DOCUMENT RESUME

ED 412 221 TM 027 412

AUTHOR	Palaszewski, Bo
TITLE	Multidimensional Models for Matrix-Sampled Reading Tasks.
PUB DATE	1997-05-22
NOTE	22p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Adults; Foreign Countries; International Education; International Studies; Item Response Theory; *Literacy; *Matrices; Multivariate Analysis; *Reading; *Sampling; *Scaling; Test Items
IDENTIFIERS	*International Adult Literacy Survey; Missing Data; *Multidimensional Models; Speededness (Tests); Sweden; Unidimensionality (Tests)

ABSTRACT

The International Adult Literacy Survey (IALS) was conducted in 1994 in seven European countries with the aim of producing literacy profiles to be used in individual national analyses as well as in international comparisons. It was assumed that the test items could be divided into three unidimensional scales corresponding to the aspects "Prose" (ability to read and comprehend text), "Doc" (document, comprehension of texts with charts and figures), and "Num" (numeracy, the ability to make calculations from text information). The assumption of unidimensionality is fundamental in the item response theory (IRT) estimation technique used to rescale the results of test subjects into a common scale. This paper explores this assumption for IALS results using models for multivariate data. It is shown that it is possible to identify a general dimension, measuring a general literacy ability, as well as more specific dimensions. Data from the Swedish IALS for 2,919 subjects were analyzed with a structural modeling approach that coped with the structural missing data introduced by the balanced incomplete block design. The final nested-factor model fitted on the passage scores obtained a reasonably good fit. The general literacy dimension in the Swedish test, defined as "Lit," did explain most of the variability of the data. The general dimension did not differ across gender, but the numerical dimensions exhibited different patterns for males and females. The analysis supports the existence of several latent dimensions in addition to the three originally postulated. The existence of block factors, probably related to speededness, is not surprising. Results indicate discernible passage effects that need not be a serious reliability problem, but that do merit further study. Two appendixes present standardized estimates and an analysis based on item levels. (Contains 1 figure, 8 tables, and 13 references.) (SLD)

Multidimensional Models for Matrix-Sampled Reading Tasks

Bo Palaszewski

University of Göteborg

Address:

Bo Palaszewski

E-mail: Bo.Palaszewski@ped.gu.se

Department of Education and Educational Research

Göteborg University

P.O. Box 1010

S-431 26 Mölndal

Sweden

May 22, 1997

Introduction

The International Adult Literacy Survey (IALS) was conducted in 1994 with the aim of producing literacy profiles to be used in individual national analyses as well as in international comparisons. The study is a large scale comparative assessment of adult skill. The first round of this survey was conducted in seven OECD countries, USA, Canada, Germany, Poland, Switzerland, Netherlands and Sweden. The Swedish subset of the IALS data file is reanalysed in this paper. The IALS study aimed at assessing three aspects of adult literacy: prose literacy, document literacy and quantitative literacy. The sampled population in each country was aged 16-64 years. In the analysis, made by Educational Testing Service (ETS), the assumptions were that the test items could be divided into three unidimensional scales corresponding to the three aspects *Prose*, *Doc* and *Num*. The *Prose* dimension refers to the ability to read and understand texts of story type or descriptions or explanations of something. The *Document* (*Doc*) dimension refers to the abilities related to understanding texts consisting of structured collections of information in the form of charts, graphs, lists and tables. They typically involve locating information or following directions, rather then reading of connected texts. *Numeracy* (*Num*) is a dimension referring to the ability to make calculations with assistance of information given in texts or documents of different kinds.

The assumption of uni-dimensionality is fundamental in the IRT-based estimation technique used to rescale the result of the test subjects into a common scale. This assumption will be scrutinised in this paper by using models for multivariate data. We will show that it is possible to identify a general dimension, measuring a general literacy ability, plus some more specific dimensions. Some of these more specific dimensions correspond to the postulated dimensions, although the support seems to be quite weak.

The more specific dimensions could also be due to context specific contributions within each passage of items, but also as a result from the design of the study. These design effects could be caused by e.g. speededness in the test. There seems to be some support for an end of text factor, probably a factor related to speededness in the test.

Evidence are also in support for the hypothesis that there exists gender differences in some dimensions.

Constructs and measurements

The Swedish subset of the IALS data was analysed in this work assuming the existence of a general dimension, *Literacy* (*Lit*), together with more specific dimensions *Num* and *Doc*. The assumption of a hierarchical model with a general dimension at the highest level and with more specific ability dimensions at the lower levels is in accordance with results from theoretical psychometrical research results, concerning the structure of the cognitive abilities. Some results (Caroll, 1993; Gustafsson, 1988) show that there are three levels of generality of cognitive abilities. The highest level is a general intellectual ability, consisting of an analytical, non-verbal, reasoning dimension. The second highest level consists of dimensions like verbal and knowledge abilities, spatial ability and memory ability. At the third and lowest level there is a large set of specific abilities, corresponding to the "primary mental abilities" as defined by Thurstone (1938).

