DOCUMENT RESUME

ED 295 979 TM 011 800

AUTHOR Jones, Patricia E.
TITLE Assessment of Dimensionality in Dichotomously-Scored Data Using Multidimensional Scaling: Analysis of HSMB Data.
PUB DATE Apr 88
NOTE 33p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Discriminant Analysis; *Factor Analysis; Language Tests; Mathematics Tests; *Multidimensional Scaling; Preschool Children; Preschool Education; *Preschool Tests; Social Development; Statistical Analysis
IDENTIFIERS *Head Start Measures Battery

ABSTRACT

Data from three subtests (language, mathematics, and social) scales of the Head Start Measures Battery (HSMB) were analyzed using principal components analysis (PCA) and non-metric multidimensional scaling (MDS). The HSMB measures preschool development in language, mathematics, nature and science, perception, reading, and social development. The sample included 1,000 children (36-60 months old) who participated in the Head Start Program and received all language, mathematics, and social routing items in fall of 1985. Loadings on the first factor were high, and eigenvalues obtained from the PCA suggested that one (or possibly two) dimensions were present in the data. Loadings of rotated principal components suggested that at least three factors corresponding to the subtests were present. The structure was clarified by MDS plots, showing that items were located in distinct sectors corresponding to their subtest. However, items having the highest IRT discrimination parameters were clustered toward the center, suggesting that the measures have a strong common factor and unique variance related to each subtest. Practically, this justifies both a common scale for all subtests when an overall measure of achievement is needed and individual subtest scaling when information on particular skills is required. Agreement between PCA and MDS can be used to reinforce the validity of the principal components model. Seven tables and four figures are provided. (SLD)

ASSESSMENT OF DIMENSIONALITY IN DICHOTOMOUSLY-SCORED DATA USING MULTIDIMENSIONAL SCALING: ANALYSIS OF HSMB DATA

Patricia B. Jones

Department of Instructional Computing
Center for Computing and Information Technology
The University of Arizona

Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

April, 1988

Abstract

Data from three subtests of the Head Start Measures Battery were analyzed using principal components analysis and nonmetric multidimensional scaling (MDS). Loadings on the first factor were high and eigenvalues obtained from the principal components analysis suggested that one, or possibly two, dimensions were present in the data. On the other hand, loadings of rotated principal components suggested that at least three factors corresponding to the subtests were present. The structure was clarified by MDS plots which showed that items were located in distinct sectors corresponding to their subtest. However, items having the highest IRT discrimination parameters were clustered toward the center. This suggests that the measures have a strong common factor as well as unique variance associated with each subtest. Practically, this justifies both a common scale for all subtests when an overall measure of achievement is needed and individual subtest scaling when information on particular skills is required. Agreement between principal components analysis and MDS can be used to reinforce the validity of the principal components model.

Introduction

Although many methods have been developed for determining the underlying latent structure of a data set, no single procedure has emerged which is satisfactory in all cases. Hattie (1985) has presented an extensive review of procedures used to examine the unidimensionality assumptions of latent trait theory; many of these methodologies are unsatisfactory for analysis in general situations because they rely on restrictive assumptions (e.g., additivity of latent components, distributional assumptions) which may not be appropriate. The methodologies discussed by Hattie are used to test for unidimensionality; they do not determine the nature and extent of multidimensionality present in the data.

While factor analysis and principal components analysis represent the most commonly used techniques for examining the latent structure of a data set, problems may be encountered in their use. First, factor analytical techniques assume that the latent components are additive. If the model is not appropriate, conclusions regarding the latent structure of the data set may be in error. Second, when dichotomized data are factor analyzed, factors related to the marginal distributions of the items (the item difficulties) may be present. Such statistical artifacts can result from the type of correlation coefficient used, or the type of data analyzed, but do not have any theoretical meaning. Spurious influence of item difficulty has been noted primarily when the phi correlation coefficient (Lord and Novick, 1968) is used to form the correlation matrix (Ferguson, 1941; McDonald and Ahlawat, 1974). Since the range of the phi coefficient may be restricted because of the differences in marginal distributions of the item pairs, substitution of the tetrachoric correlation (Lord and Novick, 1968) has been recommended. However, use of the tetrachoric correlation coefficient requires that item pairs share a bivariate normal distribution and that the probability of a chance correct response be zero. Moreover, the tetrachoric correlation is difficult to estimate and matrices of tetrachoric correlations may be non-Gramian (contain one or more negative eigenvalues). Finally, Guilford (1941) and Jones, Sabers, and Trosset (1987) have noted the presence of difficulty-related factors even when tetrachoric correlations were analyzed.

Multidimensional scaling (MDS) is a technique which can be used to assess dimensionality without recourse to the restrictive assumptions of the techniques of factor analysis. MDS techniques are based on a distance model. The algorithm attempts to locate objects in a representational space of specified dimensionality such that interpoint distances (measures of similarity or dissimilarity which are input to the procedure) are reproduced as faithfully as possible. The more closely the number of dimensions in the representational space matches the number of dimensions in the space which contains the items, the more closely the approximated distances will match the original distances. In nonmetric MDS, the scaling is carried out such that only the rank orders of the interpoint distances are preserved. The fit of the scaling is measured by a statistic such as STRESS (Kruskal, 1964).

MDS avoids the additivity requirements of the factor analytical model. Furthermore, input to the procedure may be any type of distance measure. Unlike principal components or factor analysis, correlation coefficients are not required. The interpretation of the axes in the MDS representational space may differ from that for principal components analysis where a vector of responses is plotted for each case with axes corresponding to items. The principal components form a subspace with axes which correspond to linear combinations of the items and which are often interpreted as latent variables. In MDS, the items are data points which lie in a space of unknown dimensionality. A parsimonious representation of that space is sought. The specified axes need not represent latent variables and even if they do, they need not represent the same latent variables obtained in the principal components analysis. Since MDS is based on a distance model, any orientation of the axes in the representational space preserves the distances, and their origin is arbitrary.

