DOCUMENT RESUME

ED 220 506 TM 820 521.

AUTHOR McKinley, Robert L.; Reckase, Mark D.
TITLE Multidimensional Latent Trait Models.
INSTITUTION American Coll. Testing Program, Iowa City, Iowa.
SPONS AGENCY Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE Mar 82
CONTRACT N0014-81-K0817
NOTE 28p.; Paper presented at the National Council on Measurement in Education (New York, NY, March 20-22, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Analysis; Factor Analysis; Goodness of Fit; *Latent Trait Theory; *Mathematical Models; Measurement Techniques; *Multidimensional Scaling; Research Methodology
IDENTIFIERS *Rasch Model

ABSTRACT

The usefulness of the general Rasch model for multidimensional data, from the most simple formulations to the more complex versions of the model, is explored. Also investigated was whether the parameters of the models could be readily interpreted. Models investigated included: (1) the vector model; (2) the product term model; (3) the vector and product term model; (4) the reduced vector and product term model; and, (5) the item cluster model. Of the models investigated, all but the reduced vector and product term model and the item cluster model were rejected as incapable of reasonably modelling realistic multidimensional data. The item cluster model appears to be a useful model, but its applications may be limited in scope. The reduced vector and product term model was found to be the most capable of modelling realistic multidimensional data. Although the estimation of the parameters of the reduced vector and product term model may be more difficult than it would be for other models, this model appears to be the model that is most worth pursuing. (Author/PN)

Multidimensional Latent Trait Models

Robert L. McKinley
Mark D. Reckase

American College Testing Program

Latent trait theory has become an increasingly popular area for research and application in recent years. Areas of application of latent trait theory have included tailored testing (McKinley and Reckase, 1980), equating (Marco, 1977; Rentz and Bashaw, 1977), test scoring (Woodcock, 1974), and criterion-referenced measurement (Hambleton, Swaminathan, Cook, Eignor, and Gifford, 1978). While many of these applications have been successful, they are limited to areas in which the tests used measure predominantly one trait. This limitation is a result of the fact that most latent trait models that have been proposed assume unidimensionality, and require tests that have one predominant factor. Because of this, in many situations latent trait models have not been successfully applied. For example, in achievement testing the goal is not to measure a single trait, but to sample the content covered by instruction. Therefore, most latent trait models are inappropriate since tests designed for this purpose generally cannot be considered to be unidimensional.
Even when the goal is to measure a single trait, if dichotomously scored items are used no generally accepted method exists for forming unidimensional item sets, for determining the dimensionality of existing item sets, or for testing the fit of the model to the data.

An alternative to trying either to construct unidimensional item sets or to fit a unidimensional model to already existing item sets is to develop a multidimensional latent trait model. Several such models have been proposed (Rasch, 1961; Samejima, 1974; Sympson, 1978; Whitely, 1980), but little research has been done on them. Some work has been completed on the estimation of the parameters of the multidimensional Rasch model (Reckase, 1972), and the Whitely model (Whitely, 1980), but no extensive research has been completed on the characteristics and properties of any of these models. The purpose of this paper is to present the results of research on the characteristics and properties of the multidimensional Rasch model.

Method

Design

The general design of this research was to start with the most simple formulation of the multidimensional Rasch model, and then to investigate increasingly more complex versions of the model. At each stage the properties of the model were investigated, and the reasonableness and usefulness of the model explored. This was done primarily by simulating test data to fit

Paper presented at the annual meeting of the National Council on Measurement in Education, New York, March, 1982. This research was supported by Contract Number N0014-81-K0817 from the Personnel and Training Research Programs of the Office of Naval Research.

the particular form of the model being investigated, and analyzing that data in an attempt to assess how well the characteristics of the data match the characteristics of real data.

The most general formulation of the model investigated in this research is the model described by Rasch (1961). This model is given by:

$$P(x|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp[\phi(x)'\theta_j + \psi(x)'\sigma_i + \chi(x)\theta_j\sigma_i + \rho(x)], \qquad (1)$$

where θ_j represents the ability parameters for Person j, σ_i is the difficulty parameter for Item i, $P(x|\theta_j, \sigma_i)$ is the probability of response x to Item i by Person j, ϕ, ψ, χ, and ρ are scoring formulae which are functions of x only and $\gamma(\theta_j, \sigma_i)$ is a normalizing constant necessary to make the probabilities sum to 1.0. In order to apply this model to multidimensional data, θ_j, σ_i, ϕ and ψ are interpreted as column vectors, $\phi(x)'\theta_j$ and $\psi(x)'\sigma_i$ are interpreted as inner products of vectors, $\chi(x)$ is interpreted as a matrix, and $\theta_j'\chi(x)\sigma_i$ is interpreted as a homogeneous bilinear form in the elements of θ_j and σ_i (Rasch, 1961). The more simple formulations of the model used in this research were obtained by eliminating different terms from the model statement by setting the appropriate scoring functions equal to zero.

For each model statement that was obtained, simulation test data were generated to fit the model. Using the known parameters and model statement, predictions were made as to the dimensionality of the data and the characteristics of the items. Analyses were then performed on the simulation data in order to test the predictions. If it were found that a model statement could not be used to generate realistic data, in terms of either dimensionality or item characteristics, then the model was rejected, and a different model statement was investigated. This involved altering which terms of the general model (Equation 1) were retained, and which were zeroed out. In some cases, all of the terms in a particular rejected model statement were retained, and one or more additional terms from the general model were added. The models that were obtained were labeled as the vector model, the product term model, the vector and product term model, the reduced vector and product term model, and the item cluster model.