Thus, it is assumed that the data analyses could be based on a nested-factor model. A nested-factor model is a model where all latent variables are directly related to the manifest variables, but the degree of generality differ, in such a way that there is a general factor which is related to all manifest variables, along with other factors related to subsets of the manifest variables. Gustafsson & Balke (1993) and Gustafsson (1995) defined models with different degrees of generality, e.g. a general factor, and latent variables which are of a more specific kind related to subsets of the manifest variables, as nested-factor models. The results in Gustafsson (1995) gave some limited support for using a hierarchical variance decomposition, when re-analysing the reading achievement data in the IEA study. The results also suggested that it is reasonable to consider a model with a general reading achievement factor since some of the *Text* and *Doc* factors were highly correlated. The use of a (common) general factor could be motivated from the existence of highly correlated factors due to the common requirement of some level of literacy ability to be able to solve items representing any domain.

The three scales assumed in the analysis performed by ETS could be expected to co-vary to a large degree since they all require general reading skills, but still there should be some unique variance associated with each scale if the validity of the test is sufficiently high. The unique variability associated with each scale is due to the fact that subjects' reading performance is to some extent a function of the subjects prior knowledge and experience with the type of content matter involved in an exercise. What could be questioned is whether the observed relationship between the constructs correspond to different literacy factors. Other questions of interest are whether there exists more dimensions representing text effects, gender effects, etc..

Passage scores were used as the base for the analysis since preliminary analysis results gave support for the existence of block factors related to the common text for a set of items, i.e. a passage of items. A passage generally consisted of a mix of item types (representing Literacy, Document and Numeracy domains), with emphasis on one domain. We will also search for the existence of other block effects. The possible existence of block factors corresponding to end of test factors, i.e. factors accounting for the speededness of the text, will also be considered in the analysis.

The data will be analysed using a structural modelling approach which is able to cope with the structural missing data structure introduced by the balanced incomplete block design (BIB) used in this study. This BIB design was adopted in order to obtain maximum information value for the available test time.

Data structure

The Swedish IALS data consists of a total of 2,919 subjects. Each group was administered one out of seven booklets according to Table 1. The number of subjects in booklet 1 is artificially high, since non-responding subjects was included in this group.

Table 1. Number of subjects in the 7 booklets

Booklet	n
1	527
2	348
3	415
4	436
5	521
6	365
7	307

It was not possible to administer every test item to every test person due to the very large number of test items in the study. To ensure broad coverage of the measured literacy domains and not to overburden individual assessment takers, the test battery was divided into partially overlapping subsets, blocks, which were given to subsets of the subjects. The blocks were distributed across booklets according to Table 2. Each block of items was devised to demand roughly the same amount of time, and to include items of all difficulty, domains and content. The average respondent was assumed to use 30 minutes to answer background questions and an additional 60 minutes to answer the given block of items.

The balanced incomplete block (BIB) design controls for possible effects due to *different content* in the 7 blocks. Thus the design is balanced with respect to the number of times each block occur together with other blocks, and also at which location they occur since each block occurs once at every possible location inside booklets. Effects caused by *the order* in which blocks are administered to test subjects, e.g. that the exposure of a block of items to a subject may influence the motivation to answer the items in the following blocks, i.e. a carry-over effect, is not controlled for in this balanced incomplete block design.

Table 2. Distribution of blocks within booklets

Booklet	Block
1	1,2,4
2	2,3,5
3	3,4,6
4	4,5,7
5	5,6,1
6	6,7,2
7	7,1,3

The distribution of items over the different dimensions in the Swedish IALS are given in Table 3.

