On Obtaining Estimates of the Fraction of Missing Information From Full Information Maximum Likelihood

Victoria Savalei & Mijke Rhemtulla

Published online: 31 Jul 2012.

Submit your article to this journal ⬚

Article views: 527

View related articles ⬚

Citing articles: 18 View citing articles ⬚

Ψ **Psychology Press**
Taylor & Francis Group

TEACHER'S CORNER

On Obtaining Estimates of the Fraction of Missing Information From Full Information Maximum Likelihood

Victoria Savalei
University of British Columbia

Mijke Rhemtulla
University of Kansas

Fraction of missing information λ_j is a useful measure of the impact of missing data on the quality of estimation of a particular parameter. This measure can be computed for all parameters in the model, and it communicates the relative loss of efficiency in the estimation of a particular parameter due to missing data. It has been recommended that researchers working with incomplete data sets routinely report this measure, as it is more informative than percent missing data (Bodner, 2008; Schafer, 1997). However, traditional estimates of λ_j require the implementation of multiple imputation (MI). Many researchers prefer to fit structural equation models using full information maximum likelihood rather than MI. This article demonstrates how to obtain an estimate of λ_j using maximum likelihood estimation routines only and argues that this estimate is superior to the estimate obtained via MI when the number of imputations is small. Interpretation of λ_j is also addressed.

Keywords: fraction of missing information, full-information maximum likelihood, incomplete data, multiple imputation

The two most popular modern approaches to handling missing data in the context of multivariate techniques such as regression, path analysis, and structural equation modeling (SEM) are full information maximum likelihood (FIML, also called direct ML) and multiple imputation (MI;

Little & Rubin, 2002; Schafer & Graham, 2002). These methods come from somewhat different traditions and can thus have important differences. For example, MI is usually performed within the Bayesian paradigm and involves sampling the values to be imputed from the posterior distributions of parameters. Additionally, MI effectively separates missing data treatment and data analysis, and can allow for different imputation and analysis models. Despite these differences, MI and FIML will frequently produce equivalent results (Allison, 2002; Graham, Olchowski, & Gilreath, 2007). The methods will tend to be equivalent when all of the following conditions are met: Noninformative priors are used in the MI imputation routine, the imputation and the analysis models are the same and equal to the data model assumed by FIML, and the number of imputations is large. Under these conditions, and if the assumed data model is correct (i.e., the data are multivariate normal), the FIML estimator is asymptotically fully efficient, whereas the MI estimator is only fully efficient when the number of imputations is infinite. Thus, unless one is in possession of strong prior information to conduct a Bayesian analysis, FIML is expected to produce better results.[1]

Even though MI is several decades old (Rubin, 1987), historically SEM researchers have tended to favor the FIML approach to missing data (Arbuckle, 1996; Jamshidian & Bentler, 1999; Muthén, Kaplan, & Hollis, 1987). This approach remains the only available modern missing data method in many SEM programs, although some, such as M*plus* (Muthén & Muthén, 2010), now implement both FIML and MI. Thus, the choice between FIML and MI is frequently made not on the basis of a theoretical preference for one method or the other, but on the basis of the availability of one or the other method in the program of choice. It is important, therefore, that researchers are familiar with how to obtain important missing data diagnostics via either method.

In this article, we show how to obtain an estimate of fraction of missing information for the jth model parameter, or λ_j, from the output of an SEM program that employs FIML rather than MI. Fraction of missing information, not to be confused with proportion of missing data, is an important diagnostic that communicates how much the estimation of a particular parameter is affected by nonresponse (Allison, 2002; Longford, 2005; Rubin, 1987). More specifically, it measures the inflation in the variance of the parameter estimate relative to what this variance would have been had all the data been observed. The concept of fraction of missing information is particularly important in MI because the relative efficiency of the MI estimator to the FIML estimator is determined by the quantity $(1 + \lambda_j/M)^{-1}$, where M is the number of imputations. An estimate of λ_j can therefore be useful in determining how many imputations are needed to achieve reasonable relative efficiency (Allison, 2002; Rubin, 1987; von Hippel, 2005).

Several authors have also suggested that λ_j is important in its own right as a diagnostic of the impact of missing data on estimation (Bodner, 2008; Davey, Savla, & Luo, 2005; Schafer, 1997), and should routinely be reported along with the analyses. In most scenarios, λ_j is a complicated function of the univariate and bivariate proportions of missing data, the specific nature of the missing data mechanism, the size of the correlations among the variables, and the chosen model parameterization. Thus, estimates of λ_j can serve as useful missing data diagnostics that capture the impact of missing data on estimation in the context of a particular

---

[1]Other Bayesian approaches to incomplete data in the context of SEM are possible (Lee, 2007; Song & Lee, 2002).

model of interest. It might be that a lot of data are missing, yet the estimation of a regression coefficient of interest, for example, is relatively unaffected. On the other hand, it could be that relatively few missing data values result in a large loss of efficiency in the estimation of the same regression coefficient. The measure λ_j communicates this loss of efficiency directly.