Analyses

The first analysis performed on the simulation data for any model statement was a factor analysis. Factor analysis, in this case, is not being used as a means of validating the models, but as a means of determining whether the data generated from the models have characteristics similar to those of real test data. All of the factor analyses performed in this research were performed using the principal components procedure on phi coefficients. When the obtained and expected dimensionality did not match, follow-up analyses were performed in an attempt to determine why the obtained dimensionality was different from what was predicted.

Follow-up analyses included plotting the true item parameters against the factor loadings and against traditional item statistics such as proportion-correct difficulty values and point biserial discrimination values. These analyses were performed using both the unrotated factor loading matrix and the factor loading matrix rotated to the varimax criterion. The purposes of the analyses were two-fold. One purpose was to determine whether the obtained factor structure of the data was a result of the model statement, the values used for the model parameters, or both. The second purpose was to facilitate interpretation of the model parameters and to determine whether the model yielded items with reasonable characteristics. In many cases it was necessary to generate additional data, using different values for the parameters of the model, in order to answer specific questions about a particular model statement.

A final analysis performed on each model statement involved an attempt to predict, if a model statement were rejected, what changes would yield a more acceptable model. What was acceptable was defined not only in terms of whether the simulation data factor structure was realistic, but also in terms of whether the model parameters had reasonable and useful interpretations.

Results

Vector Model

The first model that was investigated was a simple vector parameter model. The $\chi(x)\theta_j\sigma_i$ and $\rho(x)$ terms were eliminated, yielding the model given by

$$P(x|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\phi(x)'\theta_j + \psi(x)'\sigma_i), \tag{2}$$

where all the terms are as defined for Equation 1, and θ_j, σ_i, ϕ and ψ are vectors. This model was selected first because it appeared to be a straightforward extension of the unidimensional Rasch model (Rasch, 1960) to the multidimensional case. The expectation was that data generated according to this model would have a dimensionality that would vary with the number of elements in the parameter vectors. For instance, when data were generated using two elements in both the item and person parameter vectors, it was expected that the data would yield a two-factor solution when factor analyzed. This was not the case, however. Regardless of the number of elements in the parameter vectors, this model would yield one predominant factor. This was true regardless of what the actual values of the parameters were, or what values were used in the scoring functions.

Table 1 shows the first three eigenvalues from a typical principal component solution for the vector model. As can be seen, there is a dominant first factor, with two minor factors. Table 1 also shows the unrotated factor loading matrix obtained for this particular data, as well as the proportion-correct difficulty and the inner product of the item parameter vector and scoring function for each item. As can be seen, the two minor factors are difficulty factors.

Insert Table 1

Once it was ascertained that the vector model would not yield multi-factor data, it was not difficult to determine why. Equation 2 can be written as

$$P(x|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\alpha_j + \beta_i), \tag{3}$$

where $\alpha_j = \phi(x)'\theta_j$ and $\beta_i = \psi(x)'\sigma_i$. Equation 3 is the unidimensional Rasch model, with inner products of vectors as parameters. Therefore, regardless of what values the model parameters take on, the model is still a unidimensional model. The factor analyses typified by the solution shown in Table 1 serve as an empirical demonstration that the vector model is a unidimensional model. It can also be empirically demonstrated that the inner products of the scoring function and parameter vectors serve as parameters for the model. Figure 1 shows a plot of proportion-correct difficulty by the inner product of the scoring function and item parameter vectors. As can be seen, there is an almost perfect relationship between the inner products and the proportion-correct scores. When data were generated using the unidimensional Rasch model, with the inner products from the two-dimensional model as parameters, the exact same plot was obtained. Figure 2 demonstrates that the same was true for the person parameters.

Insert Figures 1 and 2

<u>Product Term Model</u>

It was clear from the results so far reported that using parameter vectors in an otherwise unidimensional model did not make it a multidimensional model. Therefore, at this point the vector model was rejected as a multidimensional model. The next model that was investigated had only the $\theta_j'\chi(x)\sigma_i$ term in it. This model was investigated next because it involved more than simple inner products of scoring and parameter vectors, but was more simple than using both inner products and the $\theta_j'\chi(x)\sigma_i$ term.

When θ_j and σ_i are vectors, $\chi(x)$ must be a matrix. The term $\theta_j\sigma_i'$ represents a matrix of products of all possible pairs of the elements in the θ_j and σ_i vectors. For two-dimensional θ_j and σ_i vectors,

$$\theta_j\sigma_i' = \begin{bmatrix} \theta_1\sigma_1 & \theta_1\sigma_2 \\ \theta_2\sigma_1 & \theta_2\sigma_2 \end{bmatrix}. \tag{4}$$

The $\chi(x)$ matrix, then, is a scoring matrix having an element for each element of the $\theta_j\sigma_i$ matrix. If the $\chi(x)$ matrix for the matrix in Equation 4 were

$$\chi(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (5)$$

for a particular response x then the numerator of the model statement for that response would be $\exp(\theta_1\sigma_1 + \theta_2\sigma_2)$. It is clear from this that by selectively using zeros in the $\chi(x)$ matrix, various combinations of θ_j and σ_i elements can be selected. Varying the values of the nonzero elements in $\chi(x)$ assigns different weights to different combinations. Thus, the product term model, given by

$$P(x|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\theta_j{}'\chi(x)\sigma_i), \qquad (6)$$

is a very rich model in terms of alternative formulations of the model that are available. Unfortunately, when data were simulated using some of these alternatives, a serious problem was discovered with the model. Regardless of which formulation of this model was used, and regardless of what values were taken on by the item parameters, the item proportion-correct difficulties were all approximately .5. A closer examination of the product term model indicates why this occurred. Using the item parameters shown in Table 2, data were generated using

