there is an object that is closest to O' on all dimensions, then it is the most preferred or most suitable object. However, the characteristic of most decisions that makes them nontrivial is the fact that different objects are closest to O' along different dimensions. The decision maker

[3] The multidimensional unfolding technique has the rather restrictive requirement that there be wide variation in the preference orderings of the individual judges. Multidimensional unfolding assumes a single spatial configuration of alternatives with ideal objects (in Coombs' model the judges themselves) scattered widely throughout the space (Bennett and Hays, 1960, p. 78). The technique used in this paper contains the possibility of a different configuration for each judge, or a single configuration with judges widely dispersed, or a single configuration with a single ideal for all judges. Furthermore, the unfolding technique yields nonmetric results, whereas the Shepard-Kruskal technique yields metric results: the interpoint distances in the configuration.

[4] The procedure that constructs a spatial representation from the rank ordering of the similarities data is a version of nonmetric multidimensional scaling developed by Roger Shepard (1962a, b) and improved by Joseph Kruskal (1964a, b). We are indebted to Messrs. Shepard and Kruskal for providing listings and operating instructions for their computer programs.

[5] C. F., the evolution of the research program in clinical judgment at the Oregon Research Institute, summarized by Hoffman (1968).

[6] In this paper we use both vector notation and attribute-value notation for describing the objects. The latter is convenient when we want to explicitly name the dimensions under consideration. Thus, we might describe an object as $X = (a_1{:}v_1, a_2{:}v_2, \ldots, a_n{:}v_n)$, where the a's are names of attributes and the v's are values of those attributes. These values can range in specificity from ratio scales to equivalence classes. For example, a personnel selection decision might have O' = (age: mid-40's, sex: male, IQ: 128, experience: over 20 years, . . .). For preferences for automobiles we might have O' = (top speed: 90, cost: 0, color: red, . . .). In most of the empirical work on decision making, O' has the attributes of probability, payoff, and events.

must assign to each of the alternatives a number or label according to its overall distance from O'. This is functionally equivalent to mapping some of the multidimensional objects from E into either the set of reals or a set of equivalence classes. The object that obtains the minimum value from the evaluation process is the selected object. The equivalence classes may be as simple as an accept-reject dichotomy, in which case the ideal object is equivalent to a set of acceptable levels along each attribute.

We postulate some evaluation function $F(S_i - O')$ that yields a measure of the proximity of the i^{th} alternative to O'. That is, in a choice between alternatives S_1 and S_2, the decision maker chooses the S_i that minimizes $F(S_i - O')$, $i = 1, 2$; where O' is the ideal object and F is the *preference evaluation function*.

We further postulate a *similarity evaluation function* that is the same as the preference evaluation function. That is, if two pairs of alternatives, (S_a, S_b) and (S_x, S_y), are judged as to relative similarity, the decision maker will designate pair (S_a, S_b) as more alike than pair (S_x, S_y) if $F(S_a - S_b)$ is less than $F(S_x - S_y)$. We can test this position by empirically obtaining both similarity and preference data on a set of alternatives. The two measures should be related to the following way. From the set of similarity measures we can construct a spatial configuration in which each point in the space represents one of the alternatives and in which the points are arranged so that the inverse rank order of interpoint distances in the space corresponds to the rank order of similarities given in the input data. In this configuration the two closest points (*i.e.*, least interpoint distance) correspond to the two alternatives that were judged most similar, the two points farthest apart correspond to the two alternatives that were least similar, *etc.*

Assume that we have obtained such a *similarity* generated configuration; what should we expect to find when we compare this similarity configuration to the *preference* data? If we locate the ideal object in this same space, then we should find that the preference ordering of the alternatives is directly related to the ordering of distances in the space from the ideal object to each alternative. For prediction of preferences we need only a set of similarity judgments and the specification of the ideal object. In the next sections we describe such an approach.

III. Definition and Scaling of the Decision Environment

A. The Decision Makers

The five subjects used throughout this study were the four males and one female comprising the admissions staff at an undergraduate college of engineering and science. Their experience and operating procedures tended to foster a high degree of consensus in their everyday decision making.[7]

B. Construction of Basic Environment

The admissions officers were asked to prepare a list of the most important attributes used in making admissions decisions. For each attribute mentioned they were asked

[7] They had been admissions officers at this school from one to ten years. In a typical year they processed almost two thousand applicants. In the year preceding the testing described below, each officer individually read and evaluated all applications. If there was disagreement as to the appropriate action, the officers met and discussed the applicant. Some applications were put into a contingency category and were reread upon the basis of some subsequent information. These meetings led to some tacit communication among the officers (especially from the Director to the other officers) as to the relative weightings that were being put on the various attributes of the applicants.

TABLE 1

Environment Table

Attributes (in alphabetical order)	— Range of Values of Attributes —				
	Greatest Value				Least Value
Alumni Interview	highest	very high	high	above average	average
Campus Interview	9	7	5	3	1
College Board Scores	800	700	600	500	400
Extra-curricular Leadership/Membership	several/ several	some/ several	none/ several	none/ some	none/ none
High School Grade Average	A	B+	B	B−	C
High School Recommendation	superior	excellent	very good	good	average
IQ	150	140	135	120	110
Rank in Senior Class	top 5%	top 10%	top 20%	top 25%	top 33%

to list typical high, average and low values. From the officers' responses we constructed a table of eight attributes with five discrete values for each attribute (see Table 1).

This Environment Table (E-Table) was constructed on the basis of three heuristics. One was simply to use the most frequently mentioned attributes. The second was to represent both values of two dichotomous meta-attributes: quantitative-qualitative values, and nonacademic-academic attributes. (Of course these are not independent, and three of the academic attributes—Boards, IQ, and Rank—have quantitative values.) The third was to use a set of relatively independent attributes. It is difficult to meet this criterion with the academic attributes, but this heuristic excluded such derived attributes as predicted grade point average. It should be noted at this time that the E-Table is used only as a "rough" approximation to the decision environment. No assumptions are made as to the relative importance of attributes, their actual interdependence or the relative value of each entry alone as an attribute. The crucial assumption—that there is no attribute that *is* very important that is *not* included in the E-Table—was supported by agreement by all the officers that the E-Table realistically represents the decision environment with which they were typically faced.

C. Unidimensional Scaling of Attributes and Values

The measurements described below determined the relative importance of the eight attributes in the E-Table and the relative importance of the five values of each attribute. Each judge completed rating forms that measured his assessment of the relative importance of attributes and the relative importance of the values of each attribute.

1. *Attribute Scaling.*—In order to obtain a ratio scale of importance along which we can place attributes, we used a procedure proposed by Comrey (1950) and elaborated by Torgerson (1958, pp. 104–112). The basic task for the subject in this approach is to observe a pair of stimuli and, with respect to some attribute, "divide 100 points between them in accordance with the absolute ratio of the greater to the lesser" (Comrey, 1950, p. 317). In our case the attribute is actually the meta-attribute of "overall importance," and the stimuli are the eight attributes in the E-Table.

Judges were then presented with all twenty-eight pairs of attributes (eight attributes, two at a time). The position of each attribute in the pair was randomized. Judges were instructed to compare the *maximum* values of each pair of attributes. Thus, they compared not simply IQ to Rank, for example, but rather they compared the relative importance of an IQ of 150 to a Rank in Class of Top 5 per cent.

TABLE 2

Scaled Values of Nominal Values Averaged over Judges

Alumni Interview	nominal value	highest	very high	high	above average	average
	scale value	1.0	.66	.45	.31	.23
Campus Interview	nominal value	9	7	5	3	1
	scale value	1.0	.67	.42	.25	17
College Boards	nominal value	800	700	600	500	400
	scale value	1.0	.47	.29	.17	.08
Activities	nominal value	several/ several	some/ several	none/ several	none/ some	none/ none
	scale value	1.0	.49	.28	.19	08
Grades	nominal value	A	B+	B	B−	C
	scale value	1.0	.49	.39	.24	.15
High School Recommendation	nominal value	superior	excellent	very good	good	average
	scale value	1.0	.81	.52	.38	.27
IQ	nominal value	150	140	135	120	110
	scale value	1.0	.58	.49	.26	.13
Rank	nominal value	top 5%	top 10%	top 20%	top 25%	top 33%
	scale value	1.0	.64	.27	.19	.09

The data were scaled by the Comrey-Torgerson procedure. This method yields a ratio scale of attributes with an arbitrary origin, and we chose the origin such that the maximum value on the scale was 1.0.

