AN INDIRECT FACTOR ANALYSIS OF A 300 ITEM ACHIEVEMENT TEST

William H. Resha, University of Alabama

Introduction

The use of factor analytic techniques has been known since the latter half of the nineteenth century. However, due to the mathematical complexities of this technique, factor analysis on large data matrices have been limited to factoring subscales or using Q-methodology when test items exceed more than 80 to 100 variables. Such substitute methods raise many questions concerning potential inaccuracies.

Due primarily to the work of Barker and Barker in the 1970's, computer programs have been developed to handle large data matrices, using an indirect factor analysis model presented by Horst (1965). Horst's (1965) model is based on three steps:

1. The sets of variables is reduced to a limited number of totals by grouping individual variables in some unspecified order.
2. Factor analysis is applied to the matrix of totals.
3. Matrix operations are used to estimate the factor loads for individual variables.

However, until recently there has been no way to measure how practical his model was. The problems in developing a computer program appeared insurmountable, but Barker succeeded in implementing a computer program which has been refined and updated over a period of years. An article by Barker and Barker (1975) gives a complete summary of the problems and successes they have encountered.

The General Board of Examining Chaplains of the Episcopal Church has developed a 300 item instrument which is administered to seminary students during the last year of their academic training prior to ordination. The instrument, General Ordination Examination (GOE), has been used for the past three years (1975, 1976, 1977). Thus far there has been no validation nor reliability studies done on this instrument.

The purpose of this study is to provide some initial data on this instrument by factor analyzing it using Barker's indirect factor analysis programs based on Horst's (1965) model.

The study hypothesis is that the six subtests of the GOE measures six distinct areas of Bible related content and knowledge the factor analysis of this instrument will also function as a theory generator for the possible future refinement of the GOE.

Methodology

The Population:

Subjects of this study were students who were trying to become ordained Episcopal priests. The majority of the students were third year seminary students in attendance at the 11 Episcopal theological seminaries in the United States.

Population represented the vast majority of all of the students who had taken the GOE since its inception in 1975. The population consisted of 786 students. Cattell (1966) has set a criterion for number of subjects required in factor analysis at 100 subjects plus the number of variables. If using this criterion, factor analysis of the GOE would require a minimum of 400 subjects.

The GOE was administered in a standardized form to all of the students in group settings. Security of test questions was maintained at a maximum.

General Ordination Examination:

The GOE is a 300 item objective multiple choice examination. In its present form, it has six subtests covering the following areas: Old Testament (60 questions); New Testament (60 questions); Church History (60 questions); Theology (60 questions); Ethics (30 questions); and Liturgics (30 questions). There are four alternatives to each question. Subtest questions are intermixed with one another. However, there is a pattern in the way the questions are listed. The scale is set up in the following pattern, 1 2 3 4 5 6 1 2 3 4 1 2 3 4 5 6 1 2 3 4, etc.

The GOE has not been normed nor has it been examined for reliability and validity. The Educational Testing Service (ETS) served as consultants during the development of the instrument.

The GOE is basically an achievement test since its primary function is to estimate the person's present knowledge of previous course work (Anastasi, 1976). Further, the test is considered norm-referenced since its results are interpreted in relation to other students performance who took the test (Anastasi, 1976; Aiken, 1971). There was no deliberate weighting of test items in reference to item difficulty. Test items are treated as quantitative variables with the data being considered as dichotomized normal. The test items will be interpreted as discontinuous dichotomous data (Ghiselli, 1964; Edwards, 1972).

The General Board of Examining Chaplains of the Episcopal Church has not formally stated a theory for the way this test is set up. It will be assumed that the theory that is implied is that the GOE tests the student's knowledge in the area of church history, literature and vocabulary of the christian tradition. This theory is encompassed in the six subtests that have been previously mentioned.

Methodology:

The 786 subject data base was factor analyzed by an indirect method to arrive at an estimate of a conventional principle axis solution, and this solution was rotated to a varimax criterion.

The raw data were punched on IBM cards and were then transposed to dichotomous data using a scoring key developed by the church. The dichotomized scored data was in the form where 1's equaled an incorrect or omitted item and 2's represented a correct response.

A SPEC01 program (Barker and Barker, 1977) listed and numbered all of the students. This allowed for a visual check on all data cards to insure they were in proper order, had the proper number of cards per student (five cards per student) and were aligned in the proper columns. In an effort to arrive at homogeneous items for clustering into totals, a Q-analysis program (CORR98) was used on 20 subjects across the 300 items. The factor loads for the items were then estimated as in R-methodology. This 20 subjects factorization provided a rough grouping rationale for the totals used to start the iterative processes involved in the indirect factor analytic approach (CORR99).