$$\chi(x) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (7)$$

for a correct response, and

$$\chi(x) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad (8)$$

for an incorrect response. This yields a model given by

$$P(x|\theta_j,\sigma_i) = \begin{cases} \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\theta_{j1}\sigma_{i1} + \theta_{j2}\sigma_{i2}) & \text{for a correct response,} \\ \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\theta_{j1}\sigma_{i1} - \theta_{j2}\sigma_{i2}) & \text{otherwise.} \end{cases} \qquad (9)$$

where the θ and σ terms are elements in the θ and σ vectors. From Equation 9 it can be seen that the item parameters are similar to the discrimination parameter in the unidimensional two-parameter logistic (2PL) model set out by Birnbaum (1968): In fact, if written as

$$P(x = 1|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp[\sigma_{i1}(\theta_{ji} + 0) + \sigma_{i2}(\theta_{j2} + 0)], \qquad (10)$$

the model is essentially a two-dimensional two-parameter logistic model with both of the difficulty parameters equal to zero for all items. Because the data used for Table 2 were generated using a bivariate N(0,1) distribution of ability, a difficulty parameter of zero yielded a predicted proportion-correct difficulty of .5.

The role of the item parameters as discrimination parameters in this model is indicated by comparing the item parameters shown in Table 2 with the rotated factor loading matrix, also shown in Table 2. Table 3 contains the correlations between the item parameters and the factor loadings in Table 2. As can be seen in Table 3, there was a relationship between the item parameters and factor loadings ($\underline{r}$ = .69 for σ_1 with Factor 2, $\underline{r}$ = .82 for σ_2 with Factor 1). A principal components analysis of phi coefficients yielded evidence that use of two item parameters resulted in a two-dimensional model. The first eigenvalues obtained for the data generated using the item parameter values in Table 2 were 7.1, 2.5, and .9. Further evidence is given by the plot shown in Figure 3. This is the plot of the item point biserials by the sum of the item parameters. Since the item point biserials are the correlations of the items with total test score, instead of with scores on each dimension, the factors are essentially summed. As can be seen in Figure 3, there is a strong relationship between the item point biserials and the sum of the item parameters ($\underline{r}$ = .82).

Insert Table 3

Insert Figure 3

Vector and Product Term Model

The vector model that was investigated first was essentially a unidimensional model that contained a difficulty parameter (the inner product $\psi(x)\sigma_i$) as the only item parameter. The product term model is a multidimensional model that contains discrimination parameters as the only item parameters. In order to obtain a multidimensional model which contained a difficulty parameter, the vector and product term models were combined. A combination of these two models is given by Equation 1 (with or without the $\rho(x)$ term), which is the general Rasch model.

Table 4 shows the item parameters used to generate data to fit the general Rasch model, as well as the rotated factor loadings obtained from the first two factors from the principal components analysis of phi coefficients. The first three eigenvalues from the solution are 5.26, 2.28, and 1.07.

Initial analyses indicated that this model could be used to model multidimensional data, and that item difficulties were not constant (see Table 4). However, these analyses also indicated that it was not realistic to use the same item parameters in both the parameter vectors and the product term. The problem is clearly shown by Table 5, which contains the correlations of the item parameters with the item proportion-correct difficulties and item point biserials. As was the case with the vector model, the inner product $\psi(x)'\sigma_i$ defined the difficulty parameter. The correlation between $\psi(x)'\sigma_i$ and the proportion-correct scores was $r = .98$. The $\psi(x)'\sigma_i$ term was also the discrimination parameter, having a correlation of $r = .89$ with the item point biserials. The problem is that, because $\psi(x)'\sigma_i$ was both the difficulty and discrimination parameter, the proportion-correct scores and point biserials had a correlation of $r = .94$. That is not a very realistic situation.

Insert Tables 4, 5

Reduced Vector and Product Term Model

Since the initial analyses of the general Rasch model indicated that parameters should not appear in both the parameter vectors and the product term, the scoring functions were altered so that parameters appeared in one or the other, but not both. In order to facilitate this, additional elements were inserted into the item parameter vector. The resulting model is given by

$$P(x = 1|\theta_j,\sigma_i) = \frac{1}{\gamma(\theta_j,\sigma_i)} \exp(\sigma_{i1} + \sigma_{i2} + \sigma_{i3}\theta_{j1} + \sigma_{i4}\theta_{j2}), \qquad (11)$$

where the θ and σ terms are again scalars.

The first three eigenvalues obtained from the principal components analysis for this model are 5.39, 1.30, and .99. Table 6 shows the item parameters that were used to generate the data, as well as the obtained rotated factor loadings. Table 7 shows the correlation matrix for the item parameters, loadings, and traditional statistics.

Insert Tables 6, 7

The results of the factor analysis of these data indicate that a dominant first factor is present. However, there was a second component present in the data which was strongly related to the item parameters ($r = .87$ for σ_3 and Factor 1, $r = .87$ for σ_4 and Factor 2). The item parameters in the product term, then, were related to the factor loadings, while the sum of the item parameters in the vector term behaved as a difficulty parameter, having a correlation with the proportion-correct difficulty of $r = .98$. There was not a significant correlation between the item difficulty and point biserial values ($r = .12$). The sum of σ_3 and σ_4 had a correlation of $r = .96$ with the item point biserials.

The analyses of the model set out in Equation 11 indicate that it has many desired characteristics. The rotated factor loadings are highly related to the item parameters in the product term, the item difficulty is highly correlated with the sum of the item parameter vector elements, and there is no correlation between item difficulty and item discrimination.

One problem that does exist with the data that were generated is that they have one predominant factor. From the factor analysis results it would not be difficult to conclude that the data had only one factor. One possible reason for this is that so many of the items had large values for both of the item parameters in the product term. In order to test this, data were generated for the set of item parameters shown in Table 8. As can be seen in Table 8, the first eight items have large values for σ_3 and small values for σ_4 while Items 11 through 18 have large values for σ_4 and small values for σ_3. Items 9, 10, 19, and 20 have equal values for σ_3 and σ_4. Table 8 also contains the rotated factor loadings obtained for these data.