Figure 1 shows the results of the scaling of the attributes. It contains the scaled value of each attribute for each individual judge and for the averaged data from all judges. The averaged scale is derived by averaging the ratios for each pair across all judges, then scaling the resultant average. For every judge, Attributes 3, 8, and 5 (Boards, Rank, and Grades) are always the most important; Attributes 2, 4, 6, and 7 (Campus Interview, Activities, High School Recommendation, IQ) always follow in various permutations; Attribute 1 (Alumni Interview) is always lowest on the scale.

The general conclusions we draw from these intermediate results are that:

a) There are the same three most important attributes for all judges;
b) These important attributes are much more important than the others (*i.e.*, the two highest are from two to ten times more important than the average of the others).

These scale values will be used later in the paper.

2. *Value Scaling.*—For each attribute, we wish to scale the values used in the E-Table. The judges were presented with the ten pairs of values for each of the eight attributes and instructed to assign points as in the earlier sections. The scale values were obtained by averaging the raw point allocation from all the judges. The results are shown in Table 2.

D. Construction of the Alternative Set

We seek data on a set of admissions decisions where the applicants take on a range of values along all their attributes. For this purpose we constructed a set of hypothetical applicants to be used in a series of similarity judgments and preference judgments. The alternatives were constructed by using all combinations of the highest and lowest values for the four most important nominal attributes (Boards, Grades, IQ, and Rank, indicated by the average results in Figure 1). The other four attributes were radomly assigned values from the E-Table. The sixteen alternatives thus generated are shown

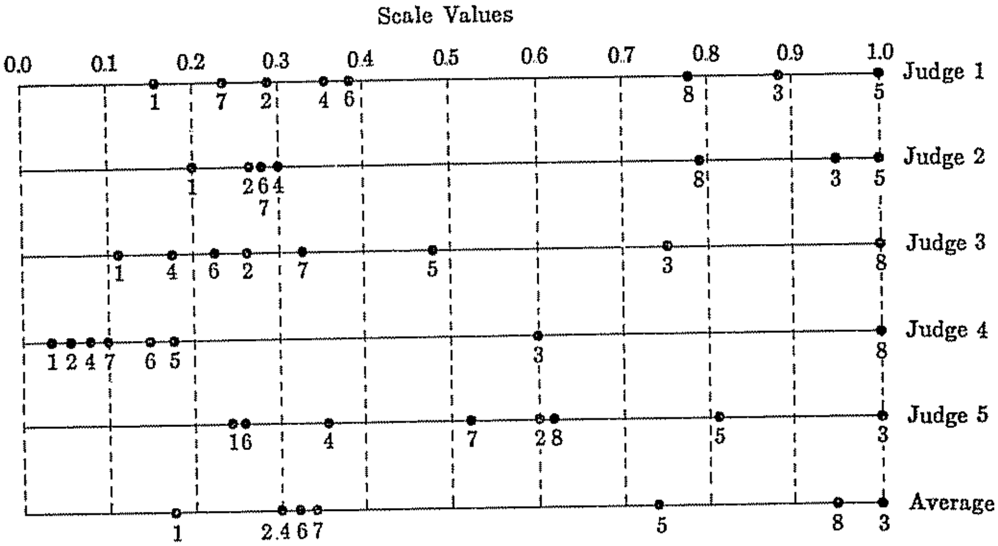


FIG. 1. Scaled attributes with maximum values

in Table 3. In the following sections we will describe the use of these alternatives; it is worth noting here that for both the similarity and the preference judgments, the alternatives are presented in pairs (120 in all) and the systematic variation in values did not seem to be immediately obvious to the judges.

IV. Predicting Preference from a Similarity Based Configuration

A. Similarity Judgments: Procedure

All distinct pairs of the sixteen alternatives described above are presented to the judges. The 120 pairs are presented on 120 slips of paper, with the order of presentation and position randomized. The judge is instructed to group the pairs into eight "degrees of similarity" by repeatedly dividing the 120 pairs into subgroups of pairs according to the relative overall similarity of one pair to another.[8]

When the judge has sorted the pairs of alternatives into eight piles, the results are scored by assigning the numbers one through eight to the slips in the first through eighth piles. The smaller the number assigned to a slip, the greater the similarity perceived by the judge between the two alternatives represented on that slip. This eight point scale of similarity is used as the prime input to the nonmetric multidimensional scaling procedure described in the next section.

B. Multidimensional Scaling: Analysis

The above procedure yields, for each judge and for the average judge, a measure of the perceived similarity of each alternative to all other alternatives. Interpreting the similarity data as a measure of proximity, we then attempt to construct a spatial configuration in which more similar alternatives are more proximate than less similar alternatives.

More formally, we seek the following: a spatial configuration where the rank of the interpoint distances is maximally inversely correlated with the rank of the similarity

[8] The slips are first divided into two piles, of high and low similarity; then each one of the piles is divided into two more; finally the four piles are divided into eight piles. This technique is described by Messick (1956).

TABLE 3

Alternative Set

Attribute	1	2	3	4	5	6	7	8
Alumni Interview	average	highest	high	high	above average	highest	above average	average
Campus Interview	5	5	1	1	5	1	3	9
College Board Scores	800	800	800	800	800	800	800	800
Activ.: Leader/Member	none/sevrl.	none/some	none/some	none/none	none/sevrl.	sevrl./sevrl.	none/sevrl.	some/sevrl.
High School Grades	A	A	C	C	A	A	C	C
High School Recommend	good	excellent	excellent	average	good	very good	average	excellent
IQ	150	110	150	110	150	110	150	110
Rank in Senior Class	top 5%	top 5%	top 5%	top 5%	top 33%	top 33%	top 33%	top 33%

Attribute	9	10	11	12	13	14	15	16
Alumni Interview	highest	high	high	very high	very high	average	very high	above average
Campus Interview	1	9	5	9	9	1	9	9
College Board Scores	400	400	400	400	400	400	400	400
Activ.: Leader/Member	none/sevrl.	some/sevrl.	sevrl./sevrl.	none/sevrl.	none/sevrl.	none/none	none/sevrl.	none/sevrl.
High School Grades	A	A	C	C	A	A	C	C
High School Recommend	average	good	excellent	good	good	superior	very good	good
IQ	150	110	150	110	150	110	150	110
Rank in Senior Class	top 5%	top 5%	top 5%	top 5%	top 33%	top 33%	top 33%	top 33%

measures. We use a procedure developed by Shepard (1962a, b) and significantly improved by Kruskal (1964a, b) called "Nonmetric Multidimensional Scaling." The procedure starts with an arbitrary configuration of points and iteratively attempts to find some arrangement of the points such that the interpoint distances correspond to the input similarity data.[9]

If, in some configuration, the rank order of the interpoint distances is exactly the opposite of the rank order of the similarity measures, then we have a perfect fit. As the dimensionality of the space is reduced and the solution becomes more highly constrained, we are apt to get some departures from perfect fit. Some of the distances may be "out of order." A measure of departure from perfect fit, called the "stress" of the configuration has been developed by Kruskal (1964a); it is quite similar to a residual sum of squares. From an extensive series of empirical investigations on a variety of data, Kruskal suggests that departures from perfect fit (stress = 0) be interpreted as follows: .025-excellent; .05-good; .10-fair.[10]

The procedure we follow is to find the best fit—the minimum stress—in spaces of decreasing dimensionality. We expect minimum stress to increase as the dimensionality decreases, starting in $n - 1$ space with zero stress. The decision as to which configuration is the most appropriate representation of alternatives rests upon scientific judgment, and is not a direct output of the scaling technique. The decision depends upon the stress, the dimensionality of the space, and the meaningfulness of the final configuration.

We adopt Kruskal's suggested interpretation and seek good fit: stress $\leqq .05$. We also adopt his suggestion to accept that configuration where there is a decrease in the marginal improvement in the fit effected by inclusion of more dimensions. (In Figure 2, the plot of minimum stress versus number of dimensions, this point locates an "elbow" in the curve.)

The meaningfulness of the configuration rests upon what the points in the space represent. Our similarity data is based upon sixteen alternatives that are distinct with respect to the values they have on eight attributes. The worst we could do is find that stress was unacceptably high in all but the fifteen-dimensional configuration. In this case the scaling procedure would have told us that there is no way to represent the alternatives other than to list them with their values. Since the alternatives are described by only eight nominal attributes, we would expect that no more than eight dimensions would be necessary for a good fit. However, we have reason to expect good fit in a much lower dimensionality. Recall that the alternatives were generated by systematically varying the values of the four most important nominal attributes, and randomly varying the values of the others. We would expect, then, that no more than eight dimensions will be necessary.

