Developments in Survey Methodology
Anuška Ferligoj and Andrej Mrvar (Editors)
Metodološki zvezki, 15, Ljubljana: FDV, 2000

Assessing Reliability and Validity in the Context of Planned Incomplete Data Structures for Multitrait-Multimethod Models

Brendan P. Bunting[1] and Gary Adamson[2]

Abstract

This paper examines some designs for the effective use of the EM algorithm in minimising the burden of data collection in the context MTMM models. Within MTMM models it is possible to distinguish between reliability and validity, however, a serious limitation with this approach is that data are required on at least three occasions and with three different methods. A planned incomplete data structure is proposed which reduces substantially the amount of data required from each individual at a third point in time. The efficacy of this approach is tested via a series of simulations and the resulting parameter estimates are shown to be both precise and efficient.

1 Introduction

To assess the reliability of a *single* measure, the measure must be administered on a number of subsequent occasions: this is generally referred to as 'test-retest' reliability. However, repeated administration of an item has at least two associated disadvantages. Firstly, each repeated administration provides an additional opportunity for respondent attrition. Secondly, the potential exists for various experiences to occur unevenly across respondents during the periods between administrations, which may unduly affect the estimation of reliability. To reduce the likelihood of such experiences occurring, perhaps one simple solution is to reduce the length of the intervening periods. However, reducing the time period between administrations may produce a confounding memory effect, whereby a respondent's current response is in some way affected by a remembered one. These

[1] School of Psychology and Communication, Faculty of Social & Health Sciences and Education, University of Ulster, Northern Ireland.
[2] School of Education, Faculty of Social & Health Sciences and Education, University of Ulster, Northern Ireland.

seemingly extricable problems associated with repeated administrations provide the impetus for this paper.

In the context of Structural Equation Modelling (SEM), the above mentioned problems associated with repeated administrations are more than vexatious, since it can be shown that when a measure is repeated on three or more occasions an identifiable model can be obtained, and under mild restrictions, both reliability and stability can be assessed (Wiley, 1973). Models of this type are usually referred to as quasi-simplex (Jöreskog and Sörbom, 1989). However, a limitation with this model type is that a distinction cannot be made between reliability and validity (Saris, 1995). In order to partition variance into a reliable and valid component it is necessary to obtain more than one measure of a given trait at each point in time. An approach which permits assessment of both reliability and validity is the multitrait-multimethod (MTMM) model. In this model three traits are usually measured at three points in time, with a different method used to measure the traits on each occasion. So, in contrast to quasi-simplex models, the MTMM model offers a way of distinguishing between reliability and indicator validity (Saris and Andrews, 1991). Furthermore, the MTMM model can be parameterised in such a way that the coefficients relating to the indicator validity can be corrected for attenuation and the formulation of true score validity coefficient(s) is attractive as a clear distinction can be made between reliability and validity (Saris and Andrews, 1991). A limitation of this method is that it must be assumed that individuals have not changed on the constructs across the data collection occasions. The likelihood of realising this assumption empirically is perhaps somewhat minimal, notwithstanding the heavy burden that remains on respondents, since responses are elicited on three different occasions. Indeed, in relation to many contexts, but with regard to psychological research in particularly, it may not be possible to realise the assumption of stability fully. Nevertheless, the MTMM model would be considered more attractive in the eyes of an applied researcher if the burden of repeated administrations could be reduced.

In this paper a method is proposed which permits estimation of reliability and validity, while reducing the burden of repeated administrations for many of the participants. The proposed method concerns constructing a research design which incorporates *planned incomplete data structures*: whereby, all participants do not complete all subsequent administrations. However, successful implementation of such a research design is predicated on being able to deal effectively with the resultant planned missing data - where effectiveness in this instance refers to obtaining correct parameter estimates and standard errors in the presence of missing data.

There are several approaches for dealing with missing data in the context of planned incomplete data structures; however, an important issue, which needs to be addressed firstly, is the mechanism underlying the missing data. Rubin (1976) and Little and Rubin (1987) distinguish three categories of missing data, which are defined according to the information that they provide in respect to the observed data: missing completely at random (MCAR), missing at random (MAR), or nonignorable. Missing completely at random (MCAR) exists when missing values are randomly distributed across all respondents, and therefore, under a MCAR process, whether a variable's data are observed or missing is not thought to affect its distribution (i.e. P(Y | y missing) = P(Y | y observed)). Missing at random (MAR) exists when missing values are not randomly distributed across all respondents but are randomly distributed within one or more sub-samples (e.g., missing more among males than females, but random within each sub-sample). MAR is a more relaxed condition, assuming only that missing and observed distributions are identical, conditional on a set of predictor or stratifying variables X (i.e. P(Y | y missing, X) = P(Y | y observed, X)). There are a number of methods for dealing with missing data. One of the simplest procedures is listwise deletion, where all individuals with data missing are removed from the analysis. This has the advantage that the resulting variance-covariance matrix will be positive definite. However, with the reduction in sample size, larger standard errors can be expected, though the parameter estimates have been shown to be accurate (Satorra, 1990). The major difficulty with the listwise deletion is that the sample size can be substantially reduced and therefore this method is likely to compromise the statistical power of the analysis. On the face of it, a more attractive procedure is pairwise deletion. Pairwise deletion is conceptually attractive, since the resulting variance-covariance matrix is derived from all of the information available. However, there are various methods for implementing the pairwise deletion procedure and no strict convention exists for deciding on the final number of respondents to be used in the structural model. This is an important aspect, since sample size can have a significant bearing on the magnitude of the standard errors associated with any given parameter. In addition, pairwise deletion can sometimes produce a variance-covariance matrix that is not positive definite: standardised correlations can have values greater than one. Nevertheless, in some simulation work, parameter estimates have been found to be unbiased (Graham et al., 1996; Arbuckle, 1996) and efficient standard errors have been produced using the pairwise deletion procedure.