Insert Table 8

The first four eigenvalues from the principal components analysis obtained for these data are 3.12, 1.77, 1.08, and 1.01. As can be seen, the second component is now larger. When data were generated using the item parameters shown in Table 9, the second factor was even greater. The eigenvalues from the principal components analysis for these data are 2.49, 2.28, 1.05, and 1.03. As can be seen, when using the item parameters from Table 9 to generate data, there are two factors of approximately equal magnitude present in the data.

Insert Table 9

Item Cluster Model

Although the reduced vector and product term model appears to adequately model multidimensional data, the presence of the product term seriously complicates parameter estimation, since separation of the item and person parameters is not possible through conditional estimation. Because of this, one more model that does not have a product term was investigated. This model is the item cluster model.

One of the reasons the item vector model, given by Equation 2, does not adequately model multidimensional data is that no information about the different dimensions is preserved in the item score when the item is dichotomously scored. The elements for the different dimensions are summed, and the sums are treated as parameters. If it were possible to score the dimensions separately, then the vector model might be able to model multidimensional data. This requires, however, polychotomous item scoring. Scoring an item on each dimension would require 2^n response categories, where n is the number of dimensions. Unfortunately, most test data are not scored polychotomously.

An alternative to having polychotomous item scoring is to consider more than one item at a time. If two dichotomously scored items are clustered together, and the cluster is treated as a single unit, then the cluster has 2^2 or 4 response categories - (0,0), (0,1), (1,0), and (1,1). The model given by Equation 2 can then be applied, with the exception that the σ vector now has two elements, both representing the same cluster. Essentially each item is considered to be unidimensional, and what is modelled is a two item, two-dimensional test. It would probably be best to treat the entire test as a cluster, but if more than a few items are on the test, the computations become impractical.

The procedure by which this model was investigated is as follows. For the two-dimensional case, item parameters were selected for 20 items. These parameters are shown in Table 10. The items were paired so that Items 1 and 2 formed Cluster 1, Items 3 and 4 formed Cluster 2, and so on until 10 clusters were formed. For each cluster there were four response categories, which were scored as follows:

a) (0,0) for incorrect on both items;
b) (0,1) for first item incorrect, second item correct;
c) (1,0) for first item correct, second item incorrect;
and d) (1,1) for both items correct.

This is essentially treating the two items in a given cluster as independent. Table 10 contains the unrotated factor loadings for the first two principal components, and the first four eigenvalues are 3.61, 3.06, 1.33, and 1.21.

Insert Table 10

As can be seen, the simulation data were treated as 20 items, rather than as 10 clusters. The eigenvalues listed above indicate that there were two roughly equal components in the data. Table 10 shows that the first component was defined by the items that were placed first in the cluster, and the second component was defined by the items that were in the second position in the cluster. Consistent with the scoring functions, there were two equal independent factors.

In order to demonstrate that the factors need not be independent, the same item parameters were used to generate data using the following scoring functions:

a) (0,0) for both items incorrect;
b) (.1, .9) for first item incorrect, second item correct;
c) (.9, .1) for first item correct, second item incorrect;
and d) (1,1) for both items correct.

The principal components analysis of phi coefficients for this model yield six factors with eigenvalues greater than one [2.46, 1.83, 1.09, 1.08, 1.01, 1.00]. Table 11 shows the unrotated factor loadings. As can be seen, there are still two factors present in the data. However, the factors are no longer defined only by the items in the corresponding position in the cluster. The first component is a general factor, while the second component discriminates between the items in the first and second positions in the cluster. Clearly these two sets of items are not independent.

Insert Table 11

Discussion

The use of simulation data to study the characteristics of a model before applying it is perhaps atypical of research on latent trait models. It is not unusual in this area to adopt a model, derive estimation procedures, and apply the model without ever going through the process this study has employed. In this study this approach has been taken for two main reasons. First, it was felt that when dealing with multidimensional latent trait models much of the common knowledge about latent trait models might no longer apply. It was felt that considerable research was necessary in order to gain an understanding of how these models work and what the model parameters represent before they could be applied. This belief has been borne out several times in this study by findings indicating that the models were not behaving in the anticipated manner.

A second reason for taking this approach was that it seemed impractical to attempt to develop estimation procedures for some of these models. Specifically, the general model set out by Rasch has a very large number of parameters. It seemed impractical to try to estimate all of them, and it was hoped that research on the model could help simplify the problem by eliminating some terms of the model and by discovering restrictions on the values the parameters could reasonably take on. With these goals in mind, the results of this study will now be discussed.

Vector Model

The most simple formulation of the general model that was investigated was the vector model. This model is simply the unidimensional Rasch model, but with vectors for parameters instead of scalars. This model was found to be totally inadequate for modelling multidimensional data. When data were generated according to this model, the resulting data were unidimensional, with item characteristics determined by the inner product of the item parameter vectors and scoring functions. From this it follows that this model would fit multidimensional data no better than a unidimensional model having parameters equal to the inner products from the vector model.

Product Term Model

Because of its slight similarity to the 2PL model, it was felt that the product term model would be better able to model multidimensional data. It was anticipated that the item parameters in the product term would behave as discrimination indices, and that is just how they did behave. Unfortunately, without the vector terms in the model there were no terms playing the role of difficulty parameters. The data generated for this model had items of constant difficulty, which does not seem very realistic. From this it was concluded that this model would be useful only for modelling items of constant difficulty, and when items have varying difficulties this model is inappropriate.

