DOCUMENT RESUME

ED 357 060 TM 019 786

AUTHOR	Kaplan, David
TITLE	The Impact of BIB-Spiralling Induced Missing Data Patterns on Goodness-of-Fit Tests in Factor Analysis. Occasional Paper OP93-1.
INSTITUTION	National Center on Adult Literacy, Philadelphia, PA.
SPONS AGENCY	Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE	Jan 93
NOTE	18p.; Later version of a paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Chi Square; Computer Simulation; Correlation; *Factor Analysis; *Goodness of Fit; Mathematical Models; Matrices; Monte Carlo Methods; *Statistical Distributions; *Structural Equation Models
IDENTIFIERS	*Balanced Incomplete Block Spiralling; Dichotomous Variables; *Missing Data

ABSTRACT

The impact of the use of data arising from balanced incomplete block (BIB) spiralled designs on the chi-square goodness-of-fit test in factor analysis is considered. Data from BIB designs posses a unique pattern of missing data that can be characterized as missing completely at random (MCAR). Standard approaches to factor analyzing such data rest on forming pairwise available case (PAC) correlation matrices. Developments in statistical theory for missing data show that PAC correlation matrices may not satisfy Wishart distribution assumptions underlying factor analysis, this impacting tests of model fit. A new approach for handling missing data in structural equation modeling advocated by B. Muthen, D. Kaplan, and M. Hollis (1987) is proposed as a possible solution to these problems. The new approach is compared to the standard PAC approach in a Monte Carlo simulation framework. Simulation results show that tests of goodness-of-fit are very sensitive to PAC approaches even when data are MCAR, as is the case for BIB designs. The new approach outperforms the PAC approach for continuous variables and is comparatively much better for dichotomous variables. One table and one figure illustrate the discussion. (SLD)

The Impact of BIB-Spiralling Induced Missing Data Patterns on Goodness-of-Fit Tests in Factor Analysis *

David Kaplan
Department of Educational Studies
University of Delaware

OCCASIONAL PAPER OP93-1
January 1993

This report was developed under the following NCAL Project:

"Assessing the Factor Structure of Data Arising from Balanced Incomplete Block Spiralling Designs: A Focus on the NAEP Adult Literacy Assessment"

An earlier version of this paper was presented at the 1992 meeting of the American Educational Research Association, San Francisco, California.

This work was supported by funding from the National Center on Adult Literacy at the University of Pennsylvania, which is part of the Educational Research and Development Center Program (Grant No. R117Q0003) as administered by the Office of Educational Research and Improvement, U.S. Department of Education, in cooperation with the Departments of Labor and Health and Human Services. The findings and opinions expressed here do not necessarily reflect the position or policies of the National Center on Adult Literacy, the Office of Educational Research and Improvement, or the U.S. Department of Education.

The author wishes to thank Feng Yu for valuable research assistance and Richard Sacher for valuable programming assistance.

The Impact of BIB-Spiralling Induced Missing Data Patterns on Goodness-of-Fit Tests in Factor Analysis

David Kaplan
Department of Educational Studies
University of Delaware
Newark, DE

In studies of adult literacy it is often desirable to assess the underlying dimensions of literacy performance. However, traditional applications of such methods as factor analysis ignore the fact that assessments of adult literacy performance utilize a complex matrix sampling technique referred t. as balanced incomplete spiraling. Such complex matrix sampling can have a deleterious impact on the results of factor analysis. This paper considers the impact of the use of data arising from balanced incomplete block (BIB) spiralled designs on the chi-square goodness-of-fit test in factor analysis. Specifically, data arising from BIB designs possess a unique pattern of missing data that can be characterized as missing completely at random (MCAR). Standard approaches to factor analyzing such data rest on forming pairwise available case (PAC) correlation matrices. Developments in statistical theory for missing data show that PAC correlation matrices may not satisfy Wishart distribution assumptions underlying factor analysis, thus impacting tests of model fit. A new approach for handling missing data in structural equation modeling advocated by Muthen, Kaplan, and Hollis (1987) is proposed as a possible solution to these problems. This study compares the new approach to the standard PAC approach in a Monte Carlo simulation framework. Results of these simulations show that tests of goodness-of-fit are very sensitive to PAC approaches even when data are MCAR, as is the case for BIB designs. The new approach is shown to outperform the PAC approach for continuous variables and is comparatively much better for dichotomous variables.