Some confusion exists in the applied literature on missing data as to whether estimates of λ_j can be obtained without MI. There are two reasons for this confusion: (a) published equations for λ_j usually involve MI-obtained variance components, creating the impression that the concept of fraction of missing information is unique to MI; and (b) programs that perform FIML do not typically report estimates of λ_j for parameters by default, whereas programs that perform MI frequently do. Some authors discuss λ_j in a way that suggests it is unique to MI. For instance, Schafer (2001) writes:

> Finally, an important advantage of multiple imputation over direct maximum likelihood is that it singles out missing data as a source of random variation distinct from ordinary sampling variability.... This partition immediately yields an estimated rate of missing information, which can be quite helpful for assessing the impact of missing data on inferences for any parameter of interest. (p. 361)

In a recent book on missing data, McKnight, McKnight, Sidani, and Figueredo (2007) wrote, "One of the advantages of MI is that it allows us to estimate the amount of missing information, or statistical uncertainty, resulting from our missing data. All other missing data handling procedures tell us only the amount of missing data" (p. 207). In fact, the concept of fraction of missing information originated in the development of the missing information principle (Orchard & Woodbury, 1972), which is grounded in likelihood theory, and obtaining estimates of λ_j from ML routines is in this sense more natural.

Curiously, although in the context of MI analyses the estimates of λ_j are computed and evaluated to ensure that a sufficient number of imputations have been used, these estimates themselves can be less than precisely estimated when the number of imputations is small. Bodner (2008) recently showed that the number of imputations necessary to produce reliable estimates of λ_j is higher than the number of imputations typically needed for the main analyses. To demonstrate this, Bodner used the cholesterol data set also analyzed by Ryan and Joiner (1994) and by Schafer (1997), containing a measure of cholesterol level at three time points after myocardial infarction for 28 patients. He found that the estimate of λ_j varied widely when the number of imputations was 3 to 10, which is the commonly recommended number (Rubin, 1987), and continued to display high variability even when the number of imputations was as high as 50 or 100. He concluded that "the minimum ms to achieve greater precision [in the estimate of λ_j] can be prohibitive" (p. 670). Yet, as the number of imputations goes to infinity, under the conditions outlined earlier, MI estimates converge to FIML estimates, and estimated standard errors from both methods are also asymptotically equivalent. Thus, an alternative to running hundreds of imputations to obtain estimates of λ_j is to obtain these estimates via FIML, as is illustrated in this article. For the cholesterol data set, Bodner (2008) reported that λ_j for the mean of the third time point, estimated from 800 imputations, was .19. Our FIML estimate is .18. Thus, obtaining FIML estimates of λ_j can be useful even when MI is preferred for the main analyses.

In this article, we first summarize the missing information principle that gave rise to the concept of fraction of missing information; we then provide the definition of fraction of missing information for individual parameters based on concepts from likelihood theory; we review the equivalent definition in MI terms; and we outline a simple procedure to obtain estimates of λ_j from an SEM program capable of running FIML estimation. We present an example illustrating the details of the proposed procedure in M*plus* (Muthén & Muthén, 2010), EQS (Bentler, 2008), and an R package for SEM called *lavaan* 0.4-14 (Rosseel, 2011, 2012). Using the same example, we also demonstrate the equivalence of the new FIML-based estimates of λ_j to the MI-based estimates based on a large number of imputations.

THE MISSING INFORMATION PRINCIPLE

In statistics, the concept of *information* refers to the amount of information available for inference about a particular parameter. The information available for inference is inversely related to the size of the standard error of the corresponding parameter estimate: The larger the standard error, the less we know about the parameter's true value. Orchard and Woodbury (1972) introduced the missing information principle, which states that information available from an incomplete data set is equal to complete information minus missing information. This idea allows us to conceptually define the fraction of missing information as the ratio of missing information to complete information.