[9] It is always possible to find some arrangement of n points in Euclidean space of $n - 1$ dimensions that satisfies the similarity data (Bennett and Hays, 1960, pp. 37–38). For example, given interpoint distances of four points, we can always find some 3-dimensional configuration that satisfies the ordinal relationship among the interpoint distances. However, an $n - 1$ dimensional representation of n points is not a more parsimonious representation of the data than the original $n(n - 1)/2$ similarity measures. What we seek is a configuration of very few dimensions that still satisfies the similarity data. As we reduce the dimensionality of the configuration it becomes more difficult to find a configuration that fits the data. At the same time, any configuration that does fit is likely to be unique, for we have $n(n - 1)/2$ measures to fit only n points. In our case, we have 120 similarity judgments with which to fit sixteen points in some configuration.

[10] Estimates of the statistical significance of these stress levels have recently been derived (Klahr, 1969). They indicate that the results presented here are of high statistical significance.

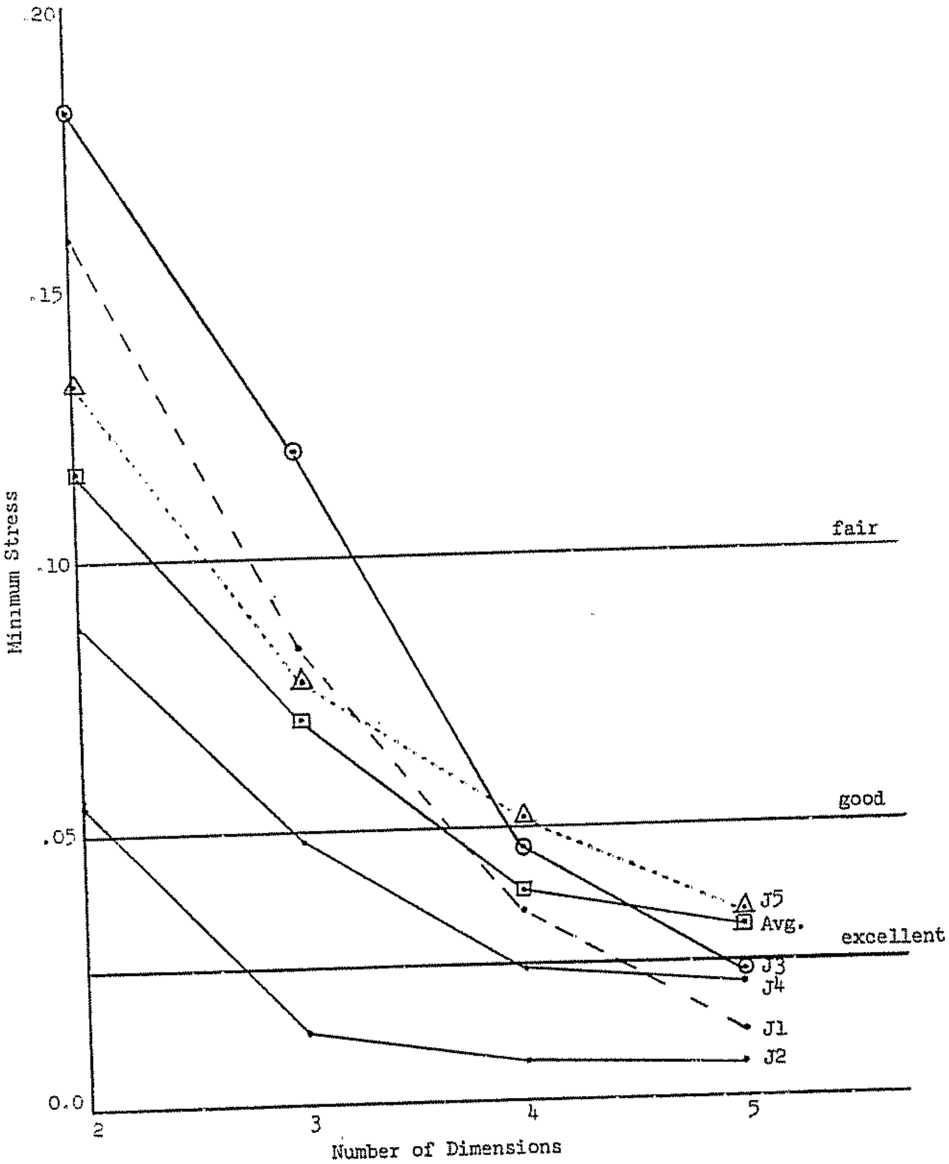


FIG. 2. Stress vs. number of dimensions (from Table 4)

C. Multidimensional Scaling: Results

In this section we describe the final configurations, their goodness of fit, their dimensionality and the relationship of the nominal attributes and values to the final configurations.

Based upon the above criteria, we found that a four dimensional configuration of the sixteen objects was the appropriate representation of the cognitive structure of the judges. This was determined by finding the minimum stress for each judge, and for the average of all judges, in spaces varying in dimensionality from five to two. The results are tabulated in Table 4 and plotted in Figure 2. Notice that Figure 2 is, in effect, a plot of goodness of fit versus number of degrees of freedom; each point represents the

TABLE 4

*Minimum Stress for Each Judge in Spaces from 2 to 5 Dimensions***

	No. of Dimensions			
	2	3	4	5
Judge 1	.159	.084	.035	.011
Judge 2	.055	.013	.007	.006
Judge 3	.182	.119	.046	.023
Judge 4	.088	.048	.024	.021
Judge 5	.132	.077	.051	.032
Average Judge*	.116	.070	.038	.031

* Note: This is not a column average. It is the minimum stress resulting from scaling the similarity data averaged over the five judges.

** Each cell represents the stress of the final configuration of the 16 alternatives. The four-dimensional configuration for the Average Judge (Stress = .038) is plotted in Figure 3.

TABLE 5

Coordinates of Alternatives in Final Four-Dimensional Configuration for Average Judge

Alternative	Dimension			
	1	2	3	4
1	0.658	0.538	0.363	0.386
2	0.589	0.644	0.299	−0.258
3	0.677	0.670	−0.248	0.320
4	0.619	0.625	−0.337	−0.303
5	0.687	−0.524	0.252	0.238
6	0.601	−0.610	0.378	−0.122
7	0.610	−0.554	−0.295	0.456
8	0.573	−0.656	−0.321	−0.388
9	−0.548	0.647	0.435	0.343
10	−0.657	0.545	0.219	−0.468
11	−0.483	0.581	−0.363	0.520
12	−0.726	0.516	−0.306	−0.320
13	−0.622	−0.611	0.188	0.251
14	−0.656	−0.601	0.552	−0.443
15	−0.623	−0.568	−0.451	0.193
16	−0.699	−0.641	−0.364	−0.404
Cluster Centers	±.627	±.596	±.336	±.339
Scatter	.058	.049	.089	.102

Note: For each dimension k, the Cluster Center is $x_k^* = \frac{1}{16}\sum|x_{i,k}|$. The Scatter is the sum of the variances of positive values about $+x^*$ and negative values about $-x^*$. Or, equivalently,

$$\text{Scatter} = (\tfrac{1}{16}\sum(|x_{i,k}| - x_k^*)^2)^{1/2}.$$

These equations are only valid when there are the same number of points in each cluster.

stress of that configuration of sixteen points in n dimensions which best fits the similarity data for a particular judge. Figure 2 does not tell us anything about what these configurations actually look like, nor how similar or different they are from one another. For all five judges and for the "average judge" four dimensions are sufficient to obtain a good fit. There is a noticeable elbow in the plot at four dimensions for all but Judge 5.

The four-dimensional configuration for the "average judge" (see discussion below) is presented in Table 5 and in Figure 3. It is important to bear in mind that the essential property of a configuration is the relationship among the interpoint distances of the sixteen points. The configuration was constructed so as to minimize the difference between the rank order of these distances and the rank order of the similarity judgments of the judges, and it can undergo any arbitrary translation, rotation or uniform stretching of axes without changing its meaning.[11] In fact, the configuration has been stretched and moved such that the centroid is at the origin and the root mean square distance from the origin to all points is equal to one.