MDS has been used to examine the structure of test data sets (Koch, 1983; Korpi and Haertel, 1984; Thomas, 1984, 1985; Allen, 1987). It is especially useful for the case of dichotomous data for which the models underlying classical or confirmatory factor analytical techniques may be inappropriate. However, few studies have explored the behavior of MDS with simulated dichotomized data of known structure. Reckase (1981) examined its effectiveness with one-, two-, three-, and nine-factor data sets for 14 similarity coefficients with and without guessing present. He used the procedure to isolate sets of homogeneous items in a two-dimensional representational space and found it useful with simulated data. However, he found that MDS results did not give the expected results for data taken from the Iowa Tests of Educational Development. The reason for the discrepancy between obtained and expected results was not clear.

Interpreting MDS Plots

The effectiveness of MDS in recovering the underlying dimensionality of one-, two-, and three-factor data sets was explored in a study prior to the current work (Jones et al., 1987). The study examined the effect of using margin-sensitive (agreement[1], phi, kappa) and margin-free (phi/phimax, Yule's Q, and the tetrachoric correlation) similarity coefficients [a discussion of these coefficients can be found in Reckase (1981) and Jones et al. (1987)]. Whereas Reckase (1981) examined the effect of guessing on MDS results, Jones et al. varied item discrimination (the correlation of the item with the latent trait) and examined the behavior of the STRESS coefficient for MDS solutions in different specified dimensionalities. Four especially important findings emerged from their study:

1. MDS plots are variations on essentially the same plot for all the similarity coefficients. An extra dimension is required in the representational space to accommodate the effect of item difficulty on margin-sensitive coefficients.
2. Orthogonal factors can be thought of as having a "center of gravity" in space located equally distant from the center of gravity for every other factor. Items related to the same factor tend to cluster about the same center of gravity. To allow for the equidistance criterion, _k_ dimensions are required in the representational space for _k+1_ factors.
3. As the correlation of a unidimensional item with its

[1] The agreement coefficient (Reckase, 1981) is the proportion of people who give identical responses on paired items. For a 2 x 2 matrix, it is the proportion of people in the (0,0) and (1,1) cells.

latent trait decreases, the distance of the item from its center of gravity increases.

4. Items related to multiple factors are located in space between the factors to which they are related.

Some amplification of the above conclusions may clarify the results of the present study.

First, items tend to group in clusters according to the factors to which they are related. If an item is related to a single factor, it is grouped with other items related to the same factor. Items with the highest item-latent trait correlations are grouped at the center of the cluster. Items are arranged around the cluster center in concentric circles (or spheres) in order of decreasing correlation such that items with the lowest correlations with the latent trait appear farthest from the cluster center. Figure 1 (Jones et al., 1987) presents results obtained from scaling a unidimensional set of 40 simulated items with varying item-latent trait correlations. The matrix input to the scaling procedure was a matrix of Yule's Q coefficients (Bishop, Fienberg, and Holland, 1975). Items numbered 1 correlated .9 with the single latent trait and appear at the center of the graph. Those correlating .6 with the latent trait are indicated with the number 2 and are scaled around the first group of items. Items which correlated .3 with the latent trait are indicated with the number 3 and appear on the outside of the circle.

If multiple factors are present in the simulated data (i.e., if some items are related to factor A, some items to factor B, etc.), the cluster centers appear at equal distances from each other. When three factors exist in the data, for instance, groups of unidimensional items are positioned at the vertices of an equilateral triangle. This configuration is necessary because items related to one of the factors are no more related to a second factor than they are to a third. For this reason, the items should be located equally distant from the cluster centers for the other two factors. In order to permit this configuration, the representational space for MDS requires two dimensions (assuming that margin-free coefficients are used for the scaling). Figure 2 (Jones et al., 1987) presents MDS results obtained for scaling a three-factor data set of 40 items using a matrix of Yule's Q coefficients. All items correlate .9 with one of the three factors. Items with (.9,0,0) factor loadings (that is, a loading of .9 on the first factor and zero loadings on the second and third factors) are indicated by the numeral 1 on the plot. The numerals 2 and 3 refer to items with (0,.9,0) and (0,0,.9) factor loadings. As the number of factors in the data increase, the number of dimensions required in the MDS representational space increase. A scaling of a simulated four-dimensional data set revealed that a pyramid (and therefore a representational space of three dimensions) is required to represent adequately the interpoint distances given by the similarity matrix. The dimensions of the MDS representational space could not be interpreted as factors. The factors corresponded to locations within the representational space.

Figure 3 (also from Jones et al., 1987), obtained for a three-factor data set, presents results illustrating the fourth conclusion described above. Items labeled 1, 2, and 3 have factor loadings of (.9,0,0), (0,.9,0), and (0,0,.9) respectively. Items labeled 4, 5, and 6 have factor loadings of (.7,.5,0), (0,.7,.5), and (.5,0,.7). Items labeled 7 have factor loadings of (.5,.5,.5). The first three sets of items form the vertices of an equilateral triangle. Since items numbered 4 are related only to the first two factors, they are placed between the item clusters corresponding to the first two factors, but somewhat more closely to the first-factor items, corresponding to their higher loading on the first factor. Similarly, item groups 5 and 6 are located between clusters related to the factors (2,3) and (3,1), respectively. Finally, the items related to all three factors are clustered in the center of the plot, the only location equally distant from all three factors.

Extension of Study

The purpose of the present study is to extend the results of Jones et al. (1987) to actual test data and to show how MDS may be effectively combined with traditional scaling and analysis techniques to inform the researcher about the structure underlying the data.