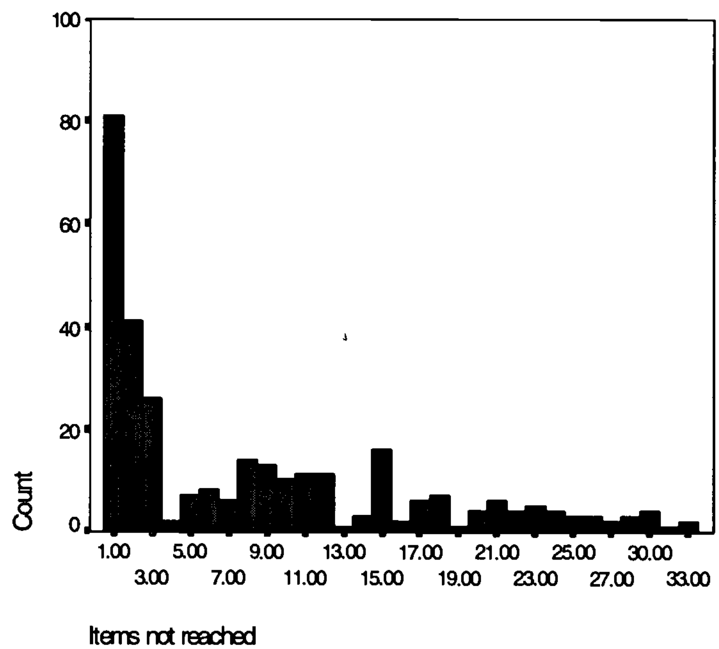
Table 3. Distribution of items over dimensions

Dimension	# items
Lit	37
Doc	39
Num	31
Sum	107

The scoring system used in this analysis scored omitted response as an incorrect response. The motivation for this procedure is that omitted responses could be due to lack of capability of giving correct answers. Since it is most common that test takers answer items in order of appearance, it could be assumed that omitted responses located amongst attained responses (correct or not), should have resulted in an incorrect response if the test taker would have given an answer.

Another possibility to interpret omitted responses is that the test might be speeded, i.e. that test takers are so slow to give responses that they do not complete all the presented items in the test because of exceeding the time limits imposed by the administration of the test. Speededness is mostly observed at the end of the test since, as previously discussed, test takers tend to answer items in a sequential order. Scoring such omitted responses as an incorrect response seems to be crude. But, as will be shown, some of this speededness could be taken into account by introducing an end of test factor in to the model.

Figure 1. Frequency of items not reached



If omitted responses after the last responded item in the booklet are counted, we get a good preliminary picture of how serious the speededness problem are. As can be seen in Fig. 1, a large proportion of test takers do not respond to items presented at the end of the booklet.

Data analysis

The analysis of the IALS data was performed by combining the 7 subsamples (books), each with a somewhat different set of variables, and using techniques developed by Muthén, Kaplan, & Hollis (1987) and Allison (1977), for analysing structural equation models with missing data structures. They proposed a maximum likelihood (ML) method for analysing data with a moderately large number of missing data patterns. Gustafsson & Stahl (1996) has developed the STREAMS system which, in combination with LISREL 8 (Jöreskog & Sörbom, 1993), has made the rather tedious task of implementing these data analysis techniques possible. The method assumes that the separate groups considered in the analysis (i.e. groups given different booklets) come from the same population. Further, it also assumes that the missing data patterns are “missing at random” or “missing completely at random” (Little & Rubin, 1987). In this case the missing data is caused by the use of matrix data sampling design described above, and could be assumed to satisfy the most restricted conditions of missing completely at random.

In essence, the method treats missing variables as latent variables. The model is then estimated simultaneously for all distinct subgroups, while appropriate equality constraints are imposed across subgroups. The method applies to any linear structural equation model within the span of LISREL models (Jöreskog, 1977). This ML approach produces estimates of parameters which are consistent, asymptotically efficient and normally distributed. The method also produces consistent estimates of the standard error of the parameter estimates. These last two results require data generated from a multinormal distribution to be valid.

The response variables in IALS are dichotomous and are thus clearly not distributed as multivariate normal distributed variables, which is required for the results from the used ML estimation to be true. This will e.g. give a tendency to overestimate the values of the χ^2 test statistic. The present analysis, however is based on “passage” scores formed as simple unweighted sum of scores (0 or 1) on items belonging to a passage. The data analysis is made more robust from deviations from assumptions implied by the used estimating technique by using passage scores. The necessity to keep the number of variables at a low level in the quite computer intensive analysis of matrix design data being used, is another reason to base the analysis on passage scores.

The existence of text dimensions

Another more important reason to base the analysis on passage scores is the indications of existence of local dependencies caused by the common text used for passages of items. The reason to adopt passage scores as the fundamental unit in the analysis is to take into account violations of the conditional independence assumption usually made. Since the blocks in IALS are made up of a number of groups of items (passage or testlet, see Sireci, Thissen, & Wainer (1991)), each related to a common text, the fundamental units to analyse the data should be the passage score obtained for the testlet, if the responses to some items (within the testlet) will tend to be more correlated than responses to other items (between testlets). This excess covariation is due to the local dependence introduced by the common text.