Before we turn to an interpretation of the configuration, we will determine whether it is meaningful to average the data across judges and discuss the "average judge." Conceptually this means that we assume that all judges are identical, and that individual differences in their responses are due to measurement error or random noise. Whether or not this is a tenable hypothesis can be determined empirically. In the case of the data we have here, we can justify aggregation across judges. The justification rests upon the result of product-moment correlations between the 120 interpoint distances in each judge's configuration and 120 corresponding interpoint distances in the "average" judge's configuration. (It is important to bear in mind that this modal configuration is obtained by averaging the raw similarity data and then performing the multidimensional analysis, *not* by averaging the configurations of the individual judges.) The measure of agreement between any two configurations that we use is the product-moment correlation between the interpoint distances in configuration a and configuration b.[12] If two configurations are identical, except for arbitrary translation, rotation and stretching and shirinking of axes, then the product-moment correlation will be one. For the five judges and for the average judge, the product-moment correlations of the interpoint distances are shown in the first five columns of Table 6. The last three columns of Table 6 contain, respectively, r_{Aj}, $j = 1, 5$: the product-moment correlation between distances in the final configurations for judge j and the average judge; σ_j^2, $j = 1, 5$: the variance of distances in the configuration for judge j; S_j^2, $j = 1, 5$: the standard error of estimate squared. The latter is a measure of how much of the variance in configuration j is unexplained by the variance in the average configuration. These large r's and low standard errors justify the use of the average judge as an estimate of the "true" configuration underlying each individual judge's similarity judgments.[13] We will confine our discussion of the configurations to the results of the average judge (Figure 3; Table 5).

The four-dimensional configuration of the averaged similarity data is plotted in Figure 3, two dimensions at a time. Each numbered point represents the location in the final configuration of the corresponding alternative in Table 3. The entire configuration in Figure 3 corresponds to a single point in Figure 2 (Stress = .038, Number of Dimensions = 4) on the line for the stress of the average judge.

If, in making similarity judgments, subjects responded only to changes in the values

[11] This is only true when a Euclidean metric is used for the distance function.

[12] This is the measure of interconfiguration similarity suggested by Shepard (1966).

$$r_{ab} = \frac{n \cdot \sum_{i<j} (d_{ij}^{a} - \bar{d}_{ij}^{a}) \cdot (d_{ib}^{b} - \bar{d}_{ij}^{b})}{([\sum_{i<j} (d_{ij}^{a} - \bar{d}_{ij}^{a})^2][\sum_{i<j} (d_{ji}^{b} - \bar{d}_{ij}^{b})^2])^{1/2}}$$

[13] This finding is consistent with the earlier comments about the tendency toward consensus that prevailed in the admissions office.

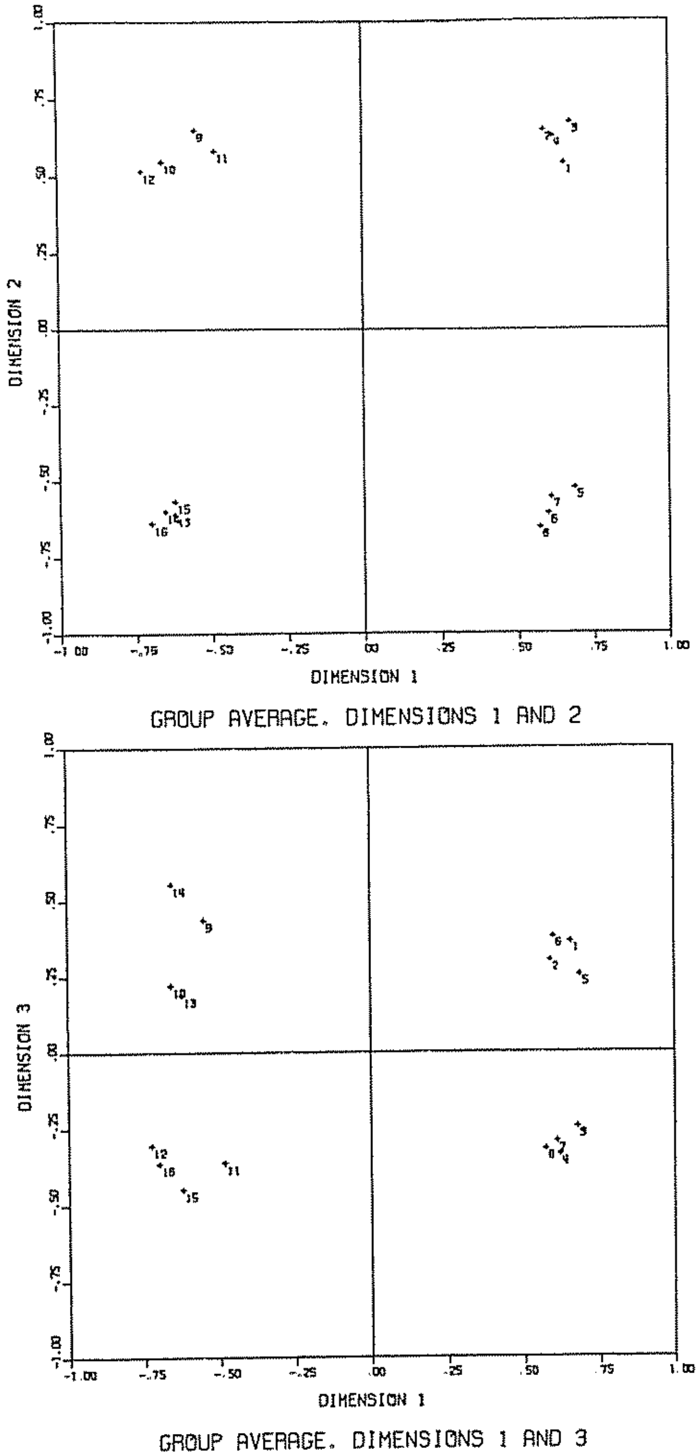
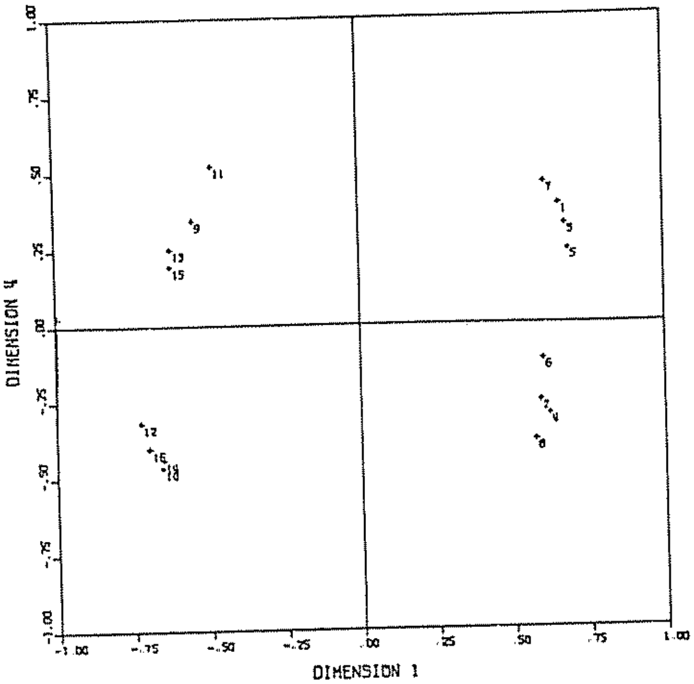
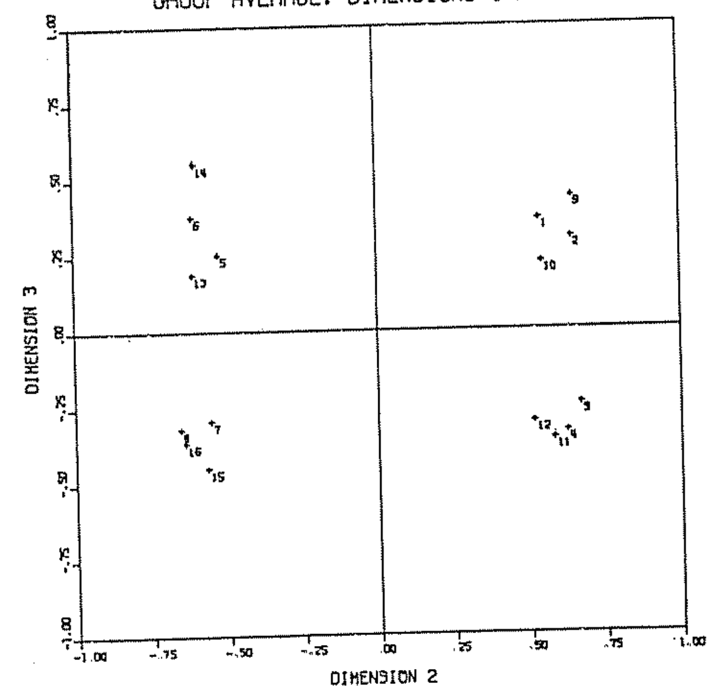


FIG. (3a)

FIG. 3. Group average; a. dimensions 1 and 3; b. dimensions 2 and 3; c. d!mensions 3 and 4.

