ORTHOGONAL ROTATION TO THEORETICAL CRITERIA:

COMPARISON OF MULTIDIMENSIONAL SPACES

Joseph Woelfel
Michigan State University

John Saltiel
Montana State University

Robert McPhee
Michigan State University

Jeffrey E. Danes
Michigan State University

Michael Cody
Michigan State University

George Barnett
Rensselaer Polytechnic Institute

Kim Serota
Arizona State University

Presented at the Annual Convention of the
Mathematical Psychology Association

Purdue, Indiana, August, 1975

The authors would like to acknowledge the help of Thomas Bailey, Edward L. Fink and Donald P. Cushman concerning theoretical and mathematical issues, and Richard Holmes for computer programming.

ORTHOGONAL ROTATION TO THEORETICAL CRITERIA:
COMPARISON OF MULTIDIMENSIONAL SPACES

Although multidimensional scaling techniques have grown in sophistication as well as number over the past decade, very little use has been made of these techniques in experimental psychological work. At least part of the reason for this may be due the ambiguities in the over-time comparison of multidimensional configurations. These ambiguities arise principally from two areas: First, in the popular nonmetric multidimensional algorithms (Kruskal, 1964 ab), the gradient or "steepest decent" interative processes which intervene between data and configuration will always be different in some measure for each of the two (or more) analyses to be compared, and there is no ready way to eliminate this artifactual source of difference. The classical multidimensional model (Torgerson, 1958) does not suffer from this problem, but even in the case where the metric or the nonmetric approach perfectly fits the data to the configuration, the orientation of the axes will be arbitrary (cf. Danes, 1975).

It is now well known in physics that there exists no "priviledged" coordinate system against which physical motions may be described, but the controversey over relative and absolute motion was one of the longest disputes of physical science. The geocentric - heliocentric solar system controversy, for example, consisted entirely in a dispute over whether the reference coordinate system for measuring astronomical motions should be fixed on the earth or relative to the sun and "fixed" stars. Neither coordinate system is given in nature, but the former, which ascribes all motions to celestial objects other than the earth, given rise to equations of motion far too complicated to be practically and theoretically useful to scientists (Einstein, 1961). In general, the choice of reference vectors against which time-ordered data will be arrayed constitutes a decision as to how apparent motions will be differentially attributed to the data points.

In multidimensional analyses, the locations of points in a cognitive space are usually associated by some theory with the "meanings" of the concepts they represent. Changes in the locations of these points over time, therefore, represent changes in meanings over time. Since the location of the points are themselves linear functions of their coordinates, transformations (in this case rotations) on these coordinates represent theories of meaning and change of meaning. Decisions about which transformations (rotations) to apply therefore cannot be made on purely mathematical grounds, but must rest on an analysis of the theories of change of meaning implicit in the various transformations.

With this in mind, rotations to congruence (e.g., "Procrustean" rotations and other related rotations) that have been suggested previously (Cliff, 1962, 1966; Gibson, 1963; Green, 1952; Horst, 1956, 1962; Hurley & Cattell, 1962; Mosier, 1939; Schonemann, 1966; Schonemann & Carroll, 1970; Tucker, 1958) when applied to time-ordered sets of multidimensional scaling data, imply theories of meaning and change of meaning which may frequently be too simple to correspond to substantive notions of cognitive change now held by most psychologists. But perhaps more precisely, applications of currently available rotation algorithms may frequently yield apparent motions too complicated to find ready substantive theoretical interpretations.