Results in appendix B show that there seems to exist within passage dependency in addition to the items dependency of literacy domains. The seven blocks of items were analysed separately, thus giving three replicates of each block. The relations between the hypothesised text factors and the items referring to that text have been assumed to be equal, while relations to the literacy factors are unrestricted. The patterns of variation give support for the assumed text factors. The text factors are, for many text passages, highly significant consistently over the replicated blocks. A significant t-value should be interpreted as a covariation between the items related to a text, which is in excess to the covariation explained by the items' relation to the general dimension and to the *Num* and *Doc* dimensions.

A variance decomposition was also performed. E.g. was the explained variation highest in the context specific dimension for the items related to the text named Le Miserable, followed in order by a factor related to a general dimension (here denoted *Lit*) and the two other postulated dimensions (*Doc* and *Num*). This pattern was not consistent over texts. For some texts the proportion explained variation was higher in *Lit* than for the context dimension. Still the exhibited patterns gives support for the existence of contextual effects.

Passage level analysis of the complete Swedish data set

Performed in this paper is thus a data analysis of structural equation models using the 35 passage scores from the 7 different blocks available as manifest variables. The data has been analysed on the assumption that there exists a common dimension for all passages(*Lit*). The *Num*-dimension was related to all passages including at least one item classified as measuring numerical abilities, and the *Doc*-dimension was related to passages including at least one item measuring the abilities of interpreting texts of documentary types.

Every addition of a latent variables to the model may be tested for significance by taking its χ^2 value, subtracting it from the previously fitted models χ^2 value. This χ^2 difference should be large in comparison to the difference in degrees of freedom between the models. In order to obtain the standard χ^2 goodness-of-fit test for this kind of models for incomplete data it is necessary first to estimate the so called "saturated" model, or the "H1-model" (Muthén, Kaplan, & Hollis, 1987). This is a model which fits all variances and covariances for a set of manifest variables. The obtained χ^2 value and its degrees of freedom for this H1-model has been subtracted from the obtained test values in order to obtain the χ^2 statistic and its degrees of freedom in the following analysis

The results in Table 4 show that a model with *Lit* as a general factor and *Num* and *Doc* as specific factors fit data rather well. The value of the chi-square statistic is significantly reduced as the *Num* and *Doc* factors are introduced into the model.

Table 4. Test values for all subjects

Model	χ^2	d.f.
Lit	2609.6	560
Lit+Num	2288.0	540
Lit+Num+Doc	2027.0	520

Gender differences

The data was also analysed to see if it was possible to detect any gender differences. The gender variable was used to divide the data set into two groups, with the following distribution across gender:

Table 5. Number of subjects in the 7 booklets, divided by gender.

Book	number of females	number of males
1	299	274
2	176	180
3	207	218
4	248	200
5	258	271
6	169	205
7	187	146

A multi-group analysis was then used to perform analysis on the data. Models were estimated with parameters restricted to be equal across gender groups, as well as models with parameters free of such restrictions. Models with restrictions on parameters between genders are denoted by including the factor name within parentheses and factors without restrictions without parentheses.

The best fitting model in Table 6 is the one that has common parameter estimates for males and females in the *Lit* dimension, but unconstrained parameter estimates between the genders in the *Num* dimension. The result also shows that the effect of the *Doc* dimension is no longer significant. This could be interpreted as evidence for a gender specific mean difference in the *Doc* dimension. The previously significant factor estimate for this dimension could then be explained as an aggregation effect, and not a real factor effect. One plausible explanation to this gender specific mean difference is given by Rosén (1995). She found that in the IEA data, the gender difference in a *Doc* factor was mainly due to context specific contributions.

Table 6. Test values for multigroup analysis for gender.

Model	χ^2	d.f.
(Lit)	3828.1	1225
Lit	3765.6	1191
(Lit)+(Num)	3506.8	1204
(Lit)+Num	3455.2	1186
(Lit)+Num+(Doc)	3452.3	1185
(Lit)+Num+Doc	Non-convergence	

The existence of block effects

To examine for the existence of possible block effects each booklet was analysed separately. For every booklet a model were fitted, including the *Lit* factor as a general dimension and a block factor accounting for a block of "passage" items. The parameter estimates within a block were restricted to be equal. The notations for this latter kind of block estimates are (Block*n*), where Block*n* represents all passages within the *n*th block. The parentheses indicate that all parameter estimates within Block*n* are restricted to be equal. These block factors were introduced one at a time in the model, and the procedure was repeated for each block within a booklet one at a time. Significant block effects were detected. These significant block effects could possibly be interpreted as speededness factors. The existence of a speededness dimension indicates to what extent the items late in the test were adequately responded or not. The predominant position for significant block effects was for the last administered block for each booklet. Six significant block effects were located at the end of each booklet. The remaining significant block effects were three for blocks in position 1 and three in position 2. Thus there seems to be some limited support for the existence of a speededness factor.