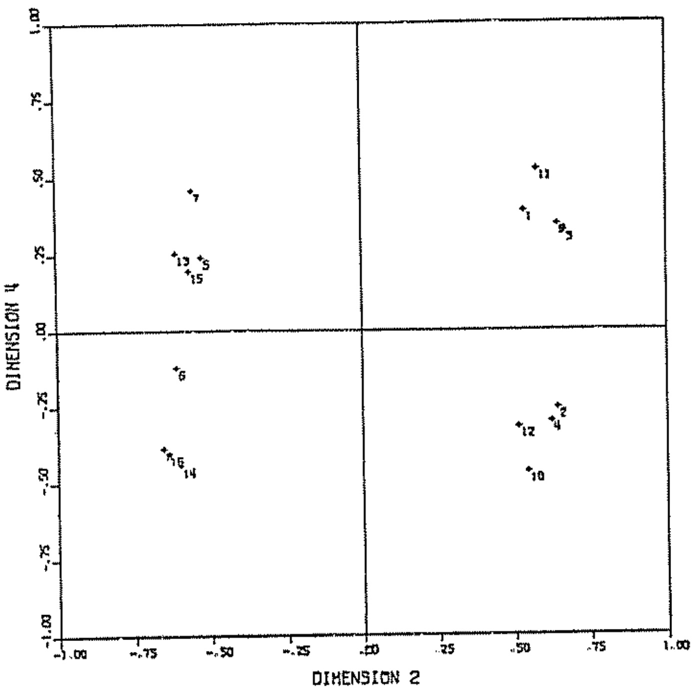


GROUP AVERAGE. DIMENSIONS 1 AND 4

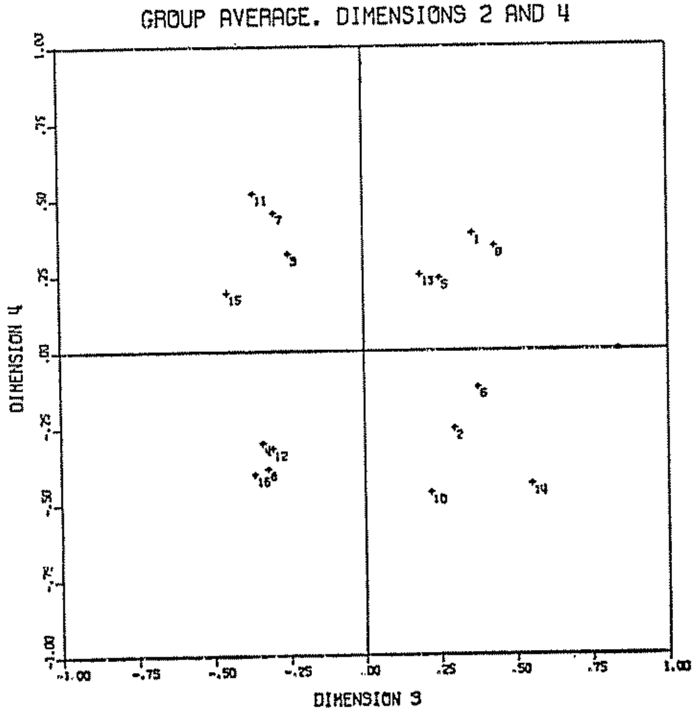


GROUP AVERAGE. DIMENSIONS 2 AND 3

FIG. (3b)



GROUP AVERAGE. DIMENSIONS 2 AND 4



GROUP AVERAGE. DIMENSIONS 3 AND 4

FIG. (3c)

TABLE 6

Product-Moment Correlations Between Interpoint Distances in all Final Four-Dimensional Configurations

Judge	Judge					σ_j^2	S_j^2
	2	3	4	5	Average		
1	.67	.71	.62	.71	.85	.13	.03
2		.58	.69	.67	.86	.17	.04
3			.56	.58	.75	.14	.06
4				.71	.85	.21	.05
5					.82	.15	.05

Notes: 1. For all correlations $n = 120$. 2. σ_j^2 is the variance of the distances in the configuration of Judge j. 3. $S_j^2 = \sigma_j^2(1 - r_{aj}^2)$, the standard error of estimate squared. 4. Each cell in the column titled "Average" contains r_{aj}—the product-moment correlation of distances in configuration for Judge j, with corresponding distances in the average judge's configuration.

of the four systematically varied attributes, and if the differential changes in the values of each attribute were judged to be equivalent, then the configuration would be a four-dimensional hypercube. If such a "perfect" configuration, consisting of an alternative at each vertex of the hypercube, were plotted in this fashion, each of the six plots in Figure 3 would have four points symmetrically placed in each quadrant. Every point would correspond to the four alternatives having the pair of values represented by the projections on the axes. Thus, for example, the upper right hand quadrant of Dimension 2 versus Dimension 1 would consist of a single point corresponding to the four alternatives that had high values of the two attributes corresponding to Dimensions 1 and 2. It is evident from Figure 3 that the points do form four clusters in each panel, although the clustering is poorer for higher dimensions.[14]

We can establish a correspondence between the dimensions in the configuration and the nominal attributes of the alternatives represented therein by examining Figure 3. In the top panel of Figure 3a, points 1, 2, 3, and 4 are simultaneously high on Dimensions 1 and 2. Alternatives 1, 2, 3, and 4 are the only alternatives that are high on both College Board Scores and Rank in Senior Class (see Table 3). In the bottom panel of Figure 3b, points 3, 4, 11 and 12 are simultaneously high on Dimension 2 and low on Dimension 3. Only Alternatives 3, 4, 11, and 12 are simultaneously high on Rank and low on Grades. Similar comparisons for the other panels of Figure 3 yield the following identifications: Dimension 1-College Board Scores; Dimension 2-Rank in Senior Class; Dimension 3-High School Grades; Dimension 4-IQ.

The distribution of points along each dimension provides information analogous to the mean and standard deviation of a correlation coefficient. The first measure we call the "cluster center" for each dimension. It is obtained by averaging the positive distances on the dimension. (Recall that the normalization of the configuration puts the centroid of the configuration at the origin, so that on each dimension $\bar{x} = 0$.) The larger the value for the cluster center, the greater the relative importance of the spread of nominal values on the corresponding dimension. The second measure is the standard

[14] All the configurations were started from an initial configuration corresponding to the "perfect" arrangement. This initial configuration converged to a final configuration that was identical, with respect to interpoint distances, to final configurations obtained from Kruskal's arbitrary initial configuration. The advantage lies in facilitating the identification of nominal attributes with the configuration axes. This technique is legitimate in Euclidean spaces (cf. Kruskal, 1964a, pp. 14, 23).

deviation of both positive and negative distances about the corresponding positive and negative cluster centers. The results of these computations are shown in Table 5, where columns 1 to 4 correspond to Dimensions 1 to 4 in Figure 3. The cluster centers for Dimensions 1 and 2 are further apart than those for Dimensions 3 and 4, indicating that the differences between maximum and minimum values of Board Scores and Rank in Class are perceived as greater than the maximum and minimum values of High School Grades and IQ. The variation around the cluster centers is also presented in Table 5 in the row labeled "Scatter." This measure indicates that Boards and Rank more consistently affect similarity judgments than do Grades and IQ.