Vector and Product Term Model

Based on the findings for the vector model and the product term model, it was hypothesized that a combination of the two models would be necessary to model items that were both multidimensional and of nonconstant difficulty. Analyses of the vector and product term model indicated that it would model multidimensional data, and that it would model items of varying difficulty. However, it was also found that, as long as the item parameter vector elements appeared both in the vector terms and in the product term, the item difficulties and discriminations would be highly correlated. Since this is rarely the case, it was concluded that this model would be useful only in a very limited number of circumstances.

Reduced Vector and Product Term Model

In order to correct the problems with the vector and product model, it was clear that a given item parameter vector element should appear only in the vector term or the product term, but not both. It was anticipated that similar problems might exist if the person parameter vector elements occurred in both the vector term and the product term, so the same correction was made for the person parameters as was made for the item parameters.

The resulting model appears to be quite successful at modelling realistic multidimensional data. It is capable of modelling correlated as well as independent factors, and the item parameters are readily interpretable. The only real problem there seems to be with this model is with the estimation of the parameters. Although there are fewer parameters to estimate than is the case with the general model, there are still a fair number to estimate. Moreover, it appears that there are no observable sufficient statistics for the parameters in the product term. These problems do not make estimation of the model parameters impossible, and probably not even impractical. However, they do make estimation more difficult.

Item Cluster Model

The item cluster model was proposed as an alternative to the vector model. This model does not involve a product term, but it still can successfully model multidimensional data. However, it does involve clustering items, which gives rise to a number of new problems. For instance, as yet it is unclear what the effect is of different combinations of items, or whether all items should be clustered with the same item. Preliminary investigations seem to indicate that the optimal clustering procedure is to cluster all items on a subtest with one item taken from a different subtest. As of yet, however, no clear results are available. While this model shows considerable promise, its usefulness is not well established, and may be limited in the types of circumstances in which it can be applied.

Summary and Conclusions

The purpose of this study was to investigate the usefulness of the general Rasch model for multidimensional data. Several formulations of the model, varying in complexity, were investigated to determine whether they could successfully model realistic multidimensional data. Also investigated was whether the parameters of the models could be readily interpreted. Models investigated included: a) the vector model; b) the product term model; c) the vector and product term model; d) the reduced vector and product term model; and, e) the item cluster model.

Of the models investigated, all but the reduced vector and product term model and the item cluster model were rejected as incapable of reasonably modelling realistic multidimensional data. The item cluster model appears to be a useful model, but its applications may be limited in scope. The reduced vector and product term model was found to be the most capable of modelling realistic multidimensional data. Although the estimation of the parameters of the reduced vector and product term model may be more difficult than it would be for other models, this model appears to be the model that is most worth pursuing.

References

Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord and M.R. Novick, Statistical theories of mental test scores. Reading, MA: Addison-Wesley, 1968.

Hambleton, R.K., Swaminathan, H., Cook, L.L., Eignor, D.R., and Gifford, J.A. Developments in latent trait theory, models, technical issues, and applications. Review of Educational Research, 1978, 48, 467-510.

Marco, G.L. Item characteristic curve solutions to three intractable testing problems. Journal of Educational Measurement, 1977, 14, 139-160.

McKinley, R.L., and Reckase, M.D. A successful application of latent trait theory to tailored achievement testing (Research Report 80-1). Columbia: University of Missouri, Department of Educational Psychology, February 1980.

Rasch, G. On general laws and the meaning of measurement in psychology. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, 1961, 4, 321-334.

Rasch, G. Probabilistic models for some intelligence and attainment tests. Copenhagen: Danish Institute for Educational Research, 1960.

Reckase, M.D. Development and application of a multivariate logistic latent trait model (Doctoral dissertation, Syracuse University, 1972). Dissertation Abstracts International, 1973, 38. (University Microfilms No. 73-7762)

Rentz, R.R., and Bashaw, W.L. The National Reference Scale for Reading: An application of the Rasch model. Journal of Educational Measurement, 1977, 14(2), 161-180.

Samejima, F. Normal ogive model for the continuous response level in the multidimensional latent space. Psychometrika, 1974, 39, 111-121.

Sympson, J.B. A model for testing with multidimensional items. In D.J. Weiss (Ed.), Proceedings of the 1977 Computerized Adaptive Testing Conference. University of Minnesota, 1978.

Whitely, S.E. Measuring aptitude processes with multicomponent latent trait models (Technical Report No. NIE-80-5). Lawrence: University of Kansas, Department of Psychology, July 1980.

Woodcock, R.W. Woodcock Reading Mastery Tests. Circle Pines, Minnesota: American Guidance Service, 1974.

Table 1

Principal Component Factor Loadings Based on Phi Coefficients with the Sums of the Item Parameters and Observed Proportion Correct for Two-Dimensional Vector Rasch Model.

ITEM	SUMS OF ITEM PARAMETERS	OBSERVED DIFFICULTY	FACTOR LOADINGS		
			FACTOR 1	FACTOR 2	FACTOR 3
1	.89	.69	.69	.29	-.10
2	-.89	.30	.73	.09	-.22
3	.43	.59	.59	.44	.12
4	2.02	.89	.33	.45	.52
5	.59	.64	.58	.45	.12
6	-.91	.28	.74	.03	-.23
7	-1.44	.19	.70	-.11	-.11
8	.47	.60	.75	.02	-.22
9	-1.05	.25	.61	-.39	.11
10	-1.76	.14	.49	-.45	.34
11	.98	.71	.71	.15	-.19
12	2.58	.93	.49	.46	.28
13	-1.31	.21	.58	-.44	.19
14	1.22	.75	.71	.25	-.19
15	-.64	.35	.67	-.35	-.01
16	.05	.52	.70	-.14	-.12
17	-2.33	.08	.41	-.43	.44
18	-.54	.38	.67	-.30	-.05
19	-.60	.37	.71	-.29	-.05
20	2.26	.92	.53	.43	.19
Eigenvalue			7.94	2.22	1.04