1. Introduction

In large scale national assessments such as the Young Adult Literacy Survey sponsored by the National Assessment of Educational Progress (NAEP) it is not always feasible or desirable for every test item to be administered to every respondent. Nevertheless, it is necessary to ensure a broad and representative coverage of the content of the assessment. One way in which such representation is accomplished is through a variant of matrix sampling referred to as *balanced incomplete block* (BIB) spiralling (Beaton, Johnson, & Ferris, 1987). In essence, a total pool of items is divided into a smaller number of blocks of tasks. The blocks are then assigned to booklets in such a way as to pair each block with every other block in some booklet. The administration of booklets is cycled (spiralled) so that each booklet is completed by a random sample of respondents. For example, Figure 1 shows the assignment of blocks to booklets for the NAEP Young Adult Literacy Survey (Kirsch & Jungeblut, 1986) where each block contains approximately fifteen literacy items. Desirable characteristics of the BIB spiralling approach are that (a) each block appears with the same frequency, (b) positional effects are controlled due to the fact that each block appears once in each of three possible positions, and (3) every pair of blocks appears with exactly one booklet (Kirsch & Jungeblut, 1986, pg II).

Figure 1. Assignment of blocks to booklets for the NAEP Young Adult Literacy Assessment

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

Given that it is of interest to study the underlying dimensions of adult literacy performance, more information must be obtained about the effects of BIB spiralled designs on results of factor analysis. One feature of data arising from a BIB design is that the resulting correlation matrix of BIB spiralled items possesses a particular pattern of missing data. This missing data pattern arises from the fact that estimates of the correlations between items within a block are based on a larger number of respondents compared to the correlations of items within different blocks. In the case of the NAEP reading data (Zwick, 1987), the ratio of these sample sizes is about 9 to 1 while for the NAEP Young Adult Literacy Survey (Kirsch & Jungeblut, 1986) the ratio is about 3 to 1. As Zwick (1987) has noted, the random sampling nature of BIB spiralled data does not guarantee that dimensionality assessment is unaffected by this pattern of missing data. In fact, it is known that pairwise available-case (PAC) methods of calculating correlation matrices can yield problems for factor analysis even if the data are missing completely at random (Little & Rubin, 1987).

To study the effects of BIB spiralled data on dimensionality assessment Zwick (1987), among other things, simulated unidimensional complete data and BIB spiralled data and applied principle component analysis and full information factor analysis (Bock, Gibbons, & Muraki, 1985) to these simulated data sets. When descriptive statistics were examined, Zwick's results showed that BIB spiralled data had little impact on the recovery of the underlying dimensionality when compared to complete data. However, an examination of the chi-square goodness-of-fit statistics rejected the hypothesis of a single underlying factor for both a complete data case and the BIB data case despite the fact that this was the model that generated the data. Zwick notes that although both cases lead to the rejection of unidimensionality, the value of the test statistic for the BIB case was approximately two-thirds the value of the complete data case. Zwick attributes this difference to the number of respondents per item wherein she finds a similar ratio.

While it is true that sample size affects the chi-square test, two alternative explanations of Zwick's are also plausible. First, Zwick generated only one finite sample replication of BIB and complete data. With one sample, the effects of BIB spiralling on the empirical distribution of the chi-square goodness-of-fit test cannot be adequately assessed. Second, with respect to the BIB case, the goodness-of-fit test may be sensitive to PAC data. This counter-explanation is plausible since PAC covariance matrices do not possess the correct statistical properties assumed by maximum likelihood (or asymptotically equivalent) theory (Brown, 1983).