One method in which the variance-covariance matrix could be computed for the entire sample would be to impute the missing values. A simple way, in which this can be achieved for example, is by replacing the missing value with the mean value of the variable. However, even when it is reasonable to assume that the data are MCAR this procedure may result in a spiked distribution with possible attenuation of item correlations and underestimation of item variance. An easy and less restrictive alternative exists to mean imputation - the use of the EM algorithm (Little and Rubin, 1987). Using this approach, optimal estimates are obtained for the missing data (Graham et al., 1996) under the less restrictive assumption that the data are MAR. One limitation with this approach is that the standard errors may be somewhat imprecise, however, the degree of imprecision appears to be marginal (Graham et al, 1996). Alternatively, precise standard errors can be obtained using the bootstrap procedure: in this event five to ten imputed data sets are required to obtain 'correct' standard errors for each model, following the procedure described by Rubin (1987) and Schafer (1997).

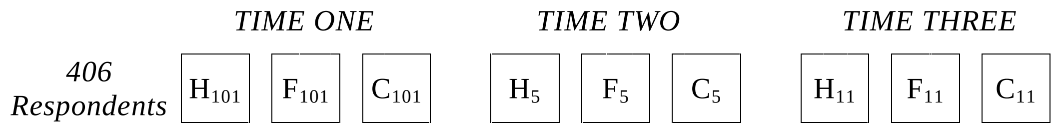


Figure 1: The MTMM research design used by Coenders et al. (1995).

Given that there appears to be an effective method for dealing with missing data, it therefore seems plausible to incorporate planned incomplete data structures into research designs. The obvious advantage is that less data is required and therefore less data needs to be collected. However, it is not at all clear how patterns of incomplete data should be structured and incorporated into research designs, particularly in longitudinal designs such as the MMTM design described previously.

In cross-sectional research designs the use of a *reference variable* has been shown to be effective in terms of obtaining the correct estimates in the context of planned incomplete data structures (McArdle, 1994; Graham et al., 1996). In such designs, a reference variable refers to obtaining complete data for a given variable, or for one variable within each factor of the research design. The use of a reference variable would also seem an attractive option in some longitudinal research designs, such as MTMM, since it could be incorporated readily across the three administrations. Specifically, in MTMM contexts a reference variable could be incorporated so that, across the three occasions, complete information can be obtained for each of three traits and methods (see Figure 3 for a schematic representation of how a reference variable can be incorporated across each of the three time periods). The effectiveness of such a research design remains to be tested empirically; however, one aim of the current paper is to assess the effectiveness of this research design using Monte Carlo simulations.

To explore the effectiveness of a reference variable in the context of patterns of incomplete data structures with the MTMM longitudinal research design, a MTMM model developed by Coenders et al. (1995), was chosen for a series of simulations. Coenders et al. (1995) formulated a true score MTMM model for the assessment of three questions relating to satisfaction with one's housing (H), finances (F), and social contacts (C). These questions were posed on three occasions, using three different methods: 101 point numeric scale ranging from 0 to 100; 5 point scale with all labelled categories from 1 to 5; and an 11 point scale ranging from 0 to 10, to a sample of 406 individuals from Catalonia in Spain (see Figure 1 for a schematic representation of this research design). Coenders et al. (1995) model specification, together with the parameters estimates (estimates associated with T_H_100, T_F_100, T_C_100 etc. are the validities; H_100, F_100, C_100 etc. are the reliabilities; estimates associated with the traits are correlations and the estimates associated with the methods are variances) and the associated standard errors (in brackets) are presented in Figure 2. While the specification of this model has no direct bearing on the current paper, other than being chosen for its MTMM structure, it is perhaps worth noting that the contribution of Coenders et al. (1995) model was a new set of constraints which achieved stability, while the model scale remained invariant - previous constraints used to improve stability made the model scale dependent and thus not estimable from a correlation matrix.