The final model

In the full model, where data from all booklets are combined in a common missing data model, existing block effects for block factors (and speededness factors) are aliased to some extent, since data with strong block effects are weighted together with data from blocks with smaller or no block effects. Still, the full model with block effects included, obtain a significantly better fit.

With estimates for the *Lit* dimension restricted to be equal across gender, and allowing for block effects, the results in Table 7 were obtained:

Table 7. Tests for block effects

Model	χ^2	d.f.
(Lit)	3828.1	1225
+(*Block1*)+	3807.5	1224
+(*Block2*)+	3806.3	1223
+(*Block3*)+	3725.9	1222
+(*Block4*)+	3708.9	1221
+(*Block5*)+	3700.6	1220
+(*Block6*)+	3700.5	1219
+(*Block7*)	3666.9	1218

Every addition of a block factor may be tested by taking it's χ^2 value subtracted from the χ^2 value from the previous model as well as the degrees of freedom. Block factors were found significant for all blocks except block 2 and 6. Tests were also made to find possible significant differences in block affects between gender. The results in Table 8 show that significant gender differences were detected only for Block3.

Table 8. Tests for block effects within gender.

Model	χ^2	d.f.
+*Block1*+	3666.9	1217
+*Block3*+	3638.7	1216
+*Block4*+	3638.2	1215
+*Block5*+	3637.8	1214
+*Block7*+	3637.7	1213

The best fitting model for the Swedish IALS was a model with a general *Lit* factor constrained across gender, block factors and a *Num* factor. Separate parameters were fitted for gender in *Block3* and *Num* only. According to our previously used notation, the model could be written as:

(*Lit*)+ (*Block1*)+*Block3*+ (*Block4*)+ (*Block5*)+ (*Block7*) +*Num*

The χ^2 test value for this model was 3,200 with 1,179 degrees of freedom. Parameter estimates for this model are found in appendix A. Only males obtained significant parameter estimates in the exhibited patterns in the *Num* dimension. They are mainly positive, 13 positive of totally 20 estimates, and only 4 with a negative sign of the factor coefficient. This pattern of differentiation in the *Num* dimension, i.e. that males but not females vary in their numeracy skills, could possibly be explained by differences in occupational training, i.e. the degree of exposition to numerical job content may vary more for males than for females. Gender specific content in the test instrument, e.g. the content is more appealing to or more easily solved by some subgroups in the population, is another possible explanation.

Conclusions

The passage score analysis used in this report is motivated by the fact that there is a local dependency between the items in every passage of items. This occurs e.g. when the correctness of the response to a previously answered item affect the probability of correct answer for the following items within a passage. The final nested-factor model fitted on the passage scores obtained a reasonably good fit. The general dimension measured by the *Lit* factor does explain most of the variability in the data, though there also seems to exist a numerical dimension. The general dimension does not differ across gender, but the numerical dimensions exhibit different patterns for males and for females. Mean differences between gender were found.

Discussion

The present analysis support the hypotheses of several latent dimensions besides the originally three postulated dimensions. The existence of block factors, probably related to speededness, is not surprising. The administered test instrument is composed of many items, despite the matrix data collection design. It could be assumed that many of the test takers are relatively unfamiliar with the sort of tasks given in the test. The variability of characteristics measuring the ability to quickly solve different posed problems is probably large in the population. The assumption of local independence for the items in the test seems to be false. The results indicate that there are discernible passage effects. This does not need to be a serious problem concerning the reliability of the test, since the number of items could always be increased. More serious is how these observed within passage dependencies affects estimation results from e.g. IRT-based analyses. It is not possible to draw firm conclusions concerning how serious this violation is, in the presented analysis. There is an obvious need of further work to illuminate this problem.

The possibility to build complex statistical models to data collected with structural missing data, as in matrix data collection designs, give the researcher new possibilities to analyse data. The theoretical framework used in extending structural missing data models to structural equation modelling is not new, but lack of good implementations in software made fitting of models of some degree of complexity prohibiting. This obstacle is now removed. But still there is a need for theoretical as well as software development within this area. The large array of the standard test statistics and diagnostic tools are not easily available when modelling incomplete data. Thus the goodness-of-fit test is more difficult to compute when modelling incomplete data, than when data is complete, and the descriptive fit statistics and the modification indices can usually not be obtained when incomplete data models are used.