Measures

Items from three of the six tests comprising the Head Start Measures Battery or HSMB (Bergan, 1986) were analyzed using both principal components analysis and MDS to determine the appropriate structure underlying the data. The HSMB is a test designed to measure pre-school development in six areas: language, mathematics, nature and science, perception, reading, and social development. It is individually administered to children between three and five years of age who participate in the Head Start Program. The test consists of free-response items. Each subtest consists of three parts: a routing test administered to every child, a level I test composed of easier items administered to children who correctly answer fewer than a preset number of items on the routing test, and a level II test composed of more difficult items administered to children whose correct responses on the routing test meet or exceed the criterion number of items. The routing tests consist of items which cover a wide range of difficulty and which generally display moderate to high item discrimination parameters. The Language, Mathematics, and Social scales were chosen for this analysis since they represent three distinctly different content areas. The routing tests were chosen because all children tested on a given scale receive the routing items; therefore, no selection for ability occurs on these items. The Language, Mathematics, and Social routing tests contain, respectively, 9, 6, and 8 items. Because the items are individually administered to very young children, guessing is not generally considered to be important and personal experience with scaling the tests has shown that a two-parameter logistic model fits most items well. Parameter estimates obtained using the Multilog scaling program (Thissen, 1984) for items of the three routing tests are presented in Table 1. Examination of the characteristics of the scales using confirmatory factor analysis prior to this investigation had suggested a separate factor for each subtest (Bergan, personal communication, June 1986). However, a matrix of intercorrelations of subtest scores for all six scales had revealed correlations among the subscales ranging from .50 to .65. This suggested the presence of a moderate to strong single factor in the measures. Particular interest in the structure of the test was prompted by a request from the national Head Start Program who wished to refer low-scoring children for additional evaluation. The fundamental problem concerned how to determine what constituted a low score on the measures. After consideration of the problem, the most appealing solution was to rescale items from several subtests on a single scale using a two-parameter model, to establish norms for ability levels for six-month age groupings, and to determine a cutting score for an appropriate percentile for each age group represented. However, if the items were not unidimensional, placing them on a single scale wo[illegible]t be appropriate and the difficulty of determining a crit[illegible]eferral would be substantially increased.

One issue related to the concept of unidimensionality should be mentioned here. Any data set which contains random error cannot, by definition, be truly unidimensional, since a dimension is required to describe fully the random error component. However, in a reliable data set, the random error component is small relative to the systematic component, so that it may be neglected without causing substantial problems. Similarly, some subsets of items may contain common effects which are small relative to an overall systematic component. Again, such components may possibly be ignored in favor of a unidimensional scaling. The problem, therefore, may not really be whether the item set is unidimensional, but whether it is sufficiently unidimensional to permit accurate scaling estimates of item parameters (see Hulin, Drasgow, and Parsons, 1983, pp. 105-108).

Subjects

The sample consisted of children who participated in the Head Start Program and who received all Language, Mathematics and Social routing items in Fall, 1985. The study was restricted to include only children between the ages of 36 and 60 months, the age range for which the scales are maximally effective. One thousand subjects were randomly sampled from more than 17,000 children who received some or all of the HSMB scales.

Data Analysis

Matrices of phi, tetrachoric correlation, agreement and Yule's Q coefficients were constructed from the raw response data. Principal components analyses were conducted on the matrices of phi and tetrachoric correlations using the FACTOR module of the SYSTAT statistical package (Wilkinson, 1984) for the IBM PC. Despite its name, FACTOR produces only principal components analyses. Scree plots of eigenvalues were examined for abrupt changes in slope. The first five principal components were then rotated orthogonally. Unrotated and rotated loadings were compared for similarity to the assumed factor structure.

Matrices of agreement, Yule's Q, and tetrachoric coefficients were scaled using the SYSTAT MDS module. The results of Jones et al. had previously suggested that these three coefficients represented appropriate similarity measures for MDS. The agreement coefficient was used because it appeared to be the most appropriate coefficient for recognizing unidimensionality. Reliable unidimensional item sets have a low STRESS value for a one-dimensional representational space when the agreement coefficient is used; this is not true for margin-free coefficients. However, when data sets contain more than a single factor, the presence of an additional dimension caused by item ordering due to difficulty tends to cloud the structure of the MDS plot. Therefore, the magnitude of the agreement coefficient was examined only to determine whether data were unidimensional. Yule's Q was used because it is not affected by item difficulty and generally yields the same results as the tetrachoric correlation, while being easier to compute. The tetrachoric correlation was used as a basis for comparison with the principal components analysis.

MDS analyses were conducted for representational spaces of one through five dimensions, using the Kruskal scaling algorithm. STRESS values were plotted against the number of dimensions used in scaling and examined for abrupt changes in slope. Two-dimensional item plots were examined to determine the configuration of the points in each representational space.

Results

Principal Components Analyses

Table 2 presents eigenvalues corresponding to the first five principal components for analyses of phi and tetrachoric correlations. Both sets of eigenvalues suggest a strong first factor and a weaker second factor. Table 3 shows the unrotated principal components loadings for the first five factors obtained by analyzing the matrix of phi coefficients. Except for the second Mathematics item (K), all items load strongly on the first principal component. The second principal component serves primarily to differentiate the Language items from the remaining items; the third principal component serves primarily to differentiate the Mathematics items from the remaining items. However, the magnitude of the third eigenvalue suggests that the third principal component is not important. The loadings for the orthogonally rotated components are presented in Table 4.

Examination of the rotated loadings for the five factors shows that items P - T load most heavily on the first component, but that the loadings for the three remaining items from the Social scale are substantially smaller. Items A - G, all Language items, load heavily on the second component, while the two remaining Language items, H and I, have their heaviest loading on the fifth principal component. Items J - O, all Mathematics items, load most heavily on the third principal component, although item K tends to load less heavily than the other items.

Tables of unrotated and rotated principal components loadings obtained from analysis of the matrix of tetrachoric correlations appear similar to those obtained for phi correlations. The loadings are presented in Tables 5 and 6.