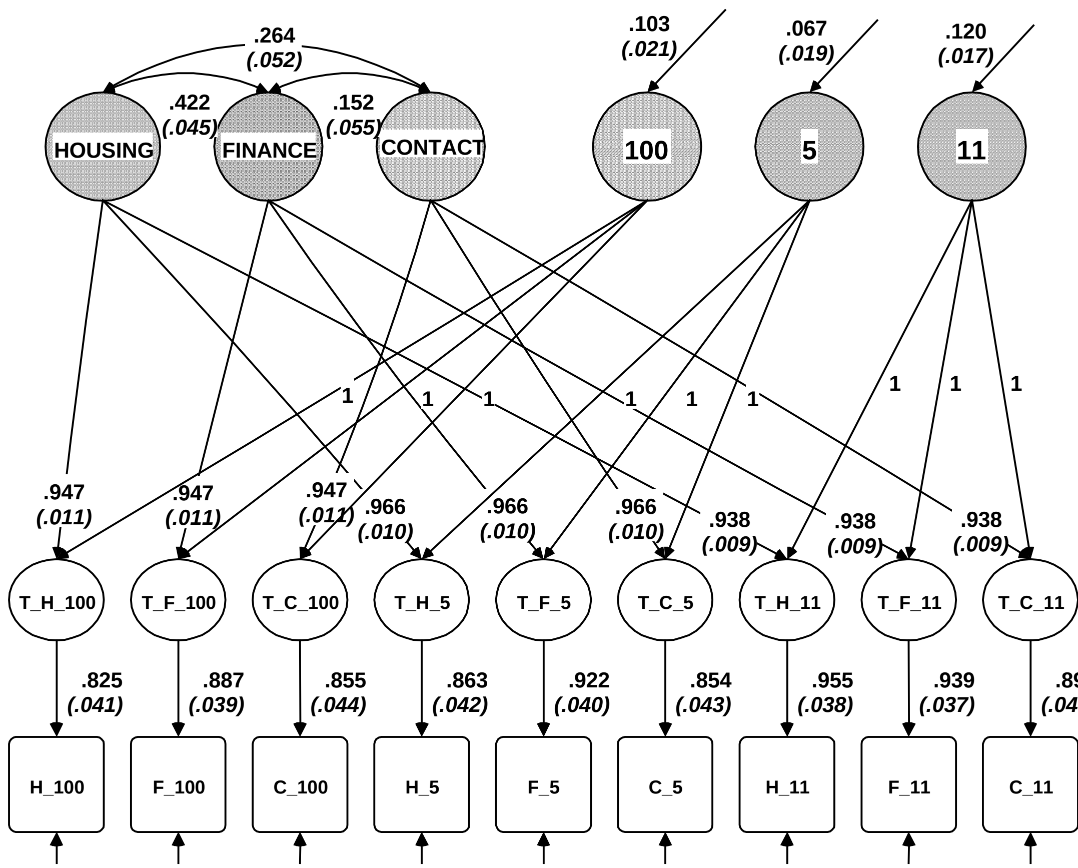


Figure 2: Coenders et al. (1995) model specification and resulting parameter estimates and standard errors (in brackets).

To clarify, the intention of this paper is to use Coenders et al. (1995) implied population variance-covariance matrix to simulate a number of data sets, then to introduce the same pattern of incomplete data into each data set. The three methods of dealing with incomplete data (i.e. Pairwise, Listwise, and EM) will then be employed. Coenders et al. (1995) original model will then be applied to the resulting variance-covariance matrices and the resulting parameter estimates (coefficients) and standard errors, for each, will be compared to the original results. In effect, the Coenders' model (and its coefficients) acts as a benchmark to test the efficacy of the various methods for dealing with missing data in the context of incomplete data structures; the objective being to assess the degree of precision and efficiency with which the parameter estimates and standard errors are re-captured, after using listwise, parwise deletion and EM imputation.

2 Method

The starting point was to generate 100 data sets, which conformed to the Coenders et al. (1995) implied population variance-covariance matrix. For practical reasons, the 100 data sets were generated using the programmability of the SPSS (version 8) statistical package, but in essence, the method and rationale detailed in the LISREL/PRELIS (version 8) manual was followed closely (Joreskog and Sorbom 1996: 6 - 12). However, rather than generating a sample size of 406 for each data set, as in the Coenders et al.. study, the sample size was adjusted to 408 in the simulations for convenience, since this can be divided by 3 and a whole number obtained. Following construction of the 100 data sets the pattern of missingness detailed in Figure 3 was introduced to each set - the blank squares indicate the portion of data which was removed.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_{101}			F_5	C_5	H_{11}		C_{11}
136 Respondents	H_{101}	F_{101}	C_{101}	H_5	F_5	C_5	H_{11}	F_{11}	C_{11}
136 Respondents	H_{101}		C_{101}	H_5	F_5			F_{11}	C_{11}

Figure 3: Planned pattern of incomplete data across 3 points in time.