Table 2

Item Parameters, Proportion Correct Scores, and Factor Loadings from a Varimax Rotated Principal Components Solution on Phi Coefficients for the Product Term Model

Item	σ_1	σ_2	p	Factor I	Factor II
1	.81	1.54	.50	.72	.30
2	1.56	.51	.49	.16	.77
3	.36	.63	.50	.51	.20
4	.56	1.62	.49	.76	.19
5	.16	.59	.53	.55	-.01
6	1.46	.39	.49	.08	.76
7	.21	1.49	.52	.75	.07
8	.35	.28	.51	.28	.35
9	.33	.88	.52	.65	.16
10	.30	1.90	.50	.80	.04
11	.53	.61	.50	.42	.36
12	1.03	1.71	.50	.72	.37
13	1.09	.38	.52	.14	.69
14	2.01	.64	.50	.19	.80
15	.22	.70	.53	.53	.13
16	.57	.46	.51	.30	.52
17	1.70	.79	.50	.30	.72
18	.50	.88	.50	.61	.20
19	.31	.45	.50	.40	.24
20	2.63	.22	.50	.02	.84

Table 3

Intercorrelation Matrix for Item Parameters, Item Statistics, and Factor Loadings for the Product Term Model

Variable	σ_1	σ_2	$\sigma_1 + \sigma_2$	Factor 1	Factor 2	P-Value	Pt. Bis.
σ_1	-	.13	.76	.49	.69	.75	.68
σ_2		-	.74	.82	-.58	.72	.66
$\sigma_1 + \sigma_2$			-	.87	.09	.98	.89
Factor 1				-	-.18	.92	.96
Factor 2					-	.08	.11
P-Value						-	.94
Pt. Bis.							-

Table 4

Item Parameters, Proportion Correct Scores, and Rotated Factor Loadings for the Vector and Product Term Model

Item	σ_1	σ_2	p	Factor 1	Factor 2
1	.230	.190	.56	.63	.12
2	.880	2.180	.77	.73	-.10
3	.900	-1.920	.35	.08	.71
4	-.900	-.110	.32	.39	-.22
5	-.640	.830	.52	.58	-.35
6	-.540	-1.040	.16	.10	.31
7	1.730	-1.350	.54	.34	.65
8	.940	.630	.69	.69	.21
9	.030	-.110	.46	.56	.11
10	-1.610	-.570	.13	.01	-.37
11	-1.170	1.260	.74	.75	.05
12	-.550	-1.070	.17	.04	.22
13	-.420	-.480	.31	.32	-.01
14	.220	-.070	.53	.60	.15
15	-.020	-1.670	.21	-.04	.55
16	2.420	.370	.78	.61	.40
17	1.230	.400	.69	.68	.28
18	.250	.410	.58	.65	.08
19	.140	.760	.64	.67	-.10
20	-1.770	.550	.30	.27	-.60

Table 5

Intercorrelation Matrix for Item Parameters, Rotated Factor Loadings, and Item Statistics for the Vector and Product Model

Variable	σ_1	σ_2	$\sigma_1 + \sigma_2$	Factor 1	Factor 2	P-Value	Pt. Bis.
σ_1	-	.13	.76	.49	.69	.75	.68
σ_2		-	.74	.82	-.58	.72	.66
$\sigma_1 + \sigma_2(\psi(x)\sigma)$			-	.87	.09	.98	.89
Factor 1				-	-.18	.92	.96
Factor 2					-	.09	.11
P-Value						-	.94
Pt. Bis.							-

Table 6

Item Parameters and Rotated Factor Loadings for the Reduced Vector and Product Model

Item	σ_1	σ_2	σ_3	σ_4	Factor 1	Factor 2
1	.206	-.503	.373	.997	.51	.01
2	-.164	.888	1.205	1.832	.60	.32
3	.448	.261	.766	.876	.34	.36
4	.814	-.008	1.321	1.714	.55	.34
5	.111	-.908	1.344	1.216	.41	.42
6	-.947	.044	1.758	1.694	.46	.51
7	-.490	.111	.687	.738	.40	.24
8	.553	-.502	.347	1.454	.61	.07
9	-.344	.639	1.307	.127	-.06	.64
10	-.257	.303	.851	.824	.26	.39
11	-.069	-.542	.472	.404	.22	.25
12	.779	.432	.392	.656	.30	.13
13	-.611	.571	.578	1.252	.59	.13
14	-.140	-1.032	.334	1.066	.60	-.04
15	-.705	.081	.821	.480	.07	.44
16	-.386	-.164	1.912	.244	.03	.71
17	-.154	.044	1.193	.537	.16	.56
18	.474	.249	1.385	1.287	.49	.44
19	.438	-.210	1.320	1.110	.42	.45
20	.294	.190	1.634	1.492	.45	.49

Table 7

Intercorrelation Matrix for Item Parameters, Item Statistics, and Factor Loadings for the Reduced Vector and Product Model

Variable	$\sigma_1 + \sigma_2$	σ_3	σ_4	$\sigma_3 + \sigma_4$	Factor 1	Factor 2	P-Value	Pt. Bis.
$\sigma_1 + \sigma_2$	-	.04	.20	.16	.11	.05	.98	.18
σ_3		-	.15	.76	-.25	.87	.02	.70
σ_4			-	.76	.87	-.27	.15	.76
$\sigma_3 + \sigma_4$				-	.40	.40	.11	.96
Factor 1					-	-.65	.06	.48
Factor 2						-	.04	.36
P-Value							-	.12
Pt. Bis.								-