Thus the principal components analysis gives apparently conflicting results. Eigenvalues suggest that only two factors, one substantially more important than the other, are present in the data. However, the magnitude of the loadings on the first principal component suggest that the data might be unidimensional, while the magnitude of the loadings on the rotated orthogonal factors suggest that as many as five factors might be present in the data.

MDS Analyses

STRESS values obtained for the 15 MDS analyses are presented in Table 7. An appropriate dimensionality in MDS may be chosen by determining the number of dimensions in the representational space at which an abrupt change in slope occurs. (The technique is analogous to examination of scree plots in principal components or factor analysis, but the appropriate number of principal components is given by the number of eigenvalues prior to the break in the plot.) An MDS representational space may also be chosen when the STRESS value is sufficiently low, but the criteria for this are not always clear. Kruskal and Wish (1978) suggest that values of .10 - .15 may be acceptably low, but the value depends on the type of similarity coefficient input to the scaling procedure. The plots for the HSMB do not contain a clear break at any dimensionality, although values appear to be acceptably low for either a two- or three-dimensional representational space when items are scaled using Yule's Q or tetrachoric coefficients.

Examination of the two-dimensional MDS plot for the routing items reveals substantial information regarding the characteristics of the items. Figure 4 shows the results obtained when items were scaled using a Yule's Q similarity matrix for a two-dimensional representational space. Items tend to be grouped in non-overlapping regions corresponding to the subscale (Language, Mathematics, or Social) to which they belong. Language items (identified with the letters A-I) appear on the left-hand side of the plot; Mathematics items are located below the Social items on the right-hand side of the plot. Such a structure would be expected from the second conclusion of the Jones et al. study; that is, that items related to the same factor tend to cluster about the same center of gravity. One item (K) from the Mathematics subscale appears quite different from the rest; although it is scaled in the sector of the plot corresponding to the remaining Mathematics items, it is widely separated from those items. As suggested by the third conclusion of Jones et al., this suggests that the item is poorly correlated with the latent trait. The structure of the items, in general, suggests that the subtests consist of shared variance which is unique to the individual subtests. The lack of strong differentiation along the second dimension in the representational space also suggests that the items might be strongly related to a common factor, since items which are related to the same common factor tend to cluster about the same center.

The MDS results help to clarify the reasons for the apparent conflict in the principal components analyses. The strong common element corresponds to the high loadings on the first factor of the unrotated principal components analysis. The unique factors correspond to the single-factor loadings on the rotated principal components.

The MDS plot also helps to understand the item parameter estimates presented in Table 1. The large distance of item K from the centroid of the mathematics items suggests that its item discrimination estimate is indicative of a substantial amount of random error. This is confirmed by items H and I which have the highest discrimination parameters of the Language items; they are scaled close to the center of the diagram. Similarly, items

M, N, and O, with the highest discrimination values of the Mathematics items, and items P, Q, T, U, and V, with the highest discrimination values of the Social items are also scaled toward the interior of the plot. In general, as the discrimination decreases, items are scaled farther away from the center, but within a region corresponding to their subtest. The closeness of the cluster centers indicates that the items contain a strong common element; the fact that the diagram can be divided into regions corresponding to each subtest suggests that each subtest also contains an additional unique element.

The structure of the item clusters in MDS space suggests that an appropriate model for item response should include θ_{Tj}, the ability of the jth individual on the common latent trait T shared by all tests, θ_{kj}, the ability on the latent trait measured by the kth subscale, and ϵ_{ij}, the error (or unique variance component) on the ith item. In practical terms, the HSMB may be scaled in two different ways to meet two different needs. Separate scalings for the subtests appropriately reflect the presence of unique latent traits present in the separable cluster domains. Yet, a single scaling of all items can also be used to gauge overall performance because of the presence of a strong common factor.

The appropriateness of the two-parameter logistic model for these data can not be determined conclusively from this analysis. However, the strong agreement between the MDS results (which were not subject to the linearity restriction nor to a particular parametric model), the principal components analysis (which was subject to the linearity restriction but not to a particular parametric model) and the item response theory analysis (which was subject to both the restriction and the model) suggest that the model is indeed appropriate.

Conclusions

Each methodology of dimensionality analysis has adherents and detractors and probably no single technique is effective in and appropriate for all situations. This study shows by example that a variety of methods can be used synergistically to examine the structure of real data for practical ends. Nonmetric MDS, a rarely used technique, can be particularly useful because of its graphic display of data structures and its freedom from restrictive assumptions. On the other hand, principal components and factor analysis yield well-defined, quantitative information under restrictive linearity conditions. The two approaches can be used together to complement each other and to extract maximal information regarding the underlying structure of the data. Substantial agreement between the methodologies can be used to suggest the appropriateness of the principal components model.

Table 1. IRT Parameter Estimates for HSMB Data.

Item ID	Test	a value	b value
A	Lang	0.937	-0.247
B	Lang	1.147	1.711
C	Lang	0.872	-0.922
D	Lang	0.931	0.588
E	Lang	1.212	-0.022
F	Lang	1.336	0.374
G	Lang	1.372	0.777
H	Lang	2.209	0.160
I	Lang	2.451	0.102
J	Math	1.017	-0.508
K	Math	0.503	-1.149
L	Math	0.959	-0.326
M	Math	1.973	0.310
N	Math	2.349	0.391
O	Math	1.758	0.928
P	Soc	1.577	0.139
Q	Soc	2.260	-0.225
R	Soc	1.442	0.543
S	Soc	1.257	-1.007
T	Soc	1.928	-0.688
U	Soc	1.679	-0.488
V	Soc	2.179	-0.765
W	Soc	0.953	0.739

Table 2. First Five Eigenvalues for Principal Components Analyses of Phi and Tetrachoric Correlation Coefficients. HSMB Data.