As mentioned earlier, an important aspect is the presence of a reference variable: Figure 3 shows that a reference variable is evident for ‘housing (H)’ at ‘time one’, ‘finance (F)’ at ‘time two’ and for ‘contacts (C)’ at time ‘three’ with the method varied across these three occasions. This means that full information for all participants is obtained on one occasion for each trait and method. Use of a reference variable as detailed in Figure 3 has obvious utility insofar as it acts as a baseline measure, which should aid the imputation process since full information is provided for all participants in relation to the three traits and the three methods, albeit on one occasion only. Another important aspect is that for those participants, on whom incomplete data is planned, the research design should conform to the assumption of either MCAR or MAR. In these simulations the pattern of missingness is considered MCAR since all participants are in theory, randomly allocated to the groups.

The next step was to implement the various methods for dealing with missing data. The methods used were listwise, pairwise, and the EM algorithm. The first two are conventional methods and are readily accommodated within the PRELIS package. One hundred variance-covariance matrices were produced by PRELIS for each of the methods. As regards the EM algorithm, the variance-covariance matrices for this imputation method were generated using the ‘NORM’ program (Schafer, 1997). Unfortunately, LISREL cannot read the output directly from

'NORM', but code written by Graham (1997), called EMFIX, amends the EM output so that LISREL can read the variance-covariance matrix produced by NORM.

An important issue when using different methods for dealing with missing data is deciding which N-value (number of participants) to use for each analysis. For example, no strict convention exists for deciding on the number of participants that should be included when the analysis involves pairwise deletion as the N-value can vary across pairs of variables. In the current study, it was decided to opt for 408 for pairwise as this choice optimises the chances of the pairwise method retrieving the best possible standard errors. Across the various methods the sample size was: 408 for the data based on pairwise deletion; 136 for the listwise deletion; and 317 for the imputation method - this last n-value was chosen because approximately 22% of the data were imputed. Finally, having decided on an appropriate n-value for each method, the Coenders et al. (1995) model was then applied to each of the generated variance-covariance matrices and the resulting parameter estimates and standard errors stored.

3 Results

An initial observation, which is noteworthy, is that a number of the models failed to converge: seven failed to converge with the pairwise deletion method; and when data were imputed using EM, two models failed to converge; while all models converged with the listwise deletion method. To explore the effectiveness of the three methods in dealing with missing data in the context of planned incomplete data structures within the MTMM research design, the estimates derived using pairwise, listwise, and EM procedures were compared with the original estimates (i.e. original parameter estimates where subtracted from the simulated estimates). Obviously, the greater the discrepancy between the original parameter estimates (as detailed in the Coenders et al. model specification in Figure 2) and the simulated estimates the more biased (less precise) the method. Moreover, the greater the spread of scores of the simulated estimates the less efficient the method. Therefore, for a method to be considered effective it should be both precise and efficient.

3.1 Reliabilities

The reliabilities computed for this model are in keeping with those defined in classical test theory. In this formulation, reliability corresponds to that portion of the variance that is free from random error.


Reliabilities
Difference between Original and Simulated
.60
.40
.20
-.00
-.20
-.40
-.60
Pairwise
Listwise
Em_317

perform relatively well in terms of point estimation, that is, they exhibit similar precision, but EM out-performs listwise in terms of efficiency (all simulated estimates are more tightly clustered around the point estimation). Overall, the EM method is rather precise and efficient, re-capturing the original reliability estimates quite effectively.

3.2 Validities

The original model estimates for the validities are detailed in Figure 2 (labelled: T_H_100, T_F_100, T_C_100, T_H_5, etc.). The differences between the simulated validities and the original are presented graphically in Figure 5 and the corresponding means and standard deviations of these distributions are displayed in Table 2.

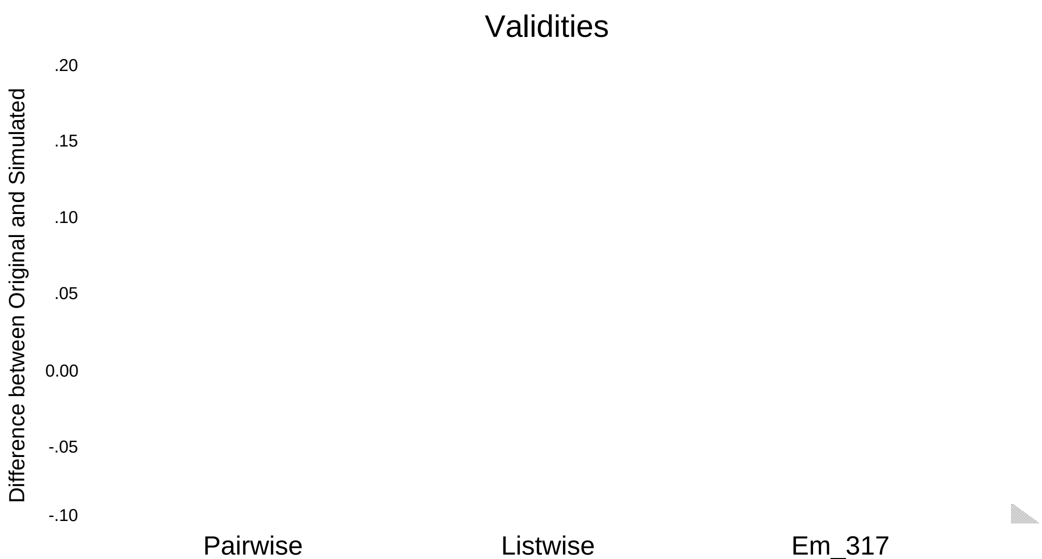


Table 2: Giving the means and standard deviations for the distribution of estimated validities around the original model values.