Appendices

Appendix A

Standardised estimates: Bold characters denote significance parameter estimates at 5% level t-test. Block (Blo*n*) estimates are not tested by t-tests, instead see previous χ^2 tests. Parameter estimates are standardised within-group estimates (see Jöreskog & Sörbom, 1989, pp. 38-40).

Male	B1P1	=	+0.70*Lit	+0.15*Blo1		+0.69*B1P1&
Female	B1P1	=	+0.70*Lit	+0.15*Blo1		+0.69*B1P1&
Male	B1P2	=	**+0.74*Lit**	+0.14*Blo1	**-0.22*Num**	+0.62*B1P2&
Female	B1P2	=	**+0.74*Lit**	+0.14*Blo1	-0.04*Num	+0.63*B1P2&
Male	B1P3	=	**+0.71*Lit**	+0.15*Blo1	**+0.12*Num**	+0.68*B1P3&
Female	B1P3	=	**+0.66*Lit**	+0.14*Blo1	-0.39*Num	+0.63*B1P3&
Male	B1P4	=	**+0.69*Lit**	+0.15*Blo1		+0.71*B1P4&
Female	B1P4	=	**+0.69*Lit**	+0.15*Blo1		+0.71*B1P4&
Male	B1P5	=	**+0.77*Lit**	+0.15*Blo1	+0.24*Num	+0.57*B1P5&
Female	B1P5	=	**+0.79*Lit**	+0.16*Blo1	+0.05*Num	+0.59*B1P5&
Male	B2P1	=	**+0.76*Lit**			+0.65*B2P1&
Female	B2P1	=	**+0.76*Lit**			+0.65*B2P1&
Male	B2P2	=	**+0.59*Lit**		+0.04*Num	+0.81*B2P2&
Female	B2P2	=	**+0.59*Lit**		-0.19*Num	+0.80*B2P2&
Male	B2P3	=	**+0.56*Lit**			+0.83*B2P3&
Female	B2P3	=	**+0.56*Lit**			+0.83*B2P3&
Male	B2P4	=	**+0.76*Lit**		**+0.10*Num**	+0.64*B2P4&
Female	B2P4	=	**+0.76*Lit**		-0.07*Num	+0.64*B2P4&
Male	B2P5	=	**+0.78*Lit**			+0.62*B2P5&
Female	B2P5	=	**+0.78*Lit**			+0.62*B2P5&
Male	B3P1	=	**+0.74*Lit**	+0.12*Blo3	+0.10*Num	+0.65*B3P1&
Female	B3P1	=	**+0.70*Lit**	+0.29*Blo3	-0.24*Num	+0.62*B3P1&
Male	B3P2	=	**+0.67*Lit**	+0.15*Blo3	**+0.21*Num**	+0.70*B3P2&
Female	B3P2	=	**+0.64*Lit**	+0.38*Blo3	+0.09*Num	+0.66*B3P2&
Male	B3P3	=	**+0.76*Lit**	+0.11*Blo3	-0.10*Num	+0.63*B3P3&
Female	B3P3	=	**+0.74*Lit**	+0.27*Blo3	-0.09*Num	+0.61*B3P3&
Male	B3P4	=	**+0.74*Lit**	+0.12*Blo3		+0.66*B3P4&
Female	B3P4	=	**+0.71*Lit**	+0.31*Blo3		+0.63*B3P4&
Male	B3P5	=	**+0.76*Lit**	+0.12*Blo3		+0.64*B3P5&
Female	B3P5	=	**+0.73*Lit**	+0.30*Blo3		+0.62*B3P5&