Coefficient	Eigenvalues				
Φ	5.398	2.101	1.392	1.181	1.165
Tetrachoric	8.090	2.704	1.448	1.213	1.189

Table 3. Unrotated loadings for principal components analysis of HSMB data (phi coefficients).

	1	2	3	4	5
A	0.423	-0.457	-0.013	-0.149	0.263
B	0.354	-0.421	0.097	-0.188	0.150
C	0.468	-0.396	-0.190	-0.050	0.096
D	0.394	-0.461	-0.088	-0.208	0.340
E	0.491	-0.472	-0.201	0.144	-0.166
F	0.483	-0.421	-0.208	0.193	-0.126
G	0.458	-0.402	0.065	0.152	-0.050
H	0.485	-0.155	0.197	0.013	-0.613
I	0.519	-0.114	0.244	0.007	-0.518
J	0.456	0.253	0.325	0.454	0.064
K	0.239	0.087	0.175	0.107	0.158
L	0.470	0.172	0.264	0.463	0.097
M	0.572	0.033	0.295	-0.042	0.271
N	0.553	0.165	0.409	-0.083	0.198
O	0.504	-0.001	0.345	-0.057	0.113
P	0.456	0.149	-0.246	-0.008	-0.072
Q	0.564	0.257	-0.404	0.031	-0.041
R	0.580	0.297	-0.269	0.135	0.120
S	0.580	0.275	-0.406	0.058	-0.025
T	0.538	0.326	-0.325	0.169	0.143
U	0.490	0.281	0.039	-0.476	-0.199
V	0.535	0.334	0.063	-0.432	-0.089
W	0.377	0.265	0.029	-0.296	0.060

Table 4. Rotated loadings for principal components analysis of HSMB data (phi coefficients).

	1	2	3	4	5
A	0.034	-0.677	0.107	-0.093	-0.008
B	-0.077	-0.571	0.095	-0.143	-0.088
C	0.229	-0.593	0.016	-0.020	-0.135
D	0.057	-0.714	0.046	-0.109	0.093
E	0.262	-0.536	-0.002	0.149	-0.417
F	0.294	-0.498	0.039	0.177	-0.374
G	0.085	-0.474	0.212	0.088	-0.342
H	0.067	-0.110	0.100	-0.185	-0.783
I	0.068	-0.122	0.186	-0.223	-0.712
J	0.220	0.086	0.705	0.043	-0.185
K	0.066	-0.059	0.346	-0.063	0.020
L	0.241	-0.001	0.670	0.090	-0.170
M	0.110	-0.312	0.534	-0.309	-0.015
N	0.059	-0.170	0.583	-0.414	-0.055
O	0.019	-0.248	0.464	-0.303	-0.142
P	0.479	-0.103	0.060	-0.178	-0.142
Q	0.701	-0.092	0.064	-0.188	-0.103
R	0.653	-0.088	0.271	-0.146	-0.003
S	0.724	-0.086	0.093	-0.177	-0.095
T	0.689	-0.054	0.247	-0.099	0.044
U	0.241	-0.029	0.026	-0.694	-0.214
V	0.280	-0.036	0.133	-0.694	-0.133
W	0.221	-0.046	0.139	-0.484	0.021

Table 5. Unrotated loadings for principal components analysis of HSMB data (tetrachoric coefficients).

	1	2	3	4	5
A	0.519	-0.523	-0.012	-0.[illegible]91	0.142
B	0.503	-0.528	-0.095	-0.267	0.009
C	0.589	-0.467	0.179	-0.071	0.075
D	0.487	-0.523	0.104	-0.357	0.195
E	0.591	[illegible].514	0.190	0.265	-0.056
F	0.587	-0.453	0.211	0.300	0.019
G	0.577	-0.439	-0.095	0.192	0.038
H	0.583	-0.153	-0.215	0.286	-0.552
I	0.624	-0.102	-0.265	0.222	-0.471
J	0.546	0.306	-0.405	0.304	0.267
K	0.299	0.108	-0.235	-0.033	0.213
L	0.561	0.205	-0.326	0.316	0.313
M	0.696	0.057	-0.280	-0.209	0.183
N	0.685	0.223	-0.373	-0.190	0.087
O	0.665	0.037	-0.306	-0.100	0.025
P	0.571	0.172	0.264	0.046	-0.071
Q	0.677	0.276	0.405	0.088	-0.001
R	0.690	0.322	0.258	0.076	0.182
S	0.702	0.296	0.391	0.105	0.032
T	0.650	0.352	0.308	0.095	0.235
U	0.579	0.313	0.043	-0.289	-0.399
V	0.632	0.374	0.009	-0.311	-0.268
W	0.468	0.322	0.055	-0.312	-0.104

Table 6. Rotated loadings for principal components analysis of HSMB data (tetrachoric coefficients).

	1	2	3	4	5
A	0.027	-0.781	-0.136	-0.136	-0.034
B	-0.055	-0.734	-0.120	-0.179	-0.154
C	0.266	-0.711	-0.054	-0.036	-0.168
D	0.079	-0.808	-0.060	-0.137	0.078
E	0.330	-0.620	-0.019	0.164	-0.449
F	0.385	-0.578	-0.052	0.203	-0.398
G	0.148	-0.566	-0.260	0.103	-0.392
H	0.107	-0.185	-0.138	-0.218	-0.827
I	0.109	-0.196	-0.234	-0.266	-0.752
J	0.253	0.056	-0.781	-0.014	-0.207
K	0.063	-0.089	-0.425	-0.099	0.027
L	0.289	-0.044	-0.733	0.056	-0.188
M	0.178	-0.366	-0.582	-0.368	-0.059
N	0.158	-0.196	-0.638	-0.465	-0.121
O	0.141	-0.299	-0.517	-0.345	-0.228
P	0.556	-0.144	-0.101	-0.244	-0.179
Q	0.771	-0.129	-0.124	-0.246	-0.137
R	0.719	-0.124	-0.330	-0.208	-0.033
S	0.792	-0.126	-0.169	-0.237	-0.132
T	0.755	-0.091	-0.315	-0.163	0.025
U	0.318	-0.051	-0.077	-0.710	-0.253
V	0.360	-0.059	-0.203	-0.712	-0.164
W	0.308	-0.057	-0.164	-0.556	0.009

Table 7. MDS Stress Values for Three Similarity Coefficients. HSMB Data.