The form of the similarity evaluation function is obtained directly from the final configuration. By constructing the configuration such that the interpoint distances agreed with the similarity judgments, we found parameters for the function:

$$F(S_a - S_b) = [\sum_{j=1}^{n} (S_{aj}' - S_{bj}')^r]^{1/r}.$$

S_{aj}' and S_{bj}' are the perceived values of perceived attribute j for alternatives a and b; n is the number of perceived relevant attributes used in the evaluation function; r is the spatial metric; and $F(S_a - S_b)$ is the perceived similarity between alternatives S_a and S_b. We have already discussed the finding that $n = 4$. The analysis of the final configuration enabled us to identify the nominal attributes with the four dimensions required in the final configurations. The value of r is 2, indicating that we are using a Euclidean distance function.[15]

The relationship between the nominal values S_{aj}, S_{bj} and the subjectively perceived values S_{aj}', S_{bj}' depends upon the unidimensional "psychophysical" functions measured in Section IIIC. That is,

$$S_{aj}' = f_j(S_{aj}).$$

(See Tversky, 1966 for a rigorous development of this concept.) Thus the most general expression for F is

$$F(S_a - S_b) = [\sum_{j=1}^{n} \{f_j(S_{aj}) - f_j(S_{bj})\}^r]^{1/r}.$$

From the data we have analyzed,

$$F(S_a - S_b) = (\sum_{j=1}^{4} \{f_j(S_{aj}) - f_j(S_{bj})\}^2)^{\frac{1}{2}}.$$

The $f_j(S_{aj})$ values correspond to the cluster centers in the final configurations. The actual form of the f_j can be determined by plotting the nominal values against the scale values of the four quantitative attributes in Table 2.

Summarizing the evaluation of the final configuration, we conclude that the underlying spatial configuration is consistent with a Euclidean metric in a space whose dimensions are in direct correspondence to the nominal attributes and values of the alternatives. In assessing the overall similarity of applicants, judges respond to the attributes Board Scores, Rank, Grades and IQ to a successively decreasing extent.

D. Preference Judgments: Procedure

All distinct pairs of the sixteen alternatives are presented to the judges. The 120 pairs are presented in a 120 page booklet, with the order of occurrence and position

[15] The scaling was done with both $r = 1$ and $r = 2$. For all judges in all dimensions, $r = 2$ always yielded a lower minimum stress than $r = 1$. Of course this does not eliminate the possibility that some other distance function might provide a better fit. The only conclusion we can draw here is that the Euclidean metric ($r = 2$) is more appropriate than the "city block" metric ($r = 1$).

TABLE 7

Preference Averaged over all Judges

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Q Overall Quality
1	0.	0.2	0.8	2.0	1.2	1.8	1.8	2.0	1.4	2.0	1.6	2.2	2.6	2.8	2.6	3.0	28.0
2	−0.2	0.	0.4	1.8	1.4	1.0	1.6	1.6	1.8	1.6	1.8	1.8	2.2	2.0	2.4	3.0	24.2
3	−0.8	−0.4	0.	0.8	0.2	1.0	1.6	1.0	0.4	0.8	1.2	1.8	1.4	1.4	2.2	2.2	14.8
4	−2.0	−1.8	−0.8	0.	0.2	−0.4	0.8	1.0	−0.2	0.6	0.8	1.6	1.0	1.0	1.0	1.8	4.6
5	−1.2	−1.4	−0.2	−0.2	0.	0.4	1.2	0.6	1.2	0.8	1.0	1.0	1.4	2.0	2.0	2.2	10.8
6	−1.8	−1.0	−1.0	0.4	−0.4	0.	0.2	0.4	0.6	1.0	1.2	1.4	1.6	1.6	2.2	2.0	8.4
7	−1.8	−1.6	−1.6	−0.8	−1.2	−0.2	0.	−0.4	−0.2	−0.8	0.	1.2	0.4	1.4	1.2	2.0	−2.4
8	−2.0	−1.6	−1.0	−1.0	−0.6	−0.4	0.4	0.	−0.8	−0.4	0.4	1.2	1.0	0.8	1.6	1.8	−0.6
9	−1.4	−1.8	−0.4	0.2	−1.2	−0.6	0.2	0.8	0.	0.2	0.6	1.0	1.0	1.4	1.0	1.8	2.8
10	−2.0	−1.6	−0.8	−0.6	−0.8	−1.0	0.8	0.4	−0.2	0.	−0.2	0.8	0.6	1.4	1.6	2.0	0.4
11	−1.6	−1.8	−1.2	−0.8	−1.0	−1.2	0.	−0.4	−0.6	0.2	0.	1.4	0.2	1.0	1.6	1.8	−2.4
12	−2.2	−1.8	−1.8	−1.6	−1.0	−1.4	−1.2	−1.2	−1.0	−0.8	−1.4	0.	0.2	0.2	0.8	1.2	−13.0
13	−2.6	−2.2	−1.4	−1.0	−1.4	−1.6	−0.4	−1.0	−1.0	−0.6	−0.2	−0.2	0.	0.8	0.6	1.8	−10.4
14	−2.8	−2.0	−1.4	−1.0	−2.0	−1.6	−1.4	−0.8	−1.4	−1.4	−1.0	−0.2	−0.8	0.	0.2	0.2	−17.4
15	−2.6	−2.4	−2.2	−1.0	−2.0	−2.2	−1.2	−1.6	−1.0	−1.6	−1.6	−0.8	−0.6	−0.2	0.	0.2	−20.8
16	−3.0	−3.0	−2.2	−1.8	−2.2	−2.0	−2.0	−1.8	−1.8	−2.0	−1.8	−1.2	−1.8	−0.2	−0.2	0.	−27.0

TABLE 8

Measures of Interpoint Distances in Final Configuration, two Additive Utility Models, and Overall Quality of Alternatives. Summary Statistics for Prediction of Overall Quality. (See Figure 4 for Plot.)

Alternative	1 Distance	2 Linear Model	3 Quadratic Model	4 Overall Quality
1	0.000	3.30	0.00	28.0
2	0.659	3.32	0.08	24.2
3	0.629	2.52	0.76	14.8
4	0.987	2.11	0.77	4.6
5	1.078	2.61	0.50	10.8
6	1.257	2.76	0.62	8.4
7	1.278	1.67	1.22	−2.4
8	1.581	1.97	1.31	−0.6
9	1.214	2.49	0.69	2.8
10	1.575	2.57	0.74	0.4
11	1.360	2.10	1.48	−2.4
12	1.692	1.68	1.45	−13.0
13	1.734	2.02	1.19	−10.4
14	1.936	1.69	1.27	−17.4
15	1.888	1.22	1.92	−20.8
16	2.094	0.92	1.95	−27.0
Correlation Coefficient	$r_{14} = -.945$	$r_{24} = .942$	$r_{34} = -.931$	$\sigma_4^2 = 222$
Error of Estimate (Squared)	$s_{14}^2 = 23.6$	$s_{24}^2 = 24.8$	$s_{34}^2 = 29.5$	$\sigma_4 = 14.9$

randomized. The judges are instructed to indicate the degree to which one member of each pair is more highly qualified than the other by placing a mark in the appropriate column of a seven point rating scale.

This procedure yields, for each judge, a measure of the perceived qualifications of each alternative compared to every other alternative. We define ${}_iP_{jk}$ as the qualification of alternative j compared to alternative k as rated by judge i. The seven possible responses are scored from -3 to $+3$. If ${}_iP_{jk} > 0$, then j was rated by judge i as being more qualified than k. We measure only for $j > k$ and define $P_{jj} = 0$, $P_{jk} = -P_{kj}$.

The results are presented for the average judge in Table 7. It contains the average of the ${}_iP_{jk}$ across judges. For each entry in Table 7, the average preference is:

$$ {}_aP_{jk} = \tfrac{1}{5} \textstyle\sum_{i=1}^{5} {}_iP_{jk}. $$

The right hand column in Table 7 represents the overall aggregation of the five sets of 120 pairwise preference judgments on sixteen alternatives. We define these row totals in Table 7 as the *Overall Quality* of the alternatives as determined by the preference judgments, that is:

$$ (\text{Overall Quality})_j = \hat{Q}_j = \textstyle\sum_{k=1}^{16} {}_aP_{jk} = \sum_{k=1}^{16} \tfrac{1}{5} \sum_{i=1}^{5} {}_iP_{jk}. $$

(For purposes of cross reference recall that the row and column numbers in this table correspond to the alternative numbers in Table 3.)