Two such rotation algorithms are most well known. The first of these (what might be called a "canonical" or "Procrustes" solution) attempts to minimize the angles between corresponding pairs of coordinate axes over time. The general solution to this problem presented independently by Cliff (1966) and Schonemann (1966) identifies two transformation matrices Λ_F and Λ_A which, when applied to two multidimensional spaces F and A, yield two transformed matrices P_F and P_A. Λ_F and Λ_A are chosen such that the scalar products of P_F and P_A are at a maximum; i.e.,

$$\phi = \sum_{i=1}^{n} P_{iKA}\, P_{iKF} = \text{MAXIMUM}.$$

Since this solution maximizes the scalar product of corresponding reference axes, individual data points are weighted into the solution as a function of their projections on these axes, and therefore points will differentially resist motion over time as a function of their distance from the origin of the space. Furthermore, since the origin of multidimensional spaces is itself arbitrary, this rotation algorithm may lead to apparent point motions which are artifactually very complex.

Assume for example a psychological experiment which manipulated information such that subjects precisely reverse their definitions of concpets B and C among the six concepts given in Figure one:

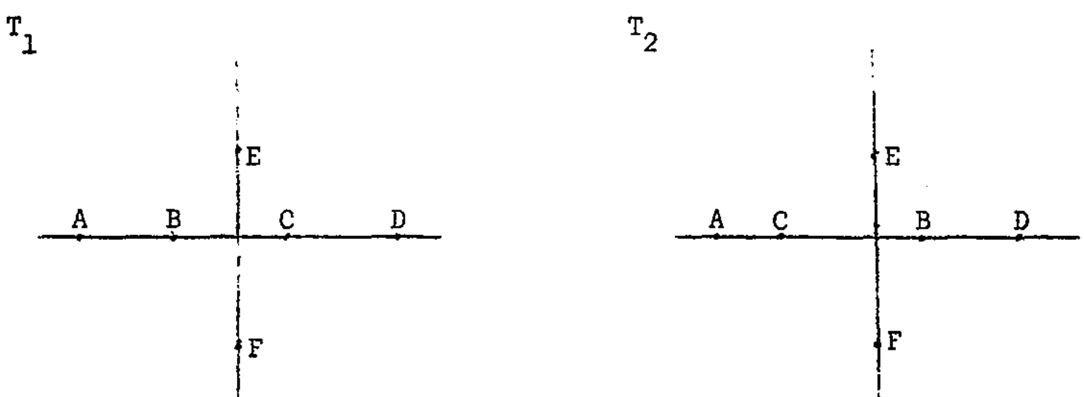

Fig. 1

If the spaces are placed one atop the other, with origins and axes corresponding, the correlation of corresponding axes will be less than unity because the order of points along the axes are different. A non-unity correlation would indicate some non-zero angle of rotation, which would in this case reduce the fit between the spaces and complicate the apparent motion.[1]

A second solution is to rotate the spaces to least squares best fit, (Woelfel, 1973; Serota, 1974) that is, to try to minimize the squared distance between matching concept-points. Such a solution uniformly attributes resistance to motion to all the data points regardless of position in the space. This method is disadvantageous because it renders highly complex the apparent change in situations where relatively simple laws could describe the "actual" change, given a more insightful rotation. For example, assume a psychological experimented which persuaded subjects to redefine only one point (X) as shown:

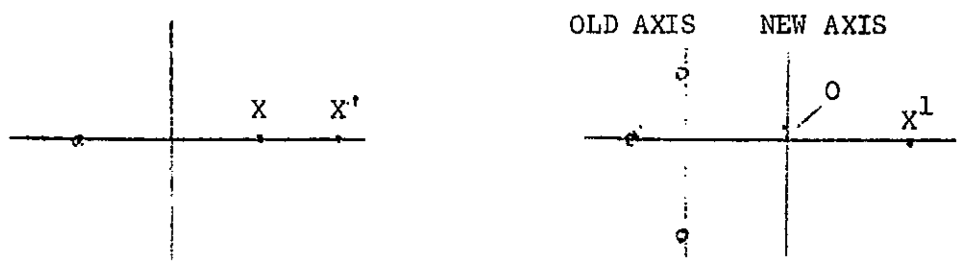

Fig. 2.