MDS STRESS Values*

Coefficient	Number of Dimensions				
	1	2	3	4	5
Agreement	28690	15251	10867	06558	04702
Yule's Q	36696	17250	11813	08263	06413
Tetrachoric	36423	16807	11524	08160	06422

* Leading decimal points omitted

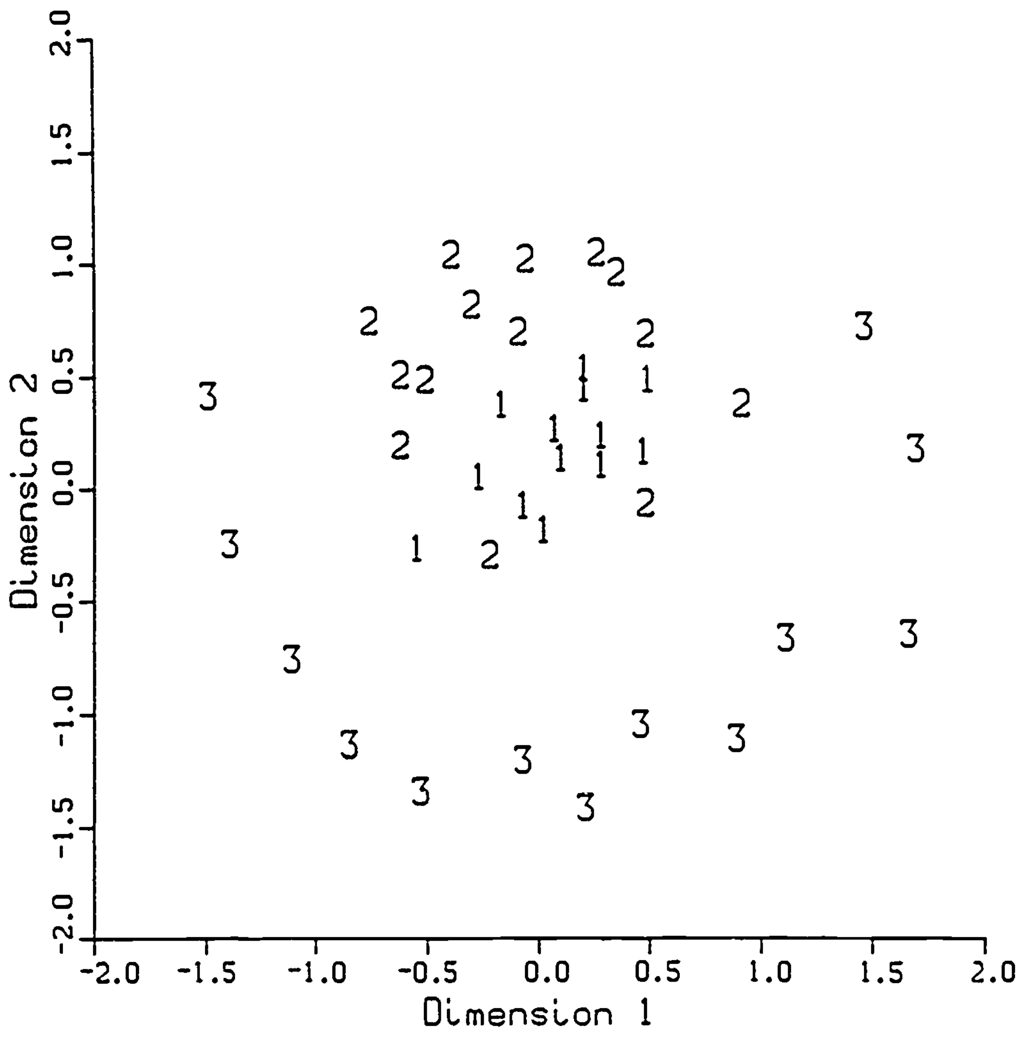

Figure 1. MDS Plot for unidimensional simulated data set, Yule's Q similarity matrix.