Given these concerns, the purpose of this paper is to more fully explore the effects of BIB spiralled data on goodness-of-fit in an expanded Monte Carlo simulation framework. Moreover, a new approach to assessing the factor structure of BIB type data is presented which is hypothesized to mitigate some of the problems associated with analyzing PAC covariance matrices. Specifically, the application of an estimator developed by Muthen, Kaplan, and Hollis (1987) in the context of structural equation modeling with missing data is proposed and compared to the standard approach of analyzing PAC correlation matrices in a Monte Carlo simulation context. The organization of this paper is as follows:

- Section 2 outlines the statistical problems associated with missing data patterns of the BIB kind.
- Section 3 presents the new estimator proposed by Muthen et al. (1987).
- Section 4 presents the design of a small Monte Carlo study comparing the new estimator against the standard approach to analyzing BIB data.
- Section 5 presents the results.
- Section 6 concludes.

2. Problems with BIB Spiralled Induced Missing Data Patterns

To motivate the proposed estimator in Section 3 below, it is useful to review the known problems of analyzing PAC covariance matrices. Following Little and Rubin (1987, pg. 42) consider the computation of jk^{th} PAC covariance between two variable Y_j and Y_k based on cases for which y_j and y_k are observed. This covariance can be written as

$$s_{ij}^{(jk)} = \frac{\sum_{(jk)} (y_{ij} - \bar{y}_j^{(jk)})(y_{ik} - \bar{y}_k^{(jk)})}{N^{(jk)} - 1}, \quad j,k = 1,2,\ldots p \qquad (1)$$

where $N^{(jk)}$ is the number of cases observed for Y_j and Y_k and $\bar{y}_j$ and $\bar{y}_k$ are sample means calculated over $N^{(jk)}$. Let $s_{jj}^{(j)}$ and $s_{kk}^{(k)}$ be sample variances of y_j and y_k for available cases, respectively. Then, an estimate of the jk^{th} PAC correlation can be expressed as

$$r_{jk} = \frac{s_{jk}^{(jk)}}{\sqrt{s_{jj}^{(j)} s_{kk}^{(k)}}} \tag{2}$$

A small artificial example reported in Little and Rubin (1987, pg. 43) shows that when $p \geq 3$ the expression in (2) can give rise to non-positive definite correlation matrices and covariance matrices (see also Browne, 1982, pg. 88).

An additional problem associated with PAC covariance matrices is that they do not maximize any proper likelihood (Muthen, Kaplan, & Hollis, 1987). Specifically, as shown by Heiberger (1977, see also Brown, 1983), the asymptotic covariance of PAC covariances in (1) can written as

$$\operatorname{acov}(s_{jk}, s_{lm}) = \frac{N_{jklm}}{N_{jk} N_{lm}} (\sigma_{jl}\sigma_{km} + \sigma_{jm}\sigma_{kl}) \tag{3}$$

where N_{jklm} is the number of complete observations on variables j, k, l, and m. It can be seen that the multiplier $(N_{jklm}/N_{jk}N_{lm})$ appears for the maximum likelihood and generalized least squares discrepancy function instead of the usual $1/N$. Moreover, as noted by Browne (1982), the discrepancy function formed from the analysis of PAC covariance matrices is not bounded by zero so that it is not a true discrepancy function. Thus, the Wishart distributional assumptions are violated and this will most probably affect the chi-square goodness-of-fit test (Bollen, 1989). One purpose of this paper is to study the affects of PAC covariance matrices on the empirical distribution of the chi-square test.

3. An Alternative Approach for Factor Analyzing BIB Spiralled Data

3.1 Theoretical Background

In this paper, a procedure is proposed that addresses the problem of missing data arising from BIB spiralling designs. The proposed method is based on recent developments in missing data theory developed by Rubin (1976, see also Little and Rubin, 1987) and applied to structural equation modeling by Muthen, Kaplan, and Hollis (1987). Muthen et al. consider the factor analysis model

$$y^{*} = \nu + \Lambda\eta + \varepsilon \tag{4}$$

where y^* is a vector of potentially observable response variables, ν is a vector of intercepts, Λ is a matrix of factor loadings, η is a vector of common factors, and ε is a vector of unique variables.