A motion of point X along the horizontal axis results in a shift in the centroid, so that, apparently, all four points moved, three of them in a direction opposite to that of X, and all of them moved <u>less</u>, seemingly, than X in fact did. Again,

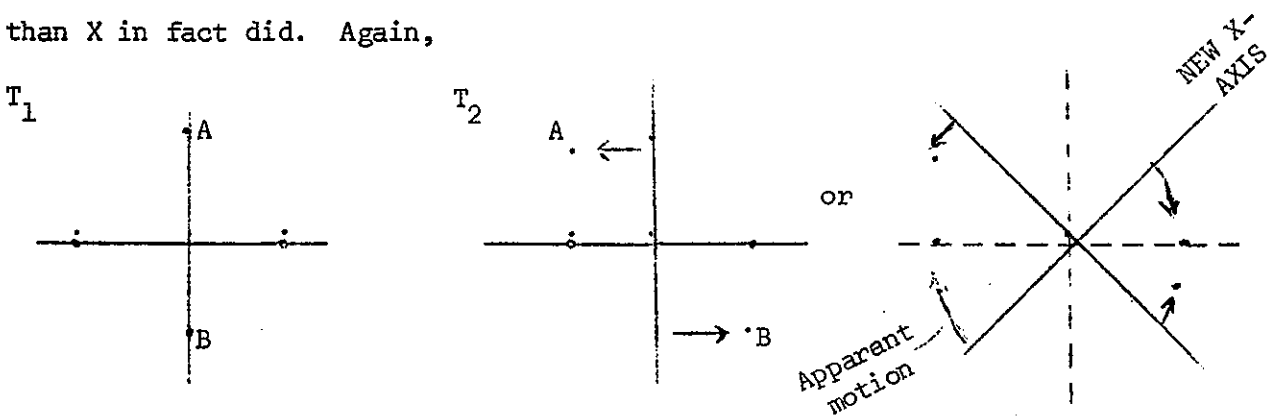

Fig. 3.

a shift in points A and B parallel to the X axis and in opposite directions will cause the least-squares axes to seem to have rotated, so that all four points will seem to have undergone rotational motion. In general, analyses of these apparent motions would not lead an experimenter to determine the simplest explanation for the change in configuration.

None of these objections are meant to detract in any way from the insight or mathematical ingenuity of these solutions, since both represent a kind of optimal solution given no theoretical information about the concepts scaled on the part of the researcher beforehand. When such information is available, however, frequently a more insightful transformation may be found. Had the researcher known in these cases, for example, that a subset of the data had been manipulated and the rest controlled, a rotation yielding far simpler apparent motions could be found[2] (McPhee, 1974).

In both the examples cited, the "extra" information required to fit the simplest apparent motions to the data may be seen to be information about the points that is independent of their coordinate values--information about how they were treated by the experimenter. Since this information is independent of the coordinate values, its value can be seen to be invariant under rotation and translation of the coordinates. The specification of such invariants under transformation is generally considered the primary task of scientific theory (Einstein, 1961; Reichenbach, 1958; Kramer, 1970; Pieszko, 1970. In this case, the invariants constitute forces applied to the concepts by the experimenter, even though those forces are expressed only as zero force or unit force in this example. An example of a theory which supplies more information is presented by Saltiel & Woelfel (1975). They postulate that the stability of the meanings of concepts is directly proportional to the amount of information about those concepts to which the individual has been exposed. Such a notion is analogous to assigning an inertial mass, represented by a scalar invariant, to each concept. Many other theories which attach a scalar invariant to the concepts can be found, including those theories which attribute differential stability to concepts on the grounds of affect, familiarity, social desirability, etc. More complex theories, which ascribe differential stability to concepts as a function of several variables, can of course also be written. We present here a solution including scalar invariants. Two cases are considered: first, where only dichotomous information is available about each concept or point, (such as whether the concept was treated or not in a psychological experiment), and second, a more general case where multivalued information describes each point (as would be implied by the "inertial mass" theories discussed earlier).

<u>The dichotomous case:</u>

The two-valued case is of interest when, for some reason, some subset of the concepts scaled is assumed not to have moved at all, as might be the case in a psychological experiment where some concepts are treated and others controlled (Saltiel, 1975).