	Pairwise		**Listwise**		**Em_317**	
	Mean	*Std. D.*	*Mean*	*Std. D.*	*Mean*	*Std. D.*
H_100	-.0277	(.0212)	.0002	(.0171)	-.0013	(.0148)
F_100	-.0277	(.0212)	.0002	(.0171)	-.0013	(.0148)
C_100	-.0277	(.0212)	.0002	(.0171)	-.0013	(.0148)
H_5	.0125	(.0210)	.0004	(.0162)	.0006	(.0131)
F_5	.0125	(.0210)	.0004	(.0162)	.0006	(.0131)
C_5	.0125	(.0210)	.0004	(.0162)	.0006	(.0131)
H_11	.0677	(.0241)	-.0009	(.0157)	-.0010	(.0109)
F_11	.0677	(.0241)	-.0009	(.0157)	-.0010	(.0109)
C_11	.0677	(.0241)	-.0009	(.0157)	-.0010	(.0109)

3.3 Correlations and Variances

The original model estimates for the correlations and variances are detailed in Figure 2 (i.e. correlations between the traits, and variances for the methods). The differences between the simulated correlations and variances and the original are presented graphically in Figure 6 and the corresponding means and standard deviations for these distributions are displayed in Table 3.

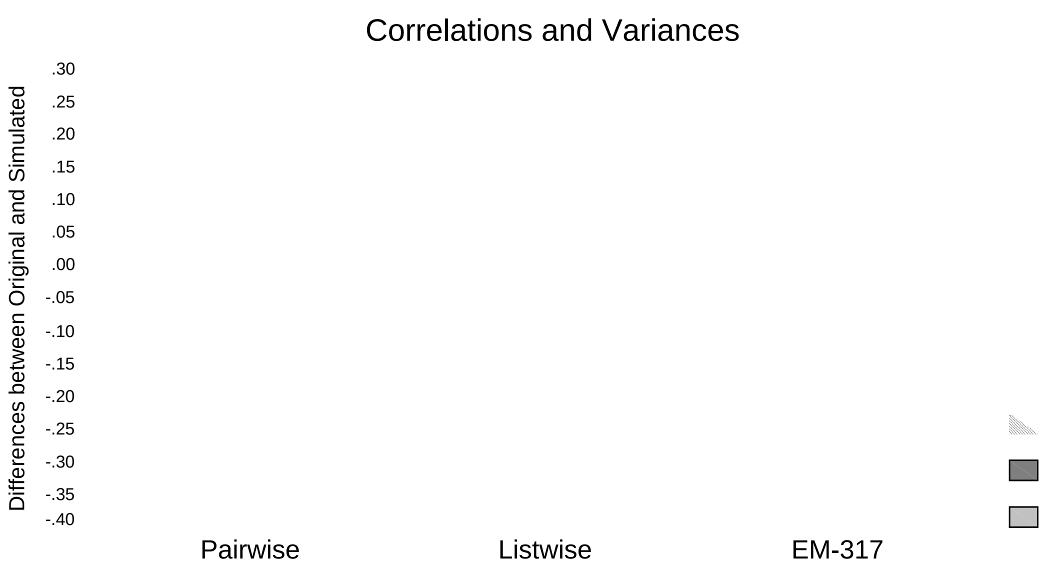
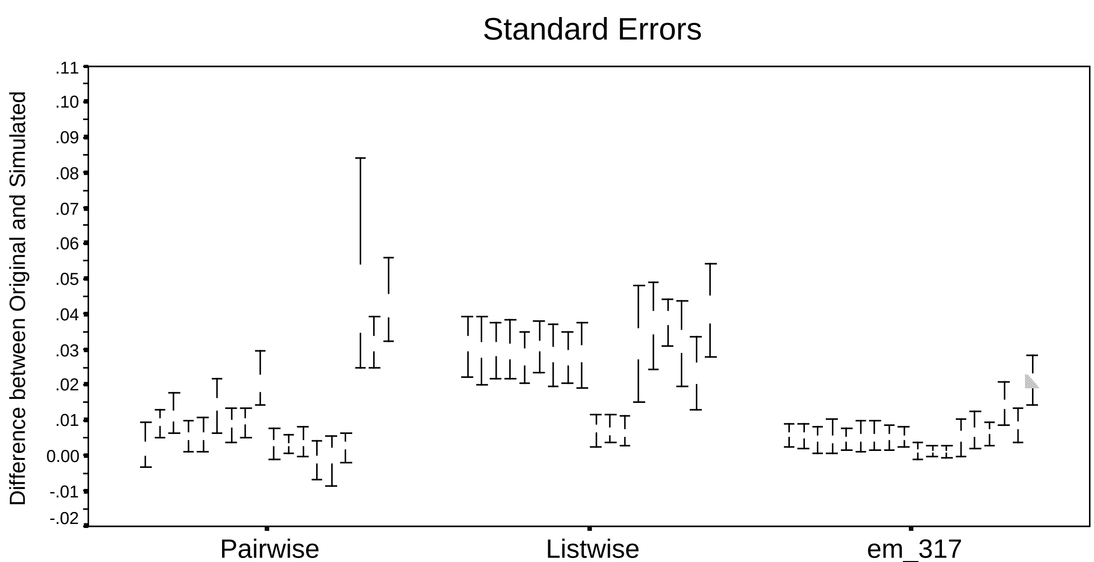