Male	B4P1	=	**+0.76*Lit**	+0.21*Blo4	**+0.23*Num**	+0.57*B4P1&
Female	B4P1	=	**+0.76*Lit**	+0.21*Blo4	+0.25*Num	+0.57*B4P1&
Male	B4P2	=	**+0.68*Lit**	+0.16*Blo4	**-0.33*Num**	+0.64*B4P2&
Female	B4P2	=	**+0.67*Lit**	+0.15*Blo4	-0.37*Num	+0.63*B4P2&
Male	B4P3	=	**+0.74*Lit**	+0.19*Blo4		+0.65*B4P3&
Female	B4P3	=	**+0.74*Lit**	+0.19*Blo4		+0.65*B4P3&
Male	B4P4	=	**+0.66*Lit**	+0.20*Blo4	**-0.10*Num**	+0.72*B4P4&
Female	B4P4	=	**+0.66*Lit**	+0.20*Blo4	-0.13*Num	+0.71*B4P4&
Male	B4P5	=	**+0.75*Lit**	+0.10*Blo4	**+0.22*Num**	+0.62*B4P5&
Female	B4P5	=	**+0.74*Lit**	+0.10*Blo4	+0.25*Num	+0.61*B4P5&
Male	B5P1	=	**+0.45*Lit**	+0.39*Blo5		+0.80*B5P1&
Female	B5P1	=	**+0.45*Lit**	+0.39*Blo5		+0.80*B5P1&
Male	B5P2	=	**+0.70*Lit**	+0.10*Blo5	**-0.19*Num**	+0.68*B5P2&
Female	B5P2	=	**+0.71*Lit**	+0.11*Blo5	-0.03*Num	+0.69*B5P2&
Male	B5P3	=	**+0.73*Lit**	+0.12*Blo5		+0.68*B5P3&
Female	B5P3	=	**+0.73*Lit**	+0.12*Blo5		+0.68*B5P3&
Male	B5P4	=	**+0.64*Lit**	+0.08*Blo5	**+0.20*Num**	+0.74*B5P4&
Female	B5P4	=	**+0.63*Lit**	+0.08*Blo5	-0.25*Num	+0.73*B5P4&
Male	B5P5	=	**+0.67*Lit**	+0.11*Blo5	**+0.16*Num**	+0.71*B5P5&
Female	B5P5	=	**+0.63*Lit**	+0.10*Blo5	-0.39*Num	+0.67*B5P5&
Male	B6P1	=	**+0.63*Lit**		**+0.15*Num**	+0.76*B6P1&
Female	B6P1	=	**+0.63*Lit**		+0.08*Num	+0.77*B6P1&
Male	B6P2	=	**+0.68*Lit**		**+0.16*Num**	+0.71*B6P2&
Female	B6P2	=	**+0.69*Lit**		-0.02*Num	+0.72*B6P2&
Male	B6P3	=	**+0.66*Lit**			+0.76*B6P3&
Female	B6P3	=	**+0.66*Lit**			+0.76*B6P3&
Male	B6P4	=	**+0.67*Lit**		**+0.21*Num**	+0.71*B6P4&
Female	B6P4	=	**+0.65*Lit**		-0.30*Num	+0.70*B6P4&
Male	B6P5	=	**+0.60*Lit**			+0.80*B6P5&
Female	B6P5	=	**+0.60*Lit**			+0.80*B6P5&
Male	B7P1	=	**+0.64*Lit**	**+0.29*Blo7**	+0.21*Num	+0.68*B7P1&
Female	B7P1	=	**+0.66*Lit**	+0.28*Blo7	-0.16*Num	+0.69*B7P1&
Male	B7P2	=	**+0.66*Lit**	+0.20*Blo7	+0.10*Num	+0.71*B7P2&
Female	B7P2	=	**+0.66*Lit**	+0.20*Blo7	-0.18*Num	+0.71*B7P2&
Male	B7P3	=	**+0.71*Lit**	+0.18*Blo7		+0.68*B7P3&
Female	B7P3	=	**+0.71*Lit**	+0.18*Blo7		+0.68*B7P3&
Male	B7P4	=	**+0.71*Lit**	+0.17*Blo7		+0.69*B7P4&
Female	B7P4	=	**+0.71*Lit**	+0.17*Blo7		+0.69*B7P4&
Male	B7P5	=	**+0.72*Lit**	+0.23*Blo7		+0.65*B7P5&
Female	B7P5	=	**+0.72*Lit**	+0.22*Blo7		+0.65*B7P5&

Appendix B

An analysis based on item level data. Models were fitted with one general literacy factor, two more specific literacy factors (*Num* and *Doc*) plus block factors corresponding to items related to a common text.

Table B1. t-values for hypotetized text and literacy factors

	BOOKLET 1	BOOKLET 2	BOOKLET 3	BOOKLET 4	BOOKLET 5	BOOKLET 6	BOOKLET 7
Block 1							
Passage 1	2.24				2.10		3.18
Passage 2	1.68				1.64		6.78
Passage 3	4.98				4.01		5.60
Passage 4	0.10				-1.49		-3.26
Passage 5	1.76				-0.35		-0.09
Num	2.04				2.08		-1.59
Doc	1.72				-3.84		0.27
Block 2							
Passage 1	5.02	7.21				6.59	
Passage 2	0.84	-0.38				0.06	
Passage 3	3.06	0.58				-0.88	
Passage 4	0.79	2.33				0.50	
Passage 5	8.36	7.65				4.00	
Num	0.16	0.32				1.10	
Doc	-0.04	0.20				0.11	
Block 3							
Passage 1		3.29	3.99				3.93
Passage 2		3.37	4.56				1.02
Passage 3		3.21	2.17				3.83
Passage 4		-0.44	-0.98				4.81
Passage 5		1.07	1.45				4.87
Num		3.02	1.42				-1.31
Doc		-0.41	0.59				-0.43
Block 4							
Passage 1	2.90		3.51	2.58			
Passage 2	11.37		8.28	6.60			
Passage 3	3.62		1.79	3.86			
Passage 4	4.50		3.38	1.07			
Passage 5	8.31		8.45	10.33			
Num	5.78		1.46	2.93			
Doc	-0.08		4.56	2.28			