In addition to the factor model in (4), Muthen et al. specify a vector of latent selection variables s^* associated with each y^* of the form

$$s^{*} = \Gamma_{\eta}\eta + \Gamma_{y}y^{*} + \delta \ . \tag{5}$$

Associated with each s^* variable is a threshold τ_j such that for the i^{th} observation,

$$s_{ij} = \begin{cases} 1 \text{ if } s^{*}_{ij} > \tau_j \\ 0, \text{ otherwise} \end{cases}$$

In other words, if $s_{ij} = 1$ then the i^{th} unit's value on y^*_j is selected to be observed as y_j and hence is not missing. If $s_{ij} = 0$, then y_{ij} is missing. Thus in this framework, each s^* variable determines if the corresponding y^* variable is to be observed as y_j, or is missing. The strength of the selectivity is determined by the elements in Γ while the amount of missingness is determined by τ.

Under the specification in (4) and (5), Muthen et al. derive the likelihood of the (not missing) observations on y^*. Specifically, Muthen et al. consider two vectors of parameters: a vector for the factor analysis model in (4) and a vector for the selection model in (5). Under general assumptions, the stacked vector of parameters for the factor model can be written as

$$\theta \equiv (\nu, \Lambda, \Psi, \Theta_\varepsilon), \tag{6}$$

where Ψ is the covariance matrix of η, and Θ_ε is the diagonal matrix of unique variances. Analogously, the parameters for the selection equation in (5) can be stacked in the vector

$$\phi \equiv (\Gamma, \Theta_\delta, \Theta_{\delta\varepsilon}, \tau), \tag{7}$$

where $\Theta_{\delta\varepsilon}$ allows for the possibility that ε and δ are correlated. The likelihood of y is a function of θ and ϕ and can be written as (see Rubin, 1976; Little, 1982, 1983; Muthen et al., 1987)

$$\log L(\theta,\phi|\mathbf{y}) = \sum_{g=1}^{G} \log \phi^g(\theta|\mathbf{y}) + \log f(\theta,\phi|\mathbf{y}), \tag{8}$$

where

$$\log \phi^g(\theta|\mathbf{y}) = const. - \frac{1}{2} N^g \log|\Sigma^g| - \frac{1}{2} N^g \, tr\Sigma^{g-1}[\mathbf{S}^g + (\bar{\mathbf{y}}^g - \mu^g)(\bar{\mathbf{y}}^g - \mu^g)']. \tag{9}$$

The expression in (8) is referred to as the "true likelihood" (Muthen et al., 1987). The first term on the right hand side of (8) is the multivariate normal likelihood for the N^g sample units displaying the g^{th} pattern of missing data, and is referred to as the "quasi-likelihood" since it ignores the missing data generating mechanism. The second term on the rhs of (8) represents the mechanism that generates the missing data.

Correct approaches to the maximization of (8) depend on whether the missing data mechanism, $\log f(\theta,\phi|y)$, is "ignorable" or "non-ignorable." In this paper we will consider the use of the quasi-likelihood estimator under one restrictive condition of "ignorability" of the missing data mechanism, namely that the missing data are missing completely at random (MCAR). Conditions for MCAR are $\Gamma_\eta = \mathbf{0}$, $\Gamma_y = \mathbf{0}$ and $\Theta_{\delta\varepsilon} = 0$. Although MCAR is a restrictive and usually unreasonable assumption for most situations it is the correct assumption for properly implemented BIB designs (Zwick, 1987). Under the assumption of MCAR maximization of the quasi-likelihood with respect to θ yields correct estimates of model parameters, correct chi-squares and correct standard errors. Muthen, et al (1987) refer to this estimator as the full quasi-likelihood (FQL) estimator.