Given m stable concepts out of n points in an r-dimensional space (m>r), the rotation procedure suggested here consists of two primary operations: (1) the establishment of a common reference system for the t_n and the t_{n+1} spaces, and (2) the rotation of the t_{n+1} space to the t_n space so that the difference between any one of the j <u>stable</u> concepts from itself is a minimum.

Establishing a common reference system consists of a straightforward translation of coordinate axes such that the origin for both spaces is at the centroid of the hypothesized stable concepts. It is important to emphasize that these concepts are hypothesized to be stable <u>relative to each other</u>; thus relative to this stable reference constellation of concepts, changes in the locations of the more volitile concepts may be calibrated.

Given the following coordinate matrices:

X = the matrix of coordinates at t_n

Y = the matrix of coordinates at t_{n+1}

The first task is to find:

A = the matrix X on the common reference system

B = the matrix Y on the common reference system

Finding the centroid of a space which is to be used as the origin for a common reference system is accomplished by first determining the average of the coordinate loadings of the j stable concepts for each of the k coordinates; that is:

$$c_k = \sum_{j=1}^{m} x_{jk}/m \qquad (1.00)$$

$$h_k = \sum_{j=1}^{m} y_{jk}/m \qquad (1.01)$$

Where, m = the number of stable concepts

x_{jk} = the projection of the <u>jth</u> stable concept on the <u>kth</u> coordinate.

The translation of the coordinate matrices from the old origin to the new "stable-centroid" origin is given by:

$$a_{ik} = x_{ik} - c_k \qquad (1.02)$$

$$b_{ik} = y_{ik} - h_k \qquad (1.03)$$

Where i refers to all of the points in the matrix. In matrix terms:

$$A = X - C \qquad (1.04)$$

$$B = Y - H$$

With both A and B coordinate matrices now located at a common reference system, the next task is to rotate the B coordinates so that the distance for any stable concept j from itself is minimized; this amounts to minimizing the following function:

$$S_{jj} = \sum_{j=1}^{m} \sum_{k=1}^{m} (a_{jk} - b^{o}_{jk})^2 = \text{MIN} \qquad (2.00)$$

Where b^{o}_{jk} are the stable projections in the rotated B matrix, denoted B^{o}.

In order to find the elements of B^{o}, b^{o}_{ik}, an orthogonal transformation matrix T is used such that:

$$B^{o} = BT \qquad (2.01)$$

The transformation matrix T may be represented as the product of a set of T of $(r^2-r)/2$ transformations T_{pq}, which perform rotations in a pq plane; i.e.,

$$T = (T_{12}, T_{13}, T_{14}, \ldots, T_{1r}, T_{23}, T_{24}, \ldots, T_{r-1(r)}) \qquad (2.02)$$

Applying the transformations T_{pq} to the $(r)(r-1)/2$ pairs of coordinates contained in matrix B (with coordinates p = 1, 2, 3, . . . r-1 and coordinates q = 2, 3, 4, . . ., r) yields:

$$(1'2', 1''3', 1'''4', \ldots, 1^{o}r', 2''3'', 2'''4'', \ldots r\text{-}1^{o}r^{o}) \qquad (2.03)$$

And, the coordinates of B^{o} are the following:

$$B^{o} = (1^{o}, 2^{o}, 3^{o}, \ldots, r\text{-}1^{o}, r^{o}) \qquad (2.04)$$

T_{pq} is defined as the two-space orthongonal tranformations commonly used in classical mechanics, that is:

$$T_{pq} = \begin{bmatrix} \cos\Theta_{pq} & -\sin\Theta_{pq} \\ \sin\Theta_{pq} & \cos\Theta_{pq} \end{bmatrix} \qquad (2.05)$$

Where, Θ_{pq} = the angles needed to minimize the distance of jth stable conceptions in matrix A from those in matrix B.