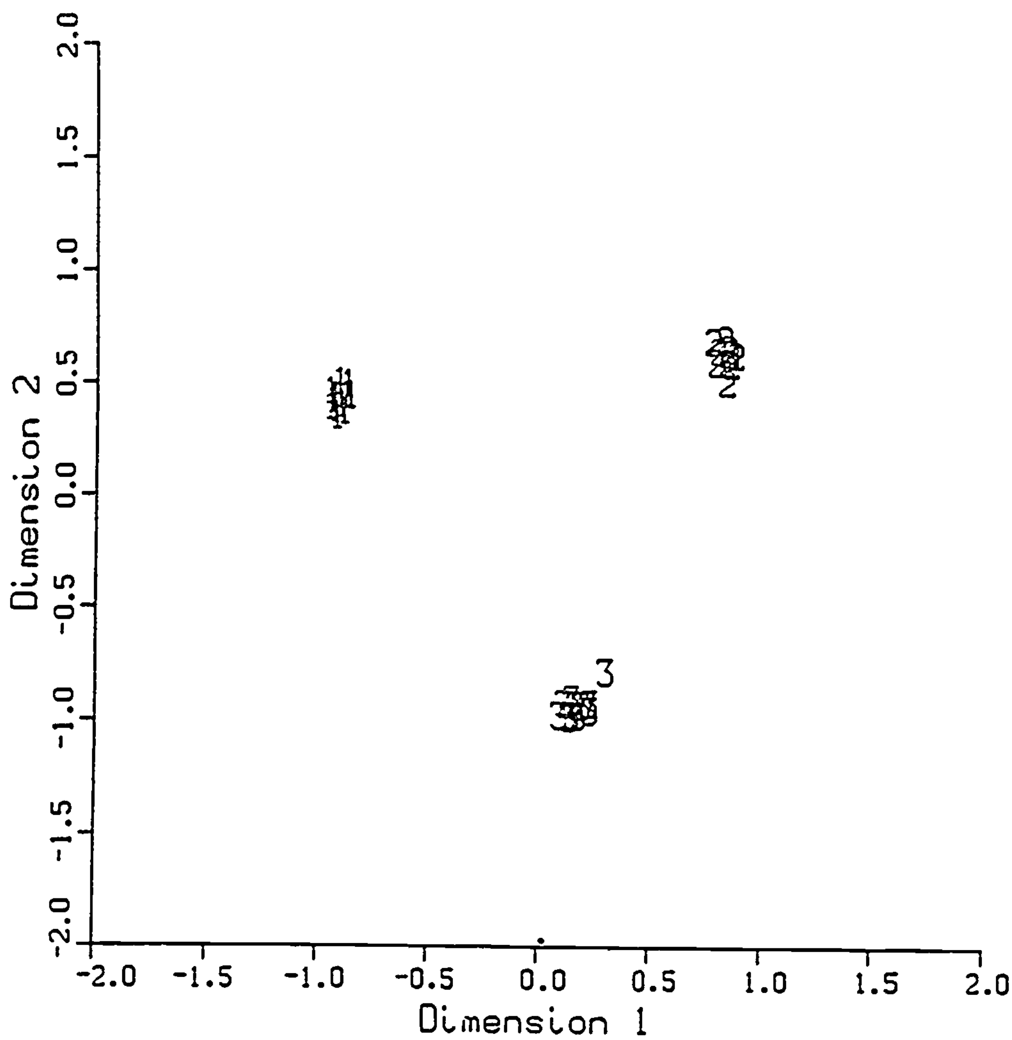

Figure 2. MDS Plot for three-dimensional simulated data set, Yule's Q similarity matrix.

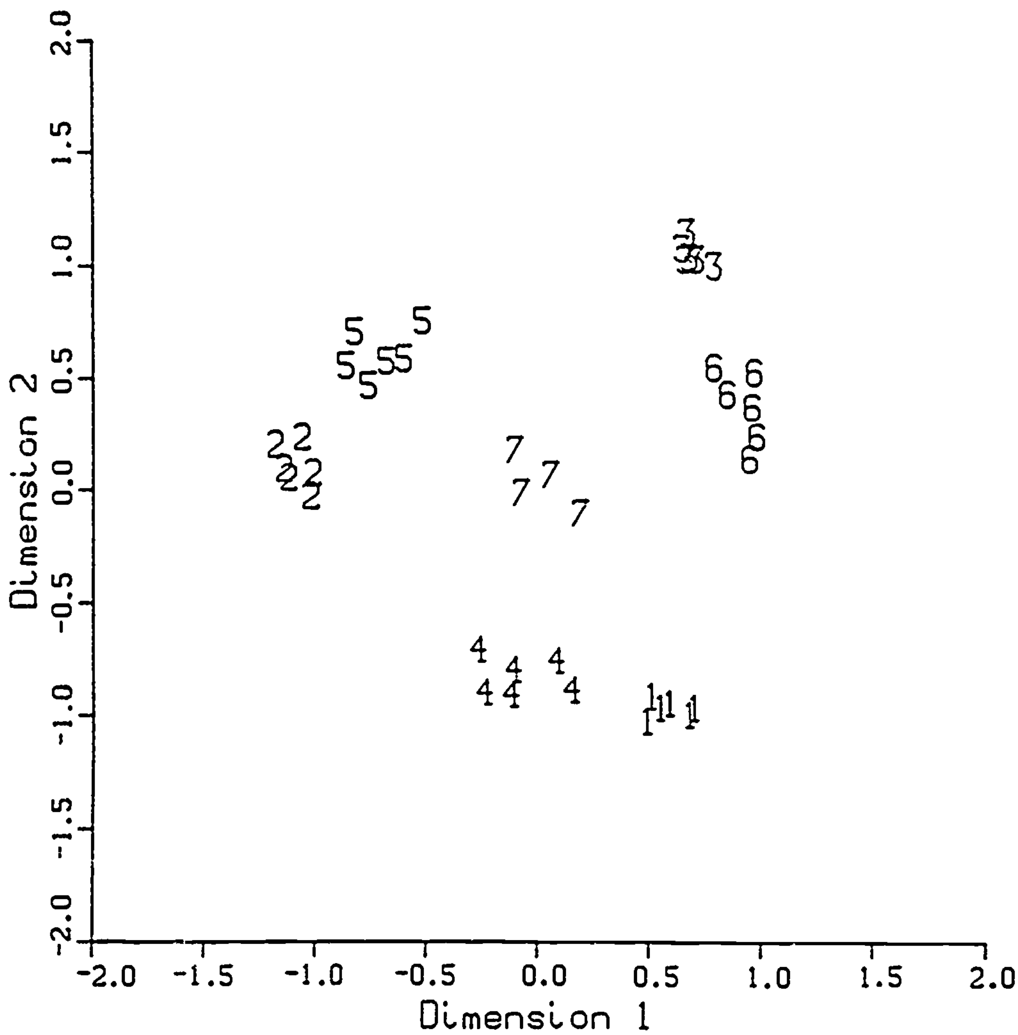

Figure 3. MDS Plot for three-dimensional simulated data set, tetrachoric correlation similarity matrix.