E. Prediction of Quality from the Spatial Configuration

The final set of averaged preferences described above (*i.e.*, the $\hat{Q}_j$'s) are taken to be the direct measure of preference for alternatives. In this section we will test the major hypothesis of this paper: that the distances in the final configuration are directly related to the preference measures. Distances in the spatial configuration from the

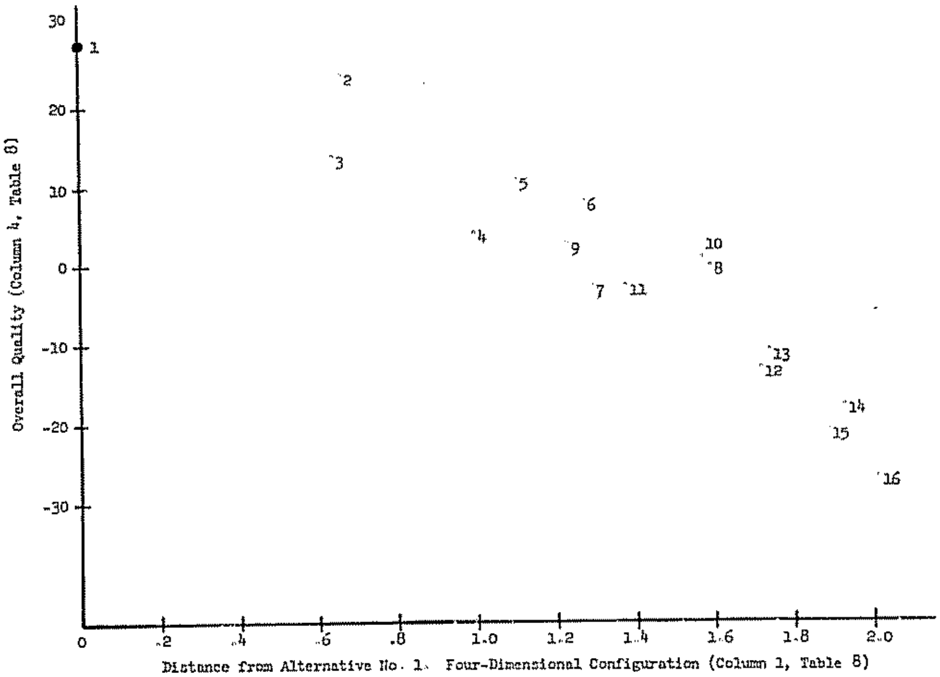


FIG. 4. Overall quality vs average distance

ideal alternative to all other alternatives are computed and compared with the quality data. Since the higher quality alternatives are posited to be closest to the most preferred alternative, we would expect a negative correlation between the two measures. We will also compare the preferences to two different additive utility models based upon the unidimensional scaling of Section IV-C.

The ideal alternative in this case is the one with the maximum value of all relevant attributes: Alternative 1. This is substantiated by the fact that for each individual judge there was no alternative preferred to Alternative 1 in the pairwise comparison, *i.e.*, ${}_iP_{1,k} \geqq 0$, $i = 1, 5$; $k = 1, 16$. The distances from Alternative 1 to all other alternatives in the final configuration are presented in column one, Table 8. The Overall Qualities are presented in column four of the same table. The product-moment correlation between the two columns is $-.94$; the variance of the quality measures is 222; the error of estimate squared of the correlation (using the distances as a predictor of the Qualities) is 23.6.[16] Thus, all but eleven per cent of the variance in Quality is "explained" by the distances in the configuration. A plot of these two measures is presented in Figure 4.

F. Prediction of Preference with Two Additive Utility Models

It is of interest to compare the prediction of Quality based upon configurational distances with Quality predictions based upon a more conventional method. We use two models of additive utility in which the overall worth of an alternative is simply the weighted sum of some function of its value along each of its nominal attributes. The weights are obtained from the scaling results of Section IIIC. We take the scale values of the attributes with maximum value (Figure 1) as the weights, a_i, of the attributes

[16] Correlation coefficients are used here because the arbitrary units of Quality and Distance would render regression coefficients meaningless.

and the scale values of the nominal values (Table 2), V_{ij}, to compute the worth of each alternative. Two additive models are used: one linear, one quadratic. The functions are computed both for individually scaled attributes and values for each judge, and for the group average. First we will discuss only the average judge. For the linear model the worth of an alternative is

$$W_j = \sum_{i=1}^{8} a_i V_{ij},$$

where

$$\begin{aligned} W_j &= \text{worth of alternative } j, \\ a_i &= \text{weight of attribute } i, \\ V_{ij} &= \text{value of alternative } j \text{ on attribute } i. \end{aligned}$$

For the quadratic model, the worth of an alternative is a function of the squared difference on each attribute between alternative one and alternative j.

$$W_1 = 0,$$
$$W_j = \sum_{i=1}^{8} a_i (V_{ij} - V_{i1})^2,$$

where W_j, a_i, V_{ij} are as defined above.

The worths of the alternatives according to these two functions are presented in columns two and three of Table 8. When these values are correlated with the Overall Quality measure, the results are essentially the same as for the correlation of configurational distances with Overall Quality. For the linear model, where a higher worth is directly related to Overall Quality, $r = .94$. For the quadratic model, where lower worths are directly related to Overall Quality, $r = -.93$. Thus we conclude that at the level of aggregation that we have treated the data (*i.e.*, aggregating across judges), the configurational *distances* based upon the analysis of similarity judgments are as good predictors of *Quality* as the more direct techniques based upon additive utility models. Furthermore both techniques are extremely accurate predictors of Quality.

Although earlier in this paper we justified treating the average judge as representative of each judge, we will next consider the relationship between quality, distance and the additive models for each individual judge. Thus we use the interpoint distances in each individual final configuration. In the additive models we use the weights from the *individual* scales in Figure 1 to compute worths for alternatives. Each of these three measures is then correlated with the average quality for the corresponding *individual* judge. The results of these individual correlations, and the aggregate level correlations discussed above, are presented in Table 9.

The results indicate that the distances in each individual's final spatial configuration are highly (negatively) correlated with that same individual's quality judgments.

TABLE 9

Correlation Between Quality and Distance, Quality and Linear Model, and Quality and Quadratic Model. Aggregate and Individual Levels

	Distance	Linear	Quadratic
Group Average	−.95	.94	−.93
Judge 1	−.77	.83	−.80
Judge 2	−.77	.94	−.97
Judge 3	−.72	.76	−.74
Judge 4	−.74	.87	−.81
Judge 5	−.89	.94	−.95

However, for all judges, both additive uy mo ar bettelitPeelsrti predictors than the distances model. Thus, although the results at the individual level support the basic hypothesis that preferences can be obtained from the spatial model, it seems that aggregating over judges obscures the relative efficacy of the spatial model versus the additive utility model.

V. Discussion

In 1954, in a review of theoretical and experimental work on individual decision making, Ward Edwards said "... it seems impossible even to dream of getting experimentally an indifference map in n-dimensional space where n is greater than 3. Even the case of $n = 3$ presents formidable experimental problems" (1954, p. 388). Ten years later, Robert Abelson, writing on theories of choice, said "there is ... a need for more theory and more experiment on the problem of the multidimensionality or multiaspect nature of choice objects" (1964, p. 263).

Both statements emphasize the fact that at the heart of all decision making behavior lies the process whereby the attributes of the alternatives are somehow combined and compared. The research described in this paper is an attempt to develop a methodology for studying this crucial aspect of decision making. The spatial configuration that is generated from the multidimensional scaling procedure consists of alternative choice objects located in a multiattribute space. Each attribute corresponds to a commodity in the usual indifference map terminology, and each alternative is an n-commodity bundle. Once the ideal object has been located in the space, it is possible to define indifference hyperspheres with origin at the ideal.

Fishburn (1967) reviews twenty-four methods for estimating additive utilities for an individual evaluator. The current research can be viewed as a twenty-fifth. The procedure, described in detail in previous sections, is to construct a spatial configuration on the basis of similarity data, to independently locate an ideal object in the configuration, and to postulate a utility or, in our terms, "quality" based only upon distance from the ideal object.

The tenability of such a method rests upon a proposition that is outside utility theory: the proposition that the utility of an object is inversely proportional to its distance from an ideal object in a subjective "decision space." The independent test of the validity of this proposition consists of obtaining an independent set of preference measures and assessing accuracy of the multidimensional configuration's distances as predictors of preference. Such a procedure was described above, and the accuracy was very high.

Two conventional forms of additive utility functions were constructed and they were both somewhat better than the spatial configuration in predicting preference. Since the spatial configuration required more effort than the additive models, one may wonder what the advantage is, especially since the configuration turned out to be fairly close to our intuitive expectations. The answer to this question lies in emphasizing the fact that the spatial configuration is not based upon nominal attributes and values, but rather upon the subjective perception of them by the decision maker. For the hypothetical alternatives that were used, there was high correlation between the nominal and the subjective. However, in more realistic situations, where the values of the attributes are not under arbitrary control of the experimenter, we may not find such a one to one correspondence between the axes of the spatial configuration and the nominal attributes.