The angles of rotation Θ_{pq} are determined by first noting that the projections of the stable concepts j on the p and q coordinates in the matrix B^o are given by:[3]

$$b^o_{jp} = b_{jp}\cos\Theta_{pq} + b_{jq}\sin\Theta_{pq} \qquad (2.06)$$

$$b^o_{jq} - -b_{jp}\sin\Theta_{pq} + b_{jq}\cos\Theta_{pq} \qquad (2.07)$$

The angle Θ_{pq} which minimizes the j concepts at two points in time in a pq plane is determined by:

$$S_{\Theta_{pq}} = \sum_{j=1}^{m}(a_{jp} - b^o_{jp})^2 + \sum_{j=1}^{m}(a_{jq} - b^o_{jq})^2 = \text{MIN} \qquad (2.08)$$

By substituting (2.06) and (2.07) in (2.08) and expanding yields:

$$S_{\Theta_{pq}} = \sum_j a^2_{jp} + \sum_j a^2_{jq} + \sum_j b^2_{jp} + \sum_j b^2_{jq} - 2\cos\Theta\,(\sum_j a_{jp}b_{jp} + \sum_j a_{jp}b_{jp}) - 2\sin\Theta\,(\sum_j a_{jp}b_{jq} - \sum_j a_{jq}b_{jp}) \qquad (2.09)$$

Taking the first derivative of $S_{\Theta_{pq}}$ with respect to the angle Θ_{pq} and setting it to zero gives:

$$\frac{dS_{\Theta_{pq}}}{d\Theta_{pq}} = \sin\Theta_{pq}(\sum_j a_{jp}b_{jp} + \sum_j a_{jq}b_{jq}) - \cos\Theta_{pq}(\sum_j a_{jp}b_{jp} + \sum_j a_{jq}b_{jp}) = 0 \qquad (2.10)$$

which readily leads to the following solutions for the angle Θ_{pq}:

$$\text{Tan}\Theta_{pq} = \frac{\sin\Theta_{pq}}{\cos\Theta_{pq}} = \frac{\sum_j a_{jp}b_{jq} - \sum_j a_{jq}b_{jp}}{\sum_j a_{jp}b_{jp} + \sum_j a_{jq}b_{jq}} \qquad (2.11)$$

In the above expression (2.10), the terms involving the angle Θ_{pq} are $\sin \Theta_{pq}$ and $\cos \Theta_{pq}$; since each of these has period 2π, so does Θ_{pq}. Thus, a solution has to be considered for Θ_{pq} between 0° and 360°; although two possible solutions for Θ_{pq} will yield an extremum, that angle which produces a minimum is the desired angle for rotation.

The Continuous Case:

More complete theories frequently provide or specify measurement operations which yield numerical information about concepts which is independent of their coordinates. Such scalar invariants are suggested by the "amount of information" or "inertial mass" theories discussed earlier. This can be accomplished by associating a scalar M_j with each concept independent of its coordinate values which controls the extent to which it should be weighted in the stable least-squares solution (Danes, 1975). Using the "inertial mass" of a concept as an indication of that concept's stability, expression (2.08) may be transformed into the following minimization procedure:

$$S_{\hat{\Theta}_{pq}} = \sum_{j=1}^{n} M_j (a_{jp} - b^{o}_{jp})^2 + \sum_{j=1}^{h} M_j (a_{jq} - b^{o}_{jq})^2 = \text{Min}, \qquad (2.12)$$

Which straight forwardly leads to:

$$\operatorname{Tan} \hat{\Theta}_{pq} = \frac{\sum_j M_j a_{jp} b_{jq} - \sum_j M_j a_{jq} b_{jp}}{\sum_j M_j a_{jp} b_{jp} + \sum_j M_j a_{jq} b_{jq}} \qquad (2.13)$$

This weighted solution implies, of course, that the initial translation of axes is also weighted; this may be conviently accomplished by computing an average which is weighted by the concepts "mass," rather than the dichotomous average given in expressions (1.00) and (1.01). A weighted value for translation is given by:

$$\hat{\bar{c}}_k = \frac{\sum_{j=1}^{n} M_j x_{jk}}{\sum_{j=1}^{n} M_j} \qquad (1.06)$$

Where, M_j = a continuously scaled measure of the "inertial mass" of the jth concept

n = the total number of concepts scaled.