The FQL estimator was compared to listwise and pairwise approaches in an extensive population study by Muthen et al. (1987). A variety of conditions were examined, including the less restrictive condition of missing-at-random (MAR; Little & Rubin. 1987) wherein ignorability did not hold. The general findings were that the FQL estimator was superior to traditional approaches even under cases where it was not correct to ignore the right hand side of (8). Missing from the Muthen et al. study was a comparison of these approaches with respect to the empirical distribution of the chi-square goodness-of-fit test.

3.2 Software Implementation of FQL Estimator

Implementation of the FQL estimator for a single population with g distinct missing data patterns makes use of the multiple group factor analysis approach available in LISREL (Joreskog, 1989), LISCOMP (Muthen, 1987), and EQS (Bentler, 1989). With respect to this paper, the groups correspond to the booklets. The procedure requires two steps. In the first step, the alternative hypothesis (H_1) of a pooled unrestricted covariance matrix is specified. This requires formation of g covariance matrices corresponding to g booklets but based on the item responses to all the tasks. By design, these correlation matrices will have missing data corresponding to blocks of items not represented in booklets. The variances and covariances corresponding to missing data in each booklet are coded one and zero, respectively. Once the matrices are formed, identifiable variance and covariance elements common to each booklet are constrained equal. It may be interesting to note that the chi-square test associated with this step tests whether the missing data are MCAR.

In the second step, the null hypothesis (H_0) corresponding to the k-factor model is specified. Utilizing the same multiple group setup as in the first step, the model imposes equality constraints among the common and identifiable parameters of the factor model. A large sample chi-square test of the null hypothesis of k factors is carried out by subtracting the H_1 chi-square from the H_0 chi-square with degrees-of-freedom equal to the difference in corresponding degrees-of-freedom.

It should be noted that for the purposes of this paper we are assuming there is no additional missing data apart from that induced by the BIB design. Clearly, this is not a reasonable assumption in practice since omitted responses can occur for administered items. The method for handling omits in NAEP Young Adult Literacy Survey, for example, was to treat omits after the last valid response in a block as not-reached and omits prior to the last valid response in a block as missing.

4. Design of Simulation Study

In order to compare the FQL procedure to the standard approach of analyzing PAC covariance matrices, a simulation study was employed. Two cases were studied: Case 1 consisted of continuous variables in order to test the new approach under ideal conditions. Case 2 consisted of dichotomous variables in order to test the new approach under more realistic conditions of item level data. Case 2 studies the problem of factor analyzing PAC phi correlation matrices and represents the approach take by Kirsch and Jungeblut (1986) in the analysis of the Young Adult Literacy Survey.

For both cases, one hundred replications of 1000 random variates were generated from a multivariate normal distribution with a covariance matrix that followed a nine-variable/one-factor model. All factor loadings were specified to be 0.7 in the population. This model possesses 27 degrees-of-freedom. For each replication, the data matrix was modified according to a BIB spiralled design in a manner similar to Zwick (1987). That is, the first three items formed Block A, the next three items formed Block B, and the last three items formed Block C. The first 333 examinees received Blocks A and B, the next 333 examinees received Blocks B and C, and last 334 examinees received Blocks A and C. Three BIB spiralled covariance matrices were formed in SPSS (SPSS Inc, 1989) and used as input to the multiple group specification of LISCOMP (Muthen, 1987). In addition, PAC covariance matrices were also formed using SPSS for each replication and analyzed via LISCOMP. For the PAC case two sample sizes were chosen: an average sample size of 466 and a lower bound sample size of 333.

5. Results

Recent theory given in Muthen, et al. (1987) show that under MCAR the PAC approach will yield no large sample bias. A cursory inspection of the results of this study show this to be the case as well. Thus, for ease of presentation, results are presented as they pertain to the empirical distribution of the chi-square goodness-of-fit test. Results are presented in terms of the mean and variance of the chi-square statistic over 100 replications as well as the rejection frequency for nominal Type I error levels of .05, .01, and .001.