Table 8

Item Parameters and Rotated Factor Loadings for the Reduced Vector and Product Model

Item	σ_1	σ_2	σ_3	σ_4	Factor 1	Factor 2
1	.206	-.503	.850	.150	.42	-.05
2	-.164	.888	.900	.200	.48	.03
3	.448	.261	.950	.250	.50	.10
4	-.814	-.008	1.000	.300	.46	.08
5	.111	-.908	1.050	.250	.53	.04
6	-.947	.044	1.100	.200	.49	.10
7	-.490	.111	1.150	.150	.50	.05
8	.553	-.502	1.200	.100	.58	-.06
9	-.344	.639	.500	.500	.28	.23
10	-.257	.303	.700	.700	.34	.26
11	-.069	-.542	.150	.850	.07	.49
12	.779	.432	.200	.900	.06	.44
13	-.611	.571	.250	.950	.06	.52
14	-.140	-1.032	.300	1.000	.10	.45
15	-.705	.081	.250	1.050	.06	.48
16	-.386	-.164	.200	1.100	.09	.53
17	-.154	.044	.150	1.150	.03	.55
18	.474	.249	.100	1.200	-.07	.63
19	.438	-.210	.700	.700	.37	.32
20	.294	.190	.500	.500	.31	.24

Table 9

Item Parameters for the Reduced Vector and Product Model

Item	σ_1	σ_2	σ_3	σ_4
1	.977	.258	1.000	.000
2	.359	.728	.000	1.000
3	-.322	.377	1.000	.000
4	-1.289	1.128	.000	1.000
5	-.613	.219	1.000	.000
6	1.299	.797	.000	1.000
7	.029	-.213	1.000	.000
8	-.360	-.862	.000	1.000
9	.769	-.487	1.000	.000
10	-1.447	2.092	.000	1.000
11	-1.252	-.243	1.000	.000
12	-.778	-1.426	.000	1.000
13	.668	-1.860	1.000	.000
14	2.102	-.025	.000	1.000
15	-.724	.968	1.000	.000
16	1.230	-.535	.000	1.000
17	.260	-1.216	1.000	.000
18	-1.092	-.432	.000	1.000
19	-.994	1.479	1.000	.000
20	-.206	-.525	.000	1.000

Table 10

Unrotated Factor Loadings on First Two Principal Components for the Independent Two-Dimensional Item Cluster Model

Item	Factor 1	Factor 2	Item Parameters
1	.56	.00	.893
2	.02	.65	-.850
3	.66	-.02	-.892
4	.01	.66	.690
5	.64	.00	.430
6	.02	.19	3.200
7	.36	-.04	2.016
8	.07	.22	-3.310
9	.61	-.05	.594
10	.00	.69	.470
11	.66	.01	-.913
12	.06	.55	1.220
13	.58	-.01	-1.437
14	.00	.58	-1.260
15	.65	-.07	.467
16	.04	.62	.880
17	.66	-.04	-1.048
18	-.02	.64	-.970
19	.56	.07	-1.760
20	.01	.42	-2.140

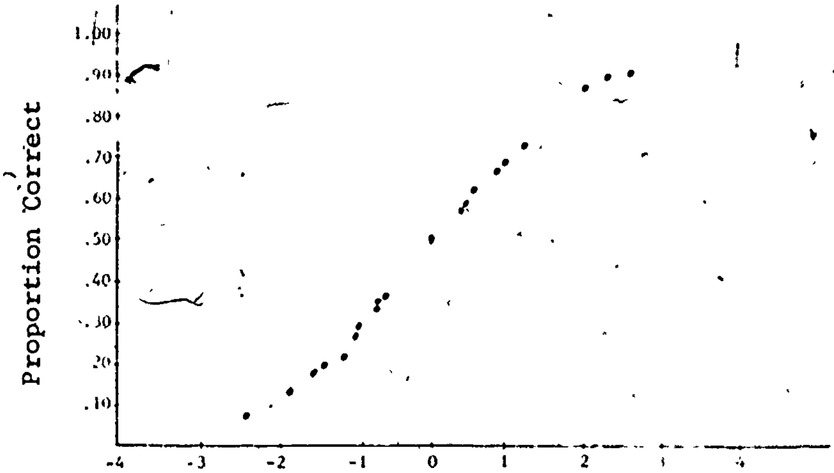
Table 11

Unrotated Factor Loadings on First Two Principal Components for the Dependent Two-Dimensional Item Cluster Model

Item	Factor 1	Factor 2
1	.40	.21
2	.33	-.36
3	.38	.24
4	.33	-.33
5	.41	.20
6	.25	-.24
7	.36	.16
8	.10	-.35
9	.40	.33
10	.26	-.39
11	.47	.30
12	.32	-.38
13	.37	.20
14	.32	-.43
15	.45	.28
16	.28	-.36
17	.44	.28
18	.28	-.37
19	.43	.10
20	.20	-.31