An Example:

While data of sufficent strength to estimate the scalar invariant for the continuous case are not available at this time, data to which the dichotomous case applies are fairly common. In this case, data were provided by a communication experiment in which subjects estimated complete paired dissimilarities comparison and for 15 concepts at four points in time by a ratio scaling technique described by Woelfel (1974). These data were averaged across subjects within time periods; two of the concepts scaled were persons (sources of information; in this case, Timothy Leary and Linus Pauling). Between times one and four, respondents received forged messages allegedly from these person advocating increasing performance of an action (the CTP) whose meaning was left deliberately vague. This action (CTP) was also among the concepts scaled. Consequently, these three concepts were manipulated in a way which should result in their being seen as increasingly similar by the subjects across the four measurements, while the other concepts not manipulated should not change (Woelfel & Saltiel, 1974).

The results of an ordinary least-squares rotation are shown in Table 1 and Figure 2. It is very unlikely that an experimenter examining this solution would detect the pattern of motion hypothesized.

In Table 2 and Figure 3, however, we present the results of a least-squares rotation excluding Leary, Pauling and the CTP from the minimization. As would be expected, these three concepts now exhibit the greatest displacements. This, however, would be expected artifactually, and so the more important evidence is provided by Figure 4. This figure shows a clear pattern of convergence among the points representing Leary, Pauling and the CTP, precisely as predicted. In this reference system, motions occur in interpretable patterns which may be related to some psychological or sociological theory perhaps more easily than in those provided by a less constrained set of transformations.[4]

FOOTNOTES

1. A canonical solution applied to patterns like this has been found to yield apparent motion involving <u>all</u> concepts in the space.

2. We might point out that the application of either of these techniques to time-ordered observations of planetary locations would similarly fail to yield the Keplerian laws of planetary motion. Contrariwise, a rotation technique which took into account properties of the planets as described in cla-sical mechanics would yield results consistent with smooth, regular planetary motion.

3. Henceforward, B^o and b^o do <u>not</u> indicate values for loadings after <u>all</u> rotations are carried out but only after the rotation in the pq <u>plane</u> is carried out.

4. Calculations for these analyses were performed using GALILEO 3.0, a metric multidimensional scaling program which provides the operations described in this paper for up to 40 x 40 concepts across any number of time periods. which is available from the authors in versions for IBM and CDC Fortran.

References

Cliff, N. Analytic rotation to a functional relationship. Psychometrika, 1962, 27, 283-295.

Cliff, N. Orthogonal rotation to congruence. Psychometrika, 1966, 31, 33-42.

Danes, J. Multidimensional scaling models for communication: Methods and applications. Unpublished Preliminary paper, Department of Communication, Michigan State University, July, 1975.

Einstein, A. Relativity: The Special and General Theory. New York: Crown, 1961.

Gibson, W. On the least-squares orthogonilization of an oblique transformation. Psychometrika, 1963, 27, 193-195.

Green, B. The orthogonal approximation of an oblique simple structure in factor analysis. Psychometrika, 1952, 17, 429-440.

Horst, P. A simple method of rotating a centroid factor matrix to a simple structure. Journal of Experimental Education, 1956, 24, 251-258.

Horst, P. Relations among m sets of measures. Psychometrika, 1962, 26, 129-149.

Hurley, J., & Cattell, R. The Procrustes Program: producing direct rotation to test a hypothesized factor structure. Behavioral Science, 1962, 7, 258-262.

Kramer, E. The Nature and Growth of Modern Mathematics (Vol. 1) Connecticut: Fawcett, 1970.