A SPEC50 program (Barker and Barker, 1977) which generates random numbers was run to obtain the random sample of 20 subjects which was used as the seed data to be initially run with CORR98, (Barker and Barker, 1977).

Using the Eigenroots from the CORR98 a Scree Test was drawn to indicate the starting point from which programming could begin. CORR98 was then programmed to rotate six factors.

CORR99 (Barker and Barker, 1977) was then used to factor analyse the GOE's 300 items, using the six rotated factors as a starting point for clustering into totals. An item had to load ±.30 or better on only one factor before it was identified with that factor. These results were compared with the theorized factor structure as presented in the GOE.

The information measure D (Relative Uncertainty Reduction) was selected to provide an objective measure of the degree to which theorized dimensions reflect estimated factor structure of the data set. In an ideal solution, all entries in the matrix of factors and item subsets would appear in the diagonal. Such a solution would indicate complete agreement of items subsets (totals) with actual factor structure. Frequently, however, certain items are found to load inappropriately. Items which fail to load as expected on factors appear as false negatives. Those which load into factors contrary to expectations appear as false positives. The D measure expresses the relationship between rows (subsets of items) and columns (factors). Use of this statistics also permits comparison between theories of degree of agreement between a priori item subsets and actual factor structure.

Results

CORR98 identified six factors as the most parsimonious initial set to use. The decision was made that a factor had to have at least three or more items loading on it before it would be considered for clustering into a total.

CORR99 was programmed for 20 iterations using six factors. The computer reached convergence at 12 iterations for maximum factor alignment.

The indirect factor analysis method identified three general factors which make up the greater part of the GOE and which contribute the most to the D value (.87). These three general factors are labeled as follows: Factor I = General Bible Content; Factor V = Historical Theology; Factor VI = Contemporary Theology.

In addition, there were three other factors that had some item loadings. The number of items loading on these factors appear to be inconsequential when the total number of items (300) on the GOE is taken into consideration.

The data indicates most impressive results, obtaining a D = .87. This D value indicates that item subsets identified by the Horst's (1965) indirect factor analysis method reflect actual factor structure. The higher D value is reflected in the lower number of false negatives (98).

There were two items which loaded at ±.30 or better on two factors (item 205 loaded on factors V and VI; item 232 loaded on factors I and V).

CORR99 was also run using scoring on subtests as hypothesized by the Episcopal

Church. Six totals consisting of 60 items each for the first 4 totals and 30 items each for the other two were used, according to the scales. The theory hypothesized by the Episcopal Church does not appear to have been substantiated by the data, which indicates a D measure of .32 which represents a weak agreement between the theorized item clusters and actual factor structure. The partial nature of the theory is reflected by the large number (206) of items, expressed as false negatives which failed to load appropriately on factors.

Conclusion

The 300 item GOE as developed by the Episcopal Church appears to have three factors which measures general knowledge in Historical Theology, Contemporary Theology and Bible Content. A total of 135 items loaded on one of six new factor structures. Of these 135 items, 84% of them (114) loaded on one of the three main factors.

The results from Horst's (1965) indirect method of factor analysis appears to be quite impressive and persuasive in reference to the possibility of either reducing the GOE down to a smaller test or revising the examination, improving the questions so more of the items load on the identified factors. Out of the 300 items, 45% loaded on one of six new factors and 38% loaded on one of three major factors. This indicates that well over 50% of the present items are not loading on any factors (subtests) with which they were initially identified. It is suggested, therefore, that until further revision is made of the GOE, caution needs to be stressed if it is used in a decision making capacity.

Results of this study support earlier studies previously cited by Barker and Barker and Hamlett (1976) in demonstrating usefulness of the indirect method as an appropriate technique in evaluating theories regarding factor structure of large data sets.

It is further noted that the overall computer cost of using the Barker and Barker programs remain at a minimum when the number of variables is considered. For example, using six factors, 300 variables, 786 subjects (five cards per subject) and 12 iterations, it took a total of in/out time of 8 1/2 minutes at a cost of $110.45. This further substantiates previous Barker articles cited in reference to the economy of Horst's (1965) model as programmed by CORR98 and CORR99 (Barker and Barker, 1977).

The present study demonstrates the usefulness of performing factor analysis on psychological tests as a first step in the overall process of test validation.