Table 3: Giving the means and standard deviations for the distribution of estimated correlations and variances around the original model values.

	Pairwise		**Listwise**		**EM-317**	
	Mean	*Std. D.*	*Mean*	*Std. D.*	*Mean*	*Std. D.*
H_F	-.1404	(.0750)	.0100	(.0841)	.0029	(.0461)
H_C	-.0566	(.0780)	-.0048	(.0931)	-.0061	(.0594)
F_C	.0283	(.0722)	.0008	(.0902)	.0006	(.0596)
M_101	.0512	(.0389)	-.0008	(.0325)	.0022	(.0279)
M_5	-.0247	(.0409)	-.0010	(.0312)	-.0014	(.0254)
M_11	-.1322	(.0483)	.0014	(.0293)	.0018	(.0204)

Similar to the previous simulated parameter estimates, pairwise performs less well, compared to both listwise and EM in terms of re-capturing the correlations and variances. While, listwise and EM perform similarly in relation to precision, EM, again, displays greater efficiency given the smaller standard deviations in Table 3.



Inspection of Figure 7 and Table 4 indicate that the results for the standard errors do not follow the pattern of results found for the parameter estimates. Unlike its performance in relation to re-capturing the parameter estimates the listwise procedure lacks both precision and efficiency in relation to the standard errors. The standard errors produced by the pairwise procedure vary somewhat across the various estimates, but generally, pairwise performs quite well. The most precise standard errors were those produced by EM, and in most cases, were only slightly larger than the original standard errors. In addition, the EM procedure performs rather well in terms of efficiency as denoted by the standard deviations in Table 4.

Table 4: Giving the means and standard deviations for the distribution of estimated standard errors around the original model values.

	Pairwise		**Listwise**		**em_317**	
	Mean	*Std. D.*	*Mean*	*Std. D.*	*Mean*	*Std. D.*
H_100	.0022	(.0035)	.0313	(.0036)	.0059	(.0015)
F_100	.0092	(.0018)	.0296	(.0039)	.0056	(.0016)
C_100	.0112	(.0026)	.0303	(.0038)	.0047	(.0019)
H_5	.0056	(.0021)	.0297	(.0038)	.0048	(.0019)
F_5	.0057	(.0022)	.0281	(.0034)	.0045	(.0013)
C_5	.0144	(.0030)	.0309	(.0034)	.0052	(.0018)
H_11	.0090	(.0021)	.0284	(.0033)	.0052	(.0018)
F_11	.0095	(.0017)	.0275	(.0034)	.0052	(.0016)
C_11	.0210	(.0042)	.0290	(.0038)	.0051	(.0015)
HFC_100	.0035	(.0017)	.0077	(.0020)	.0014	(.0010)
HFC_5	.0032	(.0013)	.0069	(.0018)	.0011	(.0008)
HFC_11	.0036	(.0017)	.0066	(.0019)	.0011	(.0009)
H_F	-.0015	(.0028)	.0313	(.0067)	.0055	(.0025)
H_C	-.0007	(.0041)	.0372	(.0050)	.0066	(.0021)
F_C	.0029	(.0018)	.0387	(.0031)	.0066	(.0014)
M_100	.0470	(.0171)	.0327	(.0058)	.0143	(.0025)
M_5	.0317	(.0034)	.0234	(.0046)	.0085	(.0020)
M_11	.0425	(.0051)	.0415	(.0064)	.0212	(.0033)