Kruskal, J. Nonmetric multidimensional scaling: A numerical method. Psychometrika, 1964b, 29, 115-129.

Kruskal, J. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika, 1964a, 29, 1-27.

McPhee, R. MDS and the search for a "proper" rotation. Unpublished manuscript, Department of Communication, Michigan State University, 1974.

Mosier, C. Determining simple structure when loadings for certain tests are unknown. Psychometrika, 1939, 4, 149-162.

Pieszko, H. Multidimensional scaling in Riemann space. Unpublished doctoral dissertation, Department of Psychology, University of Illinois, 1970.

Reichenbach, H. Space and Time. New York: Dover Publications, 1958.

Saltiel, J. A procedure for rotation in MD space. Unpublished manuscript, Department of Sociology, University of Montana, 1975.

Saltiel, J., & Woelfel, J. Accumulated information as a basis for attitude stability. Human Communication Research, 1975, in press.

Schonemann, P. A generalized solution of orthogonal Procrustes problem. Psychometrika, 1966, 31, 1-10.

Schonemann, P., & Carroll, R. Fitting one matrix to another under choice of a central dilation and a rigid motion. Psychometrika, 1970, 35, 245-255.

Serota, K. Metric multidimensional scaling and communication: Theory and implementation. Unpublished Master's thesis, Department of Communication, Michigan State University, 1974.

Torgerson, W. Theory and Methods of Scaling. New York: Wiley, 1958.

Tucker, L. An inter-battery method of factor analysis. Psychometrika, 1958, 23, 111-136.

Woelfel, J. Metric measurement of cultural processes. A paper presented to the Annual Meeting of the Speech Communication Association, Chicago, December, 1974.

Woelfel, J., & Saltiel, J. Cognitive processes as motions in a multidimensional space: A general linear model. Unpublished manuscript, Department of Communication, 1974.

TABLE 1

Distances Moved Between Time Intervals in Ordinary Least Squares Rotation, No Stable Concepts

Concept	$t_1 - t_2$	$t_2 - t_3$	$t_3 - t_4$
01 sleeping	29.096	23.817	20.100
02 dreaming	13.431	19.112	18.200
03 day dreaming	23.598	13.939	14.396
04 intense concentration	22.977	27.453	15.748
05 marijuana high	28.687	21.326	17.798
06 good	21.160	33.005	21.274
07 depression	18.591	16.979	17.851
08 alcohol high	22.245	17.851	18.903
09 relaxation	30.040	23.211	29.844
10 CTP	17.854	26.237	14.422
11 alpha wave meditation	23.303	21.125	11.371
12 transcendental mediation	16.952	14.901	10.837
13 reliable	21.752	25.938	14.330
14 Timothy Leary	22.214	25.669	19.682
15 Linus Pauling	33.463	24.067	24.209
16 me	27.979	19.377	14.375

TABLE 2

Distances Moved Between Time Intervals
for Table Concepts Rotation

Concept	$t_1 - t_2$	$t_2 - t_3$	$t_3 - t_4$
01 sleeping	20.923	15.752	19.067
02 dreaming	14.539	18.902	19.285
03 day dreaming	22.033	12.331	17.785
04 intense concentration	20.110	17.462	14.972
05 marijuana high	21.858	24.396	16.136
06 good	30.317	23.736	13.200
07 depression	20.354	16.093	12.331
08 alcohol high	25.877	11.497	14.937
09 relaxation	31.861	16.341	26.256
10 CTP	56.868*	67.708*	36.790*
11 alpha wave meditation	18.728	25.431	12.237
12 transcendental mediation	19.175	15.446	16.713
13 reliable	22.446	21.220	27.047
14 Timothy Leary	27.333	71.548*	37.551*
15 Linus Pauling	124.533*	74.952*	43.308*
16 me	24.766	25.259	23.889

* Indicates concepts not specified as stable.