The missing information principle is grounded in likelihood theory, and thus requires familiarity with maximum likelihood (ML) estimation. Let x_i be a set of scores for person i on p measures. It can be viewed as consisting of two components, $x_i = (y_i', z_i')'$, where y_i and z_i represent observed and missing scores, respectively. Viewing x as a vector random variable across people, the classic factorization of the density function for x that enables modern missing data analyses is $f(x|\theta) = f_1(y|\theta) f_2(z|\theta, y)$ (Little & Rubin, 2002; Orchard & Woodbury, 1972). That is, the density for the variables of interest x can be viewed as the product of the marginal density of the observed components y and the conditional density of the missing components z given the observed components (Dempster, Laird, & Rubin, 1977; Orchard & Woodbury, 1972). It follows that the likelihood function for a sample of size n can similarly be factored as follows:

$$L(\theta|X) = L_1(\theta|Y) f_2(Z|Y, \theta), \tag{1}$$

where X, Y, and Z represent all of the complete, observed, and missing data for n observations, respectively. Equation 1 states that the likelihood of the complete data is equal to the likelihood of observed data times the conditional density of the missing data given observed data.

From Equation 1, it follows that the score vector, or the derivative of the log-likelihood, is then partitioned as follows:

$$\frac{\partial \log L(\theta|X)}{\partial \theta'} = \frac{\partial \log L_1(\theta|Y)}{\partial \theta'} + \frac{\partial \log f_2(Z|Y, \theta)}{\partial \theta'} \tag{2}$$

The information matrix can be defined as either the negative expected value of the second derivative of the log-likelihood or as the covariance matrix of the score vector (Rao, 2002).

Following the second definition, the information about the parameters θ that would be available from complete data is the covariance of the left side in Equation 2: $J_X = \text{cov}(\frac{\partial \log L(\theta|X)}{\partial \theta'})$. The information about θ that would be available from observed data is the covariance matrix of the first term on the right side of Equation 2: $J_Y = \text{cov}(\frac{\partial \log L_1(\theta|Y)}{\partial \theta'})$. Additional information about θ that would be available from missing data, over and above information contained in the observed data, is the covariance matrix of the second term on the right side of Equation 2: $J_{X|Y} = J_{Z|Y} = \text{cov}(\frac{\partial \log f_2(Z|Y,\theta)}{\partial \theta'})$. One can write $J_{X|Y} = J_{Z|Y}$ because the information available in the complete data over and above the information in observed data is the same as the missing information. If missing data have nothing to add to the estimation of θ, this quantity would consist of zeros. The missing information principle of Orchard and Woodbury (1972) can then be stated in equation form as follows:

$$J_X = J_Y + J_{X|Y}. \tag{3}$$

It states that complete information is the sum of observed information and missing information (for proof, see Little & Rubin, 2002; Orchard & Woodbury, 1972). Equation 3 implies that $J_X \geq J_Y$; that is, that the amount of information about θ contained in complete data is greater than or equal to the information about θ contained in incomplete data.[2]

If all of the data had been observed, then $Y = X$ and Z is null, and the ML parameter estimates, which we call $\hat{\theta}_{ML}$, would be obtained by maximizing the complete data likelihood function $L(\theta|X)$. In this case, and under standard regularity conditions, the asymptotic distribution of $\hat{\theta}_{ML}$ is normal with a covariance matrix equal to the inverse of the corresponding complete data information matrix:

$$a \, \text{cov}(\sqrt{n}\hat{\theta}_{ML}) = J_X^{-1}$$

(Rao, 2002). If the data are incomplete, the corresponding ML (also called FIML) estimates, which we call $\hat{\theta}_{FIML}$, would be obtained by maximizing the incomplete data likelihood function $L_1(\theta|Y)$, which produces consistent estimates as long as the missing data mechanism is either missing completely at random (MCAR) or missing at random (MAR; Little & Rubin, 2002). The asymptotic distribution of $\hat{\theta}_{FIML}$ is normal with a covariance matrix equal to the inverse of the corresponding incomplete data information matrix,

$$a \, \text{cov}(\sqrt{n}\hat{\theta}_{FIML}) = J_Y^{-1}.$$

From Equation 3, $J_X^{-1} \leq J_Y^{-1}$, and in particular, the diagonal elements of J_Y^{-1} are larger than the diagonal elements of J_X^{-1}, implying that standard errors are larger when the data are incomplete. This is rather intuitive: To the extent that missing data contains useful information about θ, incomplete data contains less information about θ and more uncertainty exists around estimated parameters. This uncertainty manifests itself in greater variability of $\hat{\theta}_{FIML}$ relative to $\hat{\theta}_{ML}$.

---

[2] The notation indicating that one matrix is "greater" than the other is used here in the standard sense that their difference is nonnegative definite.

FRACTIONS OF MISSING INFORMATION (λ_j)

The missing information principle makes clear that the relative decrease (loss of information) in the size of the components of J_Y relative to the corresponding components of J_X captures the extent to which missing data affect the quality of parameter estimates. When the data are missing, an estimate of J_Y, the incomplete data information matrix, can be obtained either analytically or numerically. When data are missing, it is also possible to obtain an estimate of J_X (using either MI or ML procedures, as we show