5.1 Quality of Data Generation

As a check on the quality of the data generation process, complete data were studied. Recalling that the df = 27, the results show that across 100 replications, the mean $\chi^2 = 25.965$, the variance of $\chi^2 = 43.470$. The rejection frequencies for nominal Type I errors of .05, .01, and .001 were 2, 1, and 1 respectively. Thus, we find that there is a slight underestimation of the expected values for the chi-square distribution with 27 df. While these values should be kept in mind when examining the results below, it was decided that 100 replications was adequate to gauge the behavior of the chi-square distribution for this study.

5.1 Case 1: Continuous Variables

The upper panel of Table 1 presents results comparing the FQL estimator to the standard approach of analyzing PAC covariance matrices under ideal conditions of continuous normally distributed variables. It can be seen that the PAC approach yields clearly unacceptable rejection frequencies for both conditions of sample size. It should be noted that all replications converged properly, thus the results can only be explained by the violation of Wishard distribution assumptions induced by PAC covariance matrices. By comparison, the FQL estimator shows nearly perfect performance with respect to mean, variance and rejection frequency of the chi-square distribution.

5.2 Case 2: Dichotomous Variables - Phi Coefficients

In order to generate dichotomous variables, the continuous variables from the complete data were dichotomized on the basis of a 70/30 split. This corresponds to a threshold κ, say, of .52. All variables were dichotomized according to this threshold so as not to confound the possible effects of PAC data with the occurrence of difficulty factors arising from unequal thresholds in the population. BIB spiralled data were then created in the manner described in Section 4.

The results for the analysis of phi coefficients are displayed in the lower panel of Table 1. It can be seen that for dichotomous data the PAC approach yields unacceptable rejection frequencies and are worse than those found for the continuous data in Section 5.2. This result can be explained by noting that chi-square goodness-of-fit tests are sensitive to categorization (see e.g., Boomsma, 1983; Olsson, 1979) Here, the effects of categorization interact with distributional violations induced by PAC created correlation matrices. By comparison the FQL estimator performs better but rejection frequencies are still unacceptable.

Table 1.

Chi-square mean, variance, and rejection frequencies for PAC and FQL approaches on continuous and dichotomous variables[a]

Data Generation Check

Estimator	Mean	Variance	Rejection Frequencies .05	.01	.001
ML	25.965	43.470	2	1	1

Continuous Variables

Estimator	Mean	Variance	Rejection Frequencies .05	.01	.001
PAC-466[b]	50.404	274.497	68	53	36
PAC-333[c]	35.944	139.592	33	19	08
FQL	28.034	53.548	05	00	00

Dichotomous Variables

Estimator	Mean	Variance	.05	.01	.001
PAC-466	56.575	276.532	87	69	50
PAC-333	40.622	139.278	47	28	13
FQL	28.034	53.548	35	14	05

[a] $E(\chi^2) = 27$, $Var(\chi^2) = 54$

[b] Based on average sample size of 466.

[c] Based on minimum sample size of 333.

6. Conclusion

This paper was motivated by the problem of assessing the literacy skills of young adults when the data are obtained from a BIB spiraling scheme. The results of this study suggest that caution must be exercised when interpreting goodness-of-fit tests from data arising from BIB spiralled designs using PAC approaches. Specifically, this paper demonstrates that the chi-square goodness-of-fit test is extremely sensitive to violations of Wishart distribution assumptions arising from the construction of PAC covariance matrices even when the missing data are MCAR. Recall that even in the ideal case of continuous variables, 68% of the models were rejected at the 5% level. While these results may explain Zwick's (1987) finding of a large goodness-of-fit value for her simulated BIB data, she is essentially correct in stating that the random nature of the spiralling procedure does not guarantee that dimensionality is unaffected. This paper provides the statistical justification for why such problems might occur.