FIGURE 4

Ordinary Least Squares Rotation

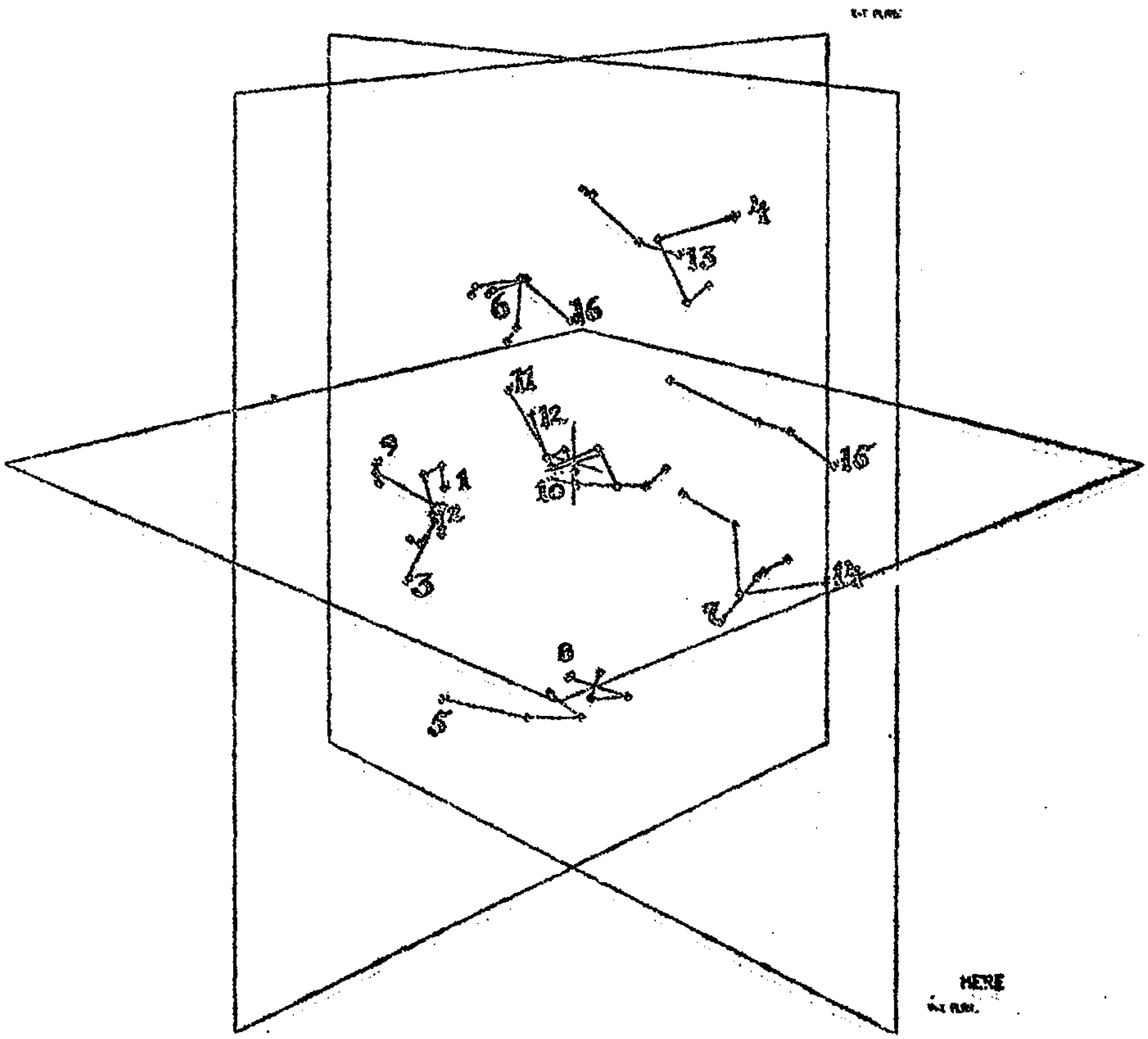

FIGURE 5

Stable Concepts Rotation