4 Discussion

In the context of the incomplete data structures used in this simulation study, the results demonstrate that the EM algorithm out-performs the more conventional forms of dealing with patterns of incomplete data. The EM procedure re-captured both the original parameter estimates and associated standard errors with a substantial degree of precision. In general, it not only produced the more precise estimates but also the most efficient. However, EM performed less well compared to listwise in one respect – the variance-covariance matrix produced by EM failed to converge on two occasions, whereas this was never the case with listwise. The listwise procedure did produce fairly precise and efficient parameter estimates, but they were generally not within as narrow a band as those produced using the EM procedure. Furthermore, the listwise procedure did not perform well in recovering the standard errors, indeed, the standard errors produced by listwise were both biased and inefficient. The somewhat surprising finding was the poor performance of the pairwise procedure. This was the only procedure where the original sample size of N=408 was maintained. Yet, the simulated parameter estimates produced by pairwise were invariably biased and inefficient, though, with the exception of some cases, the pairwise procedure did produced quite precise and efficient standard errors. The comparatively poor performance of pairwise is particularly surprising since previous research indicates that the pairwise procedure should perform fairly well under the MCAR assumption[3], but it is difficult to identify any single reason why pairwise should perform so ineffectively in this instance. Notwithstanding, the main finding of this research is that EM performs effectively in the presence of incomplete missing data structures within the context of the specified MMTM model.

In light of the effective performance of the EM procedure in dealing with the planned incomplete data, it may be fair to assume that the research design detailed in Figure 3 could be implemented empirically. While the concept underlying this design is perhaps not entirely novel in terms of research designs generally (similar to some cross-sectional research designs, e.g., McArdle, 1994; Graham et al., 1994), it nevertheless demonstrates the efficacy of the approach in relation to MMTM longitudinal research designs specifically.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_{101}	C_5	H_{11}	F_5	C_{11}			
136 Respondents	H_{101}	F_{101}	C_{101}	H_5	F_5	C_5	H_{11}	F_{11}	C_{11}
136 Respondents	H_{101}	F_5	C_{101}	H_5	F_{11}	C_{11}			

Figure 8: Design for administering the planned pattern of incomplete data, showing how the planned pattern in Figure 3 can be collapsed so that only a third of participants answer at the third period in time.

In terms of longitudinal designs, the research design featured in Figure 3 has a number of advantages: (a) not all participants are required to respond on all three occasions; and (b) very precise and efficient parameter estimates are obtained with the EM procedure. However, it also has a major disadvantage: to obtain reference information on the traits and method, at each point in time, data need to be distributed equally across all three occasions. This disadvantage can be somewhat minimised by restructuring the scheme presented in Figure 3, in such a way, that

---

[3] The authors note that the procedure used to implement pairwise deletion in Prelis/Lisrel has been modified since this paper was completed.

the total sample respond to questions at times one and two, leaving only one third of the respondents to be approached on a third occasion. Figure 8 shows how the pattern of incomplete data described in Figure 3 can be restructured so that only one-third of respondents need be approached at the third period in time. Such a restructuring of the data is only possible where it is reasonable to assume that no change in opinion or attitude has occurred between occasions.

The re-organisation of the data in Figure 8 has the obvious advantage that only one-third of the participants is approached at time 3. However, this configuration could lead to confounding method and trait effects, since some participants are required to respond using two different scaling methods on one occasion, while others respond using one method only. In order to avoid this possible confounding effect one could simply re-organise the design in such a way that responses to each question, within a given occasion, were obtained using a different method. This alternative arrangement for presenting a mixture of methods is displayed in Figure 9. However, in terms of the imputation process and analysis the data must be restored to the original incomplete data pattern and this procedure will only be effective where no change has occurred between the data collection periods: of course, with MTMM models this assumption is standard.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_5	C_{11}	H_{11}	F_{101}	C_5			
136 Respondents	H_{11}	F_{101}	C_5	H_5	F_{11}	C_{101}	H_{101}	F_5	C_{11}
136 Respondents	H_5	F_{11}	C_{101}	H_{101}	F_5	C_{11}			

Figure 9: An alternative design for avoiding potential confounding method and trait effects.

An obvious limitation with the research design presented in Figure 9 is that one third of participants need to be approached on a third occasion. In some follow-up work undertaken by the authors, the possibility of reducing the data collection to two points in time was explored. This was achieved by not having a full reference variable on each occasion; that is, omitting one part of the reference variable on each time occasion (i.e. $H_{101,}$ T1, $F_{5,}$ T2, and $C_{11,}$ T3 from Figure 3) - the resulted configuration is given in Figure 10. This configuration can then be collapsed to two points in time as in Figure 11. However, in order to identify the MTMM model it is essential, as already pointed out, to restructure the data to assume a time structure for data at three points in time, i.e. as in Figure 10.

The research design presented in Figure 11 is the configuration presented in Figure 10 collapsed to two points in time. However, as noted earlier, a potential confounding effect could exit with this configuration: each group of participants (136 in each group), on each occasion, respond to two traits using the same scaling method and respond to the third trait using a different scaling method. The different scaling methods could, again, possibly have an undetermined and differential affect on how the participants rate the traits. Consequently, a better approach would be to present a mixture of scaling methods on each occasion, as detailed in Figure 12.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_{101}			F_5	C_5	H_{11}		C_{11}
136 Respondents		F_{101}	C_{101}	H_5		C_5	H_{11}	F_{11}	
136 Respondents	H_{101}		C_{101}	H_5	F_5			F_{11}	C_{11}

Figure 10: Planned pattern of incomplete data for 2 points in time.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_{101}	C_5	H_{11}	F_5	C_{11}			
136 Respondents	H_5	F_{101}	C_{101}	H_{11}	F_{11}	C_5			
136 Respondents	H_{101}	F_5	C_{101}	H_5	F_{11}	C_{11}			

Figure 11: Planned pattern of incomplete data in Figure 10 collapsed to 2 points in time.