There has been some speculation and experimentation in the past few years to determine the appropriate metric for similarity data. (Shepard's 1964a article is one of the best on this topic.) We analyzed the similarity data in both the Euclidean and the city-block metric. In all cases the Euclidean metric provided a better fitting configuration. It has been found that in psychophysical realms such as color or tone perception the Euclidean metric provides better fit than the city-block metric (Kruskal, 1964a, p. 24). However, Shepard noted that as we move toward highly analyzable stimuli—from the perceptual to the conceptual—the Euclidean metric seems to break down. He found that a metric between the Euclidean and the city-block provides the best fit, but even it was inappropriate when shifts in attention focus occurred. Thus, we expected to find the city-block metric a better representation than the Euclidean. In fact, we found the opposite. This leads us to speculate that the Euclidean metric might be appropriate at both ends of a continuum of analyzability. At one end we have the psychophysics situation, *e.g.*, color judgments, in which it is usually impossible for the subject to directly perceive and name the attributes along which the stimuli vary. In the middle we have stimuli such as simple geometric figures varying along a few highly obvious attributes. At the other end we have the kind of "real world" complexity that we have used. We suggest that the subject, when faced with stimuli varying along many incommensurate nominal attributes, is incapable of making the sort of discrete and explicit trade-offs that yield the city-block metric, but instead forms an overall impression that is best captured in a Euclidean metric.

The collection of procedures used in this study comprise a new methodology for the study of decision making. It is therefore appropriate to present a brief methodological critique. The major part of the critique is addressed to the artifacts that reduced the realism of the decision environment. They are not inherent in the procedure, but were introduced to simplify this initial effort.

The procedure introduced a note of artificiality when, in generating the alternatives, high and low values for all four of the major attributes were generated. This led to some highly unrealistic alternatives, such as having an IQ of 150, a Grade Average of A, a College Board Score of 400, and a Rank in Class of Top 33 per cent. This forced independence of attribute values reduced their usefulness to the judges as mutual predictors. Once the judges learned that they were about to encounter these kinds of alternatives, they no longer relied as much upon the level of IQ, for instance as a predictor of College Board Scores, *etc.* The net effect of this artifact upon the multidimensional scaling tends to increase the number of dimensions in the final configuration. We would like to remove this artifact in future work by obtaining the similarity data with *actual* alternatives in whatever decision realm we study.

An important aspect of our arbitrary manipulation of the attribute values is the fact that the fourth attribute whose values we systematically varied, IQ, was far behind the other three on the individual rating forms (Figure 1). By including it with the others, and by making it independent, as noted above, we may have focused more than the usual amount of attention upon it. On the average four dimensions were required to adequately represent the decision space, so that judges were to some extent attending to the variation in IQ. However, as can be seen from Figure 3, and Table 5, the effect of IQ upon spatial distribution was both small and unreliable compared to the other three attributes. Although the arbitrary selection of four of the eight nominal attributes may have overemphasized the usual importance of one of them, the scaling technique is sensitive to those cases where the variation in an attribute was relatively unimportant.

Another problem in our experimental method occurs in the sequence with which we gathered data from the decision makers. The early steps consisted of several tests concerning their judgments about the relative importance of attributes and values of attributes. Only a few weeks later we tested the judges again on similarity and preference measures. It could be the case that the early discussion, testing and explicit questioning about the relative importance of attributes and values led the judges to be more consistent than they might have otherwise been.

Finally, the generality of these results needs to be explored in those situations where preference is not a monotone increasing function of each attribute as was the case in the environment studied here. The relation between similarity and preference in such circumstances may in general be more complex than what we have found.

VI. Conclusion

We have developed a methodology to construct a representation of the subjective decision space of a set of experts operating in a familiar environment. The space, constructed on the basis of similarity judgments, provides accurate predictions of preference. There are several major conclusions that can be drawn on the basis of this research.

1. The technique of nonmetric multidimensional scaling, recently developed for use in the psychophysics laboratory, can be utilized as a powerful tool in decision making research.
2. The underlying dimensions of a complex decision space can be discovered and related to the nominal attributes of the alternatives. To the well-known fact of limited capacity to notice and process we can add information about what parts of the environment are being selected as relevant to the decision maker.
3. The proposition that preference and similarity judgments are made in the same decision space has been subjected to empirical test, and has not been refuted.

There are several possible directions for future research along these lines. The most straightforward would be a repeat of the procedure with experts from another very different environment, *e.g.*, investment analysts or race-track handicappers. In such an extension we would remove most of the artifacts described above by using the actual alternatives with which the decision makers commonly deal.

A shift in emphasis could make this kind of analysis useful in the marketing area. If we can construct a spatial configuration for a class of products (*e.g.*, cars) an if we can also obtain a transitive preference ordering, then we can locate the ideal object in the space. This ideal object would indicate a latent demand for a certain product. The same kind of analysis might be used in conflict situations to determine the subjective spaces of opposing parties dealing with the same set of alternatives. These kinds of applications are speculative at this time, but they seem to be feasible extensions of the methodological developments contained in this research on decision making in complex environments.

References

1. ABELSON, ROBERT P. "The Choice of Choice Theories," in S. Messick and A. Brayfield (eds.), *Decision and Choice*. New York: McGraw-Hill Book Company, Inc., 1964.
2. ATTENEAVE, F. "Dimensions of Similarity," *American Journal of Psychology*, Vol. 65 (1950), pp. 516–556.
3. BENNETT, J. F. AND HAYES, W. L. "Multidimensional Unfolding: Determining the Dimensionality of Ranked Preference Data," *Psychometrika*, Vol. 25 (1960), pp. 27–43.
4. COMREY, A. L. "A Proposed Method for Absolute Ratio Scaling," *Psychometrika*, Vol. 15 (1950), pp. 317–325.
5. COOMBS, C. H. "Psychological Scaling without a Unit of Measurement," *Psychological Review*, Vol. 57 (1950), pp. 145–158. (a)

6. ——. *A Theory of Data*. New York: John Wiley and Sons, Inc., 1964. (b)
7. EDWARDS, W. "The Theory of Decision Making," *Psychological Bulletin*, Vol. 51 (1954), pp. 380–417.
8. FISHBURN, P. C. "Methods of Estimating Additive Utilities," *Management Science*, Vol. 13 (1967), pp. 435–453.
9. HOFFMAN, PAUL J. "Cue-Consistency and Configurality in Human Judgment," in B. Kleinmuntz (ed.), *Formal Representation of Human Judgment*, Wiley, (1968).
10. KLAHR, D. "Decision Making in a Complex Environment." Unpublished Ph.D. dissertation, Carnegie Institute of Technology, 1968.
11. ——. "A Monte Carlo Investigation of the Statistical Significance of Kruskal's Non-Metric Multidimensional Scaling Procedure", *Psychometrika*, (September 1969).
12. KRUSKAL, J. B. "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," *Psychometrika*, Vol. 29 (1964), pp. 1–27. (a)
13. ——. "Nonmetric Multidimensional Scaling: A Numerical Method," *Psychometrika*, Vol. 29 (1964), pp. 28–42. (b)
14. LINGOES, J. C. "An IBM 7090 Program for Guttman Lingoes Smallest Space Analysis-I," *Behavioral Science*, Vol. 10 (1965), pp. 183–184.
15. MESSICK, S. J. "An Empirical Method of Multidimensional Successive Intervals," *Psychometrika*, Vol. 21 (1956), pp. 367–376.
16. SHEPARD, R. N. "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function, I," *Psychometrika*, Vol. 27 (1962), pp. 125–140. (a)
17. ——. "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function, II," *Psychometrika*, Vol. 27 (1962), pp. 219–264. (b)
18. ——. "Attention and the Metric Structure of the Stimulus Space," *Journal of Mathematical Psychology*, Vol. 1 (1964), pp. 54–87. (c)
19. TORGERSON, W. S. *Theory and Methods of Scaling*. New York: John Wiley and Sons, Inc., 1958.
20. ——. "Multidimensional Scaling of Similarity," *Psychometrika*, Vol. 30 (1965), pp. 379–393. (b)
21. TVERSKY, A. "The Dimensional Representation and the Metric Structure of Similarity Data." Unpublished manuscript, Center for Cognitive Studies, Harvard University, Cambridge, Massachusetts, 1966.