	BOOKLET 1	BOOKLET 2	BOOKLET 3	BOOKLET 4	BOOKLET 5	BOOKLET 6	BOOKLET 7
Block 5							
Passage 1		9.31		7.44	3.34		
Passage 2		5.87		7.18	2.89		
Passage 3		4.34		3.55	2.11		
Passage 4		7.62		11.14	1.10		
Passage 5		7.42		7.92	1.45		
Num		-0.51		2.43	0.87		
Doc		1.34		-0.06	-0.60		
Block 6							
Passage 1			6.66		7.28	4.94	
Passage 2			3.69		5.29	3.08	
Passage 3			8.25		5.28	0.66	
Passage 4			7.90		7.37	-0.13	
Passage 5			0.59		-1.40	3.62	
Num			0.96		1.56	-4.68	
Doc			-2.0		-1.53	5.65	
Block 7							
Passage 1				8.68		3.87	5.16
Passage 2				4.93		0.93	1.32
Passage 3				2.15		-0.33	3.70
Passage 4				4.45		0.56	3.58
Passage 5				5.61		0.89	5.07
Num				0.81		-0.29	0.47
Doc				0.49		-0.33	-1.34

References

Allison, P. D. (1977). Estimation of linear models with incomplete data. In C. Clogg (Ed.), *Sociological methodology 1987* (pp. 71-103). San Francisco: Jossey Bass.

Caroll, J. B. (1993). *Human cognitive abilities. A survey of factor-analytic studies.* Cambridge: Cambridge University Press.

Gustafsson, J. E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 35-71). Hillsdale, New Jersey: Lawrence Erlbaum.

Gustafsson, J. E. (1995). *Alternative Hierarchical Models of Reading Achievement.* Paper presented at the Annual Meeting of the American Educational Research Association., San Francisco.

Gustafsson, J. E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioural Research, 28*, 407-434.

Gustafsson, J. E., & Stahl, P. A. (1996). *STREAMS user's guide, version 1.6*. Mölndal, Sweden: MultivariateWare.

Jöreskog, K. G. (1977). Structural Equation Models in the Social Sciences: Specification, Estimation and Testing. In P. R. Krishnainh (Ed.), *Application of Statistics* . Amsterdam: North Holland.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 reference guide*. Chicago: Scientific Software International.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis of missing data*. New York: Wiley.

Muthén, B., Kaplan, D., & Hollis, M. (1987). On structural equation modelling with data that are not missing completely at random. *Psychometrika, 52*, 431-462.

Rosén, M. (1995). *Gender differences in performance on documents as a function of cultural factors*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Sireci, S. G., Thissen, D., & Wainer, H. (1991). On the reliability of testlet-based tests. *Journal of Educational Measurement, 28*, 237-247.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

AREA 1997

TM027412

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Multidimensional models for Matrix-Sampled Reading Tasks

Author(s): Bo Palaszewski

Corporate Source:

Publication Date: 14 March 97

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce the identified document, please CHECK ONE of the following options and sign the release below.

[X] **Check here**
Permitting microfiche (4"x 6" film), paper copy, electronic, and optical media reproduction

Sample sticker to be affixed to document

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 1

Sample sticker to be affixed to document

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Level 2

[ ] **or here**
Permitting reproduction in other than paper copy.

Sign Here, Please

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Signature: [signature]	Position:
Printed Name: Bo Palaszewski	Organization: Göteborg University
Address: Göteborg University Box 1010 S-431 36 Mölndal SWEDEN	Telephone Number: (+46) 31 773 28 97
	Date: 12 March 1997

OVER

CUA

THE CATHOLIC UNIVERSITY OF AMERICA
Department of Education, O'Boyle Hall
Washington, DC 20064
202 319-5120

February 21, 1997

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a printed copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at http://ericae2.educ.cua.edu.

Please sign the Reproduction Release Form on the back of this letter and include it with **two copies** of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (523)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to: AERA 1997/ERIC Acquisitions
The Catholic University of America
O'Boyle Hall, Room 210
Washington, DC 20064

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://aera.net). Check it out!

Sincerely,

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

[1]If you are an AERA chair or discussant, please save this form for future use.