Two further conclusions appear appropriate from this study. First, the indirect factor analytic method as presented has proven its flexibility. It is noted that all of the computer runs were made during normal operating hours and it was not necessary to shut down other projects or handle the Barker program in any special manner by the computer center.

Finally, regardless of the significance of the reported findings, it is significant that someone with relatively limited training in statistics could be successful in completing such a project. The relative simplicity of the Barker programs will allow future researchers in the behavioral sciences this additional tool in their empirical investigations.

References

Aiken, L. R., Jr. Psychological and Educational Testing. Boston, Mass.: Allyn and Bacon, Inc., 1971.

Anastasi, A. Psychological Testing (4th Ed.). New York: Macmillan, 1976.

Barker, H. R., Barker, B. M. and Carlton, B. B. Accuracy Of An Indirect Factor Analytic Solution As A Function Of The Method Of Assigning Variables To Totals. Paper presented at American Statistical Association, Atlanta, 1975A.

Barker, H. R. and Barker, B. M. Summary of research on a novel, indirect factor analytic solution. Paper presented at American Statistical Association, Atlanta, 1975B.

Barker, H. R. and Barker, B. M. Behavioral Sciences Statistics Program Library. (2nd Ed.). University of Alabama, 1977.

Barker, H. R. and Barker, B. M. An indirect factor method for testing the dimensionality of large data sets. Paper presented at Southeastern Psychological Association, New.Orleans, 1976.

Barker, H. R., Fowler, R. D., and Peterson L. P. Factor Analytic Structure of the Short-Form MMPI Items. Journal of Clinical Psychology, 1971, 27, 228-233.

Cattell, R. B. Extracting the correct number of factors in factor analysis. Educational and Psychological Measurement, 1958, 18, 791-838.

Edwards, A. L. Experimental Design in Psychological Research (4th Ed.).

New York: Holt, Rinehart and Winston, Inc., 1972.

Ghiselli, E. E. Theory of Psychological Measurement. New York: McGraw-Hill, 1964.

Horst, P. Factor Analysis of Data Matricies. New York: Holt, 1965.

Hunnicutt, B. M. and Barker, H. R. Application and evaluation of a novel and indirect factor method to a very large data matrix. Paper presented at Southeastern Psychological Association, Hollywood, May, 1974.

Sloan, H. C. A simplified procedure for estimating the factor structure of large data matrices. Unpublished doctoral dissertation. University of Alabama, 1973.

Stallings, N. A. Evaluation of an indirect method of estimating the factor structure of large data matrices. Unpublished doctoral dissertation. University of Alabama, 1973.

Table 1

Sex and Population Distribution			
Year	Male	Female	Total
1975	189	30	219
1976	227	37	264
1977	224	59	303
			786