Figure 1

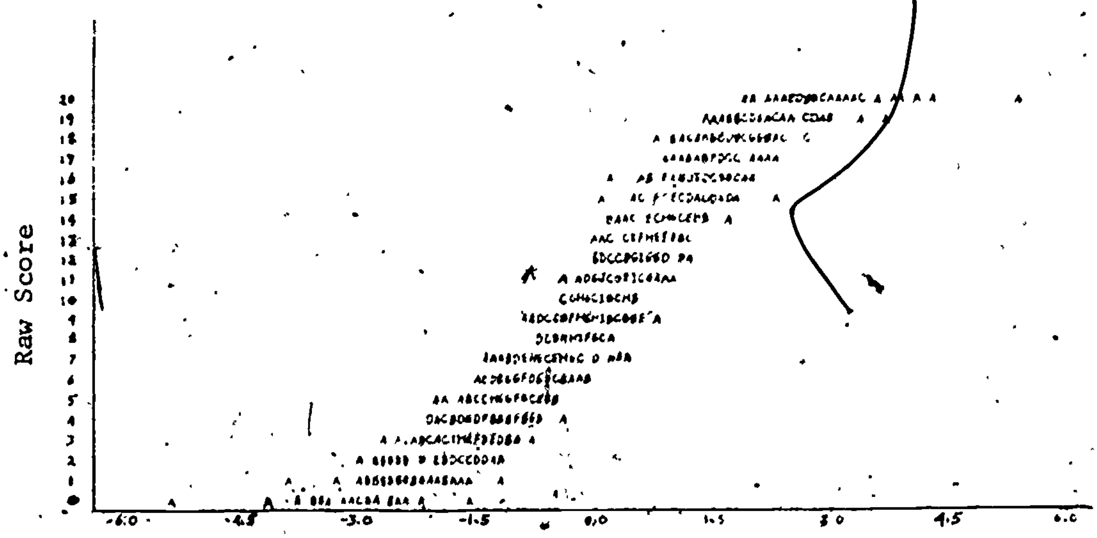
Relationship Between the Proportion Correct and the Inner Product of the Item Parameters for Twenty Items Generated Using a Two-Dimensional Vector Model



Inner Product for Item Parameters

Figure 2

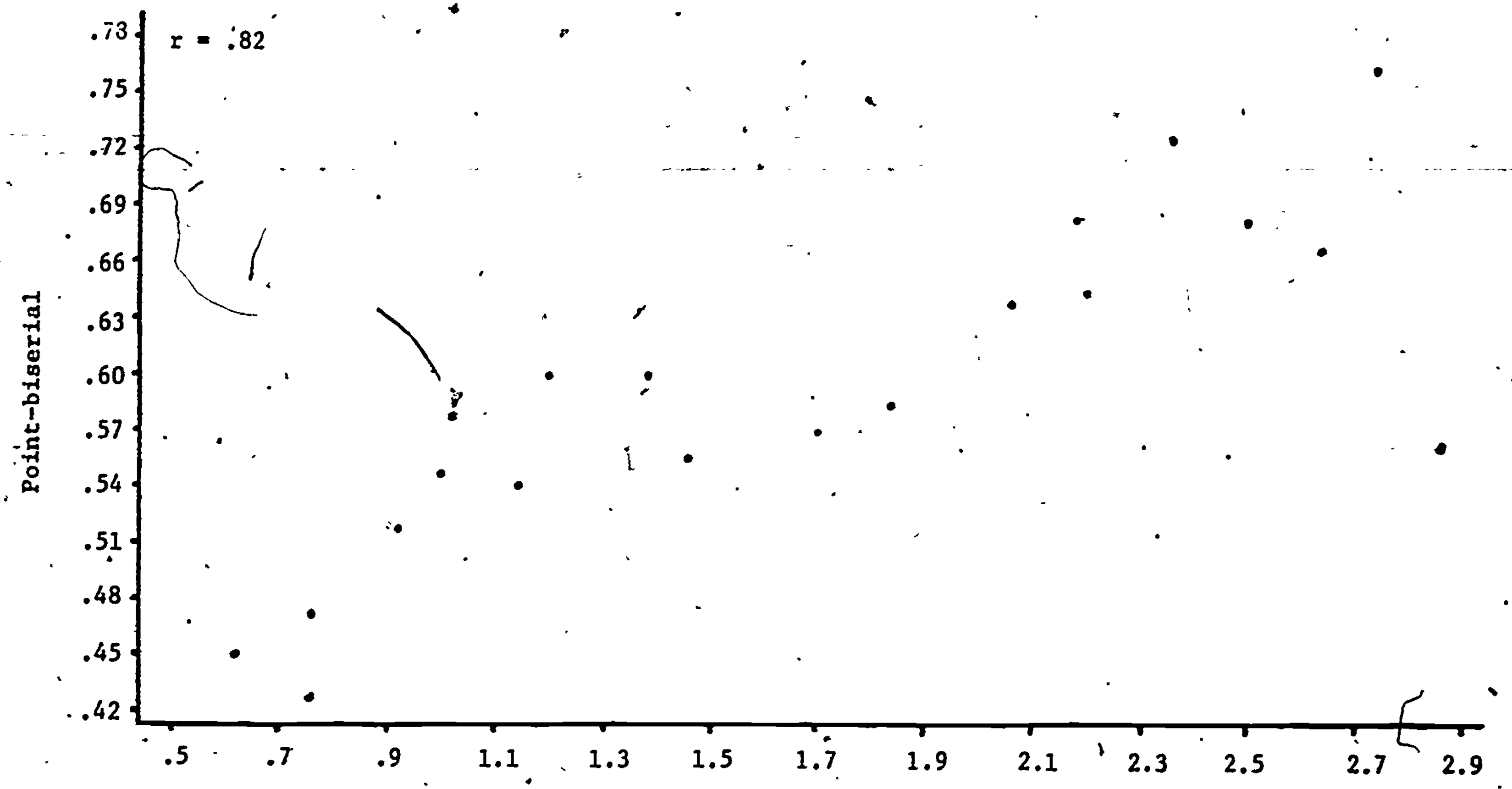
Relationship Between the Raw Score and the Inner Product of the Ability Parameters for Twenty Items Generated Using a Two-Dimensional Vector Model



Inner Product for Person Parameters

Figure 3

Relationship between the Sum of the Item Parameters
and the Point-Biserial Discrimination Index
for 20 Items Generated Using the Product Term Model



Inner Product of the Scoring Function and Item Parameters