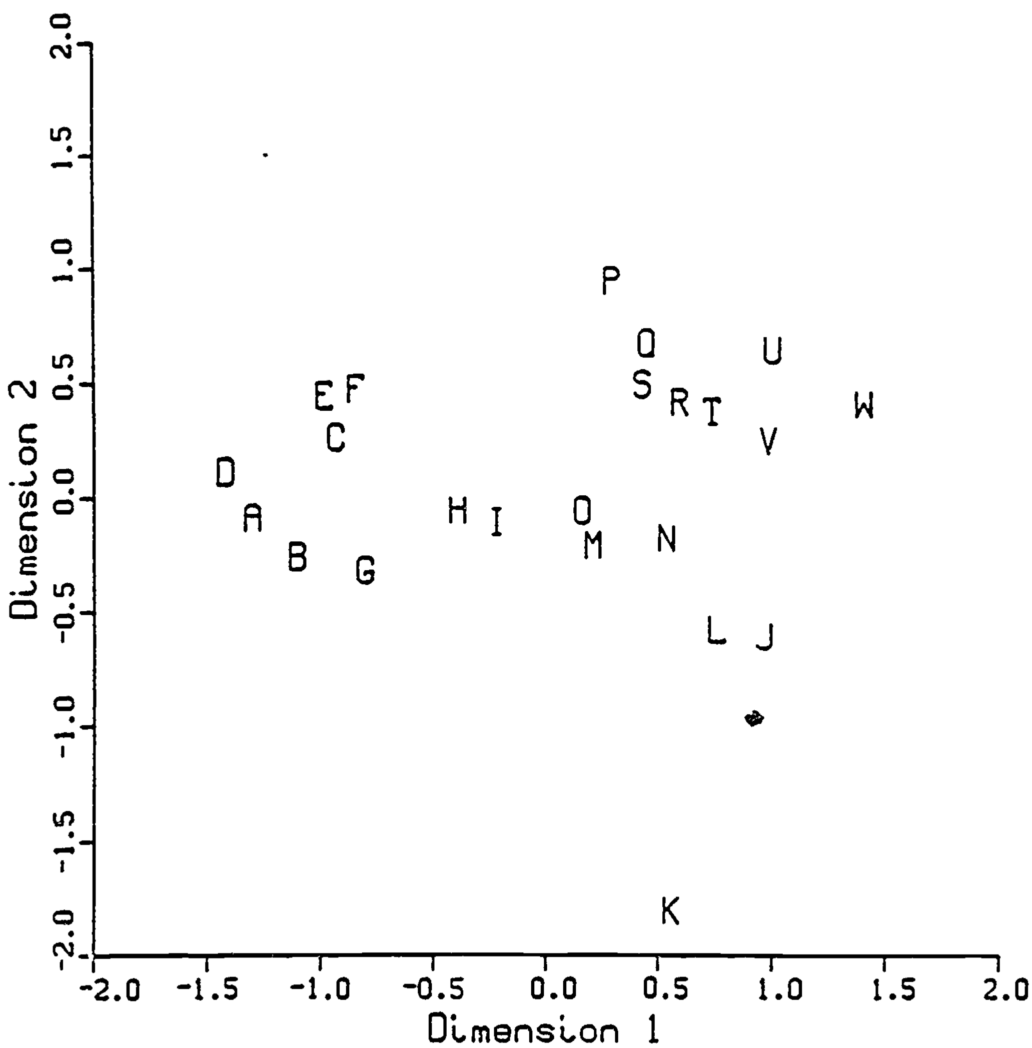

Figure 4. MDS Plot for HSMB data, Yule's Q similarity matrix.

REFERENCES

Allen, N. L. (1987). An examination of the complexity of the stimulus space of subtests from a standardized achievement battery. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D. C.

Bergan, J. (1986). Head Start Measures Battery. Test developed for the Department of Health and Human Services through the Administration for Children, Youth, and Families. Contract #HHS-105-81-C-008.

Bishop, Y. M., Fienberg, S. W., and Holland, P. W. (1975). Discrete Multivariate Analysis. Cambridge: MIT Press.

Ferguson, G. A. (1941). The factorial interpretation of test difficulty. Psychometrika, 6, 323-329.

Hulin, C. L., Drasgow, F., and Parsons, C. K. (1983). Item Response Theory: Application to Psychological Measurement. Homewood, IL: Dow Jones-Irwin.

Jones, P. B., Sabers, D. and Trosset, M. (1987). Assessment of dimensionality in dichotomously-scored data using multidimensional scaling. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D. C.

Guilford, J. P. (1941). The difficulty of a test and its factor composition. Psychometrika, 6, 67-77.

Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. Applied Psychological Measurement, 9, 139-164.

Koch, W. (1983). The analysis of dichotomous test data using nonmetric multidimensional scaling. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada. (ERIC Document Reproduction Service No. ED 235 204).

Korpi, M. and Haertel, E. (1984). Locating reading comprehension test items in multidimensional space: An alternative analysis of test structure. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 246 108).

Kruskal, J. B. (1964). Nonmetric multidimensional scaling: a numerical method. Psychometrika, 29, 115-129.

Kruskal, J. B. and Wish, M. (1978). Multidimensional Scaling. Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-011. Beverly Hills and London: Sage Publications.

Lord, F. M. and Novick, M. R. (1968). Statistical Theories of Mental Test Scores. Reading, MA: Addison-Wesley.

McDonald, R. P. and Ahlawat, K. S. (1974). Difficulty factors in binary data. British Journal of Mathematical and Statistical Psychology, 27, 82-99.

Reckase, M. D. (1981). The formation of homogeneous item sets when guessing is a factor in item responses. Office of Naval Research Report 81-5.

Thissen, D. L. (1984). MULTILOG: a user's guide. Mooresville, IN: Scientific Software.

Thomas, J. A. (1984). A comparison of factor analysis and multidimensional scaling in determining underlying structure. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 244 994).

Thomas, J. A. (1985). Evaluation of the construction of the subscales for the Piers-Harris and Tennessee Inventories. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Wilkinson, L. (1984). SYSTAT: the System for Statistics. Evanston, IL: Systat, Inc.