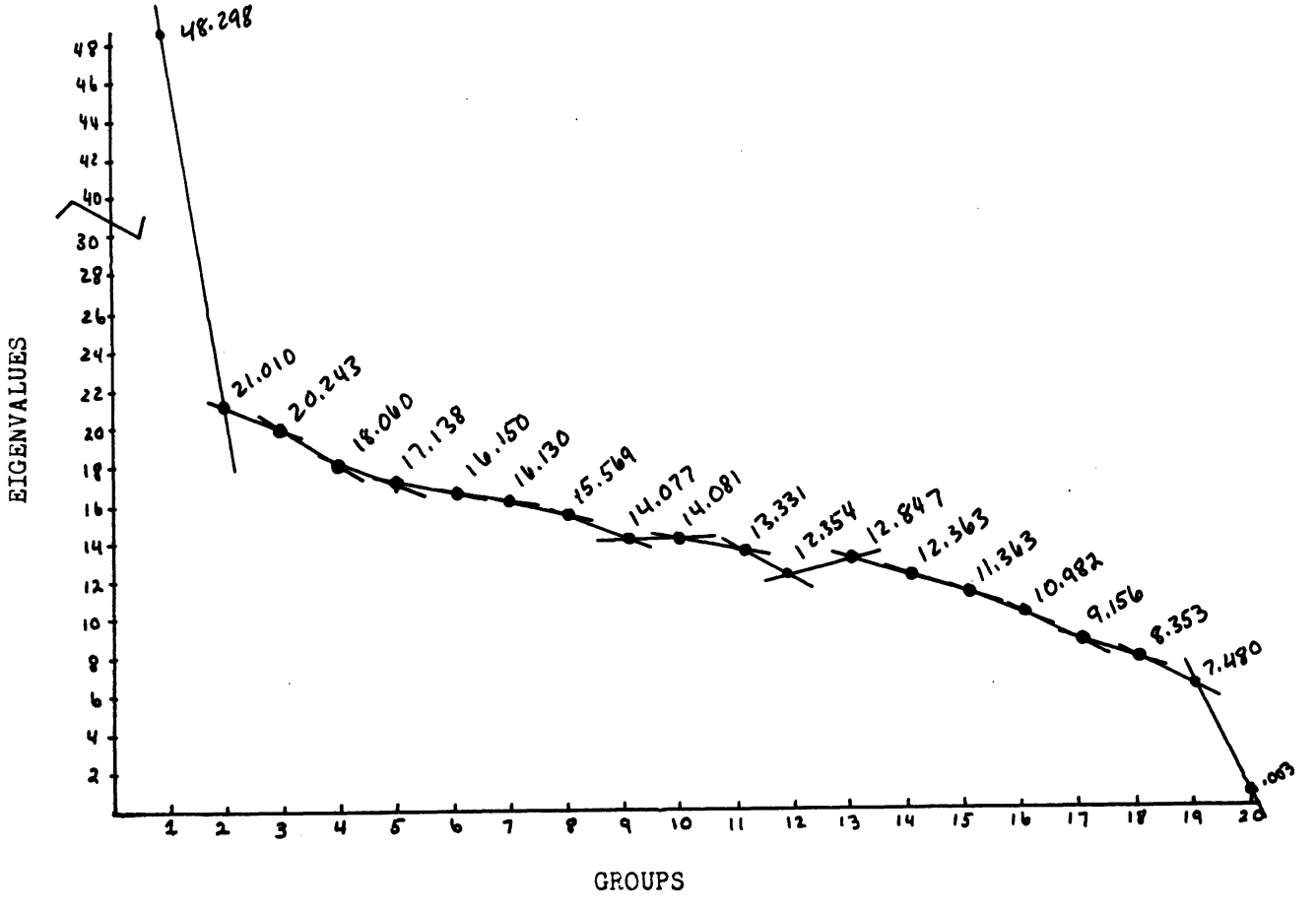

Figure 1. Scree Test on Eigenvalues of General Ordination Examination

Table 2

Year	Subject #	Total Sample
1975	51 107 121 185 203	5
1976	256 362 382 443 465	5
1977	501 508 512 517 548 551 664 680 694 724	10
		20

Table 3

Items Loading ±.3 Or Better Onto One Of Six Factors

Factor I (Gen. Bible Content)	Factor II (Specific Identification)	Factor III (Biblical Identification)	Factor IV (Church Definitions)	Factor V (Historical Theology)	Factor VI (Contemporary Theology)
1	187	12	53	9	56	284	5
8	188	21	57	105	129	287	10
11	192	44	103	133	180	289	35
17	197	72	119	139	186	293	14
27	198	256	135	144	196	294	38
31	212		141	175	204	296	45
32	217		286	223	206	298	50
37	218			255	211	299	75
48	237			275	216	300	83
51	288				229		89
61	242				230		96
97	246				231		99
98	257				235		100
102	288				248		110
107					249		120
108					253		140
111					258		170
117					263		179
118					264		184
122					265		185
127					266		189
131					268		200
138					270		210
147					271		220
148					272		221
152					273		227
157					276		240
161					277		245
172					278		260
181					281		285

Table 4

Association Between Item Subsets and Varimax Factors
(Actual Data Results)

	I	II	III	IV	V	VI	False -	Sum
1	41	0	0	0	0	0	0	41
2	0	5	0	0	0	0	0	5
3	0	0	7	0	0	0	0	7
4	0	0	0	7	0	0	0	7
5	0	0	0	0	29	0	0	29
6	0	0	0	0	0	27	0	27
False +								
	2	0	0	0	8	1	98	109
Sum	43	5	7	7	37	28	98	225

HX = 2.135704 HY = 2.214495 HXY = 2.417882 D = .87258

Table 5

Association Between Variable Subsets and Varimax Factors
(Church Hypothesized Data Results)

	I	II	III	IV	V	VI	False -	Sum
1	26	0	0	0	0	0	34	60
2	0	21	0	0	0	0	39	60
3	0	0	17	0	0	0	43	60
4	0	0	0	12	0	0	48	60
5	0	0	0	0	8	0	22	30
6	0	0	0	0	0	10	20	30
False +								
	0	0	0	0	0	0	0	0
Sum	26	21	17	12	8	10	206	300

HX = 2.521930 HY = 1.670171 HXY = 3.398042 D= .31486