	TIME ONE			*TIME TWO*			*TIME THREE*		
136 Respondents	H_{101}	F_5	C_{11}	H_{11}	F_{101}	C_5			
136 Respondents	H_{11}	F_{101}	C_5	H_5	F_{11}	C_{101}			
136 Respondents	H_5	F_{11}	C_{101}	H_{101}	F_5	C_{11}			

Figure 12: Design for administering the planned pattern of incomplete data, showing how the planned pattern in Figure 10 can be collapsed so that participants answer only at two periods in time, with the scaling method varied within participants.

The research design in Figure 12 is particularly attractive since data collection has been reduced to two points in time and different methods have been used on each of the two occasions. This approach has added value, insofar as it could minimise memory effects (i.e. because participants have to evaluate the trait on a different scale overtime) and hence shorten the required time period between administrations. As might be expected however, given the greater uncertainty introduced by having data from only two points in time (but rearranged for a three points in time data structure), together with a reduction in sample size, greater variation in both the standard errors and parameter estimates might be expected. Some of these limitations could be reduced however, by increasing the sample size. An increase in sample size should not be considered as an overriding disadvantage, since this design allows the researcher to collect data at two points in time only, and the time and effort saved, notwithstanding the financial benefits, would appear to make this design most appealing.

References

[1] Arbuckle, J.L. (1996): Full information estimation in the presence of incomplete data. In G.A. Marcoulides and Schumacker, R.E. (Eds.), *Advanced Structural Equation Modeling: Issues and Techniques*. Mahwah, NJ: Lawrence Erlbaum Inc, 243-277.

[2] Coenders, G., Batista-Foguet, J.M., and Satorra, A. (1995): Scale dependence of the true score MTMM model. In W.E. Saris and A.Munnich (Eds.), *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments*. Budapest: Eotvos University Press, 243-277.

[3] Graham, J. (1997): *EMFIX*. The methodology center: The Pennsylvania State University. Available at http://methcenter.psu.edu

[4] Graham, J.W., Hofer, S.M., and MacKinnon, D.P. (1994): Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, **31**, 197-218.

[5] Jöreskog, K.G. and Sörbom, D. (1989): *LISREL 7 User's Reference Guide.* Chicago: Scientific Software.

[6] Joreskog, K.G. and Sörbom, D. (1996): *PRELIS 2 User's Reference Guide.* Chicago: Scientific Software International.

[7] Little, R.J.A. and Rubin, D.B. (1987): *Statistical Analysis with Missing Data*. New York: J. Wiley & Sons.

[8] McArdle, J.J. (1994): Structural factor analysis experiments with incomplete data. *Multivariate Behavioural Research*, **29**, 409-451.

[9] Rubin, D.B. (1987): *Multiple Imputation for Nonresponse in Surveys*. New York: J. Wiley & Sons.

[10] Saris, W.E. and Andrews, F.M. (1991): Evaluation of measurement instruments using a structural modeling approach. In P.P. Biemer, R.M. Groves, L.E. Lyberg, N. Mathiowetz, and S. Subman (Eds.), *Measurement Errors in Surveys*. New York: Wiley, 575-579.

[11] Saris, W.E. (1995): Designs and models for quality assessment of survey measures. In Saris, W.E. and Münnich, Á. (Eds.), *The Multitrait-Multimethod Approach to Evaluate Measurement Instruments*. Budapest: Eötvös University Press, 9-37.

[12] Satorra, A. (1990): Robustness issues in the analysis of MTMM and RMM models. In W.E. Saris and A. van Meurs (Eds.), *Evaluation of Measurement Instruments by Meta-Analysis of Multitrait Multimethod Studies*. Amsterdam: North-Holland.

[13] Schafer, J.L. (1997): *Imputation of Missing Data under a Multivariate Linear Mixed Model*. Technical report 97-100, The Methodology Center, The Pennsylvania State University. Available at http://methcenter.psu.edu.

[14] Scherpenzeel, A.C. (1995): *A Question of Quality: Evaluating Survey Questions by Mulitrait-Multimethod Studies*. Doctoral dissertation, University of Amsterdam. Leidschendam, the Netherlands: Royal PTT Nederland.

[15] Wiley, D. (1973): The identification problem for structural equation models with unmeasured variables. In A.S. Goldberger and O.D. Duncan (Eds.), *Structural Equation Models in the Social Sciences*. New York: Seminar Press, 69-83