Of particular note is the behavior of the FQL estimator proposed by Muthen et al. (1987). The results indicate extremely good performance for continuous variables, and relatively better performance for dichotomous variables. Nevertheless, the performance of the FQL estimator for dichotomous variables was not adequate and more extensive theory for this case needs to be developed. In particular, the approach of Muthen et al. (1987) must be extended to include more general estimation procedures which explicitly take into account the categorical nature of item responses (e.g., Bock, Gibbons, & Muraki, 1985; Muthen, 1984).

As in all simulation studies, generalizability is limited by the choice of levels of independent variables. The purpose of this paper was to focus on a particular pattern of missing data and to provide an alternative to assessing goodness-of-fit. A factor not examined in this study was the sample size ratios. It may be the case that PAC approaches perform relatively better when less extreme sample size ratios are found. However, given that MCAR is unrealistic except for perhaps all but BIB designs, and given that the sample size ratios for this study are representative of those found in BIB designs such as the Young Adult Literacy Survey, it is felt that these results speak to the problems of ignoring the missing data pattern when factor analyzing data from BIB spiralled assessments.

A final consideration is the problem of omitted responses not induced by the BIB design. A standard approach for dealing with the problem omitted responses for such assessments as the NAEP Adult Literacy study is to treat omitted items after the last valid response within a block as not-reached and omitted responses before the last valid response as incorrect. From the standpoint of item response theory this may be a sensible approach (see Lord, 1980). However, it appears that the argument for such an approach does not rest on a formal model which specifies whether omitted responses are omitted completely at random or are functions of other omissions where perhaps ignorability is not an appropriate assumption. It may be necessary for practical purposes however, to adopt this approach because application of the FQL estimator for general patterns of omitted responses would require covariance matrices to be formed for every distinct missing data pattern. These resulting covariance matrices would likely be based on unreasonably small sample sizes. A similar concern was raised by Muthen et al. (1987) regarding the application of the FQL estimator for general structural equation models. Nevertheless, if current approaches to handling omits and not-reached responses can be justified, then the FQL estimator appears to be a superior approach for treating BIB induced missing data patterns.

References

Beaton, A. E., Johnson, E. G., & Ferris, J. J. (1987). The assignment of exercises to students. In A. E. Beaton (ed). *Implementing the new design: The NAEP 1983-1984 technical report* (pp 97-118). Princeton, NJ: Educational Testing Service.

Bentler, P. M. (1989). *EQS Structural equations program manual.* Los Angeles: BMDP Statistical Software.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1985). *Full-information item factor analysis* (MRC Report No 85-1[revised]). Chicago: National Opinion Research Center.

Brown, C. H., (1983). Asymptotic comparison of missing data procedures for estimating factor loadings. *Psychometrika*, 48, 269-291.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (ed.), *Topics in applied multivariate analysis* (pp. 72-141). London: Cambridge University Press.

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and non-normality.* Ph.D. Thesis, University of Groningen.

Heiberger, R. M. Regression with pairwise-present covariance matrix: A dangerous practice. *Proceedings of the Statistical Computing Section, 1977.* Washington, DC: American Statistical Association.

Joreskog, K.G., & Sorbom, D. (1989). *LISREL 7: A guide to the program and applications.* Chicago: SPSS Inc.

Kirsch, I. S., & Jungeblut, A. (1986). *Literacy: Profiles of America's young adults.* Princeton, NJ: Educational Testing Service.

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

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Muthen, B. O. (1987). LISCOMP: Analysis of linear structural equations using a comprehensive measurement model. Chicago: Scientific Software, Inc.

Muthen, B., Kaplan, D., & Hollis, M. On structural equation modeling with data that are not missing completely at random. *Psychometrika, 52,* 431-462.

Olsson, U. (1979). On the robustness of factor analysis against crude classification of the observations. *Multivariate Behavioral Research,* 485-500.

Rubin, D. B. (1976). Inference and missing data. *Biometrika,* 63, 581-592.

Zwick, R. (1987). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement, 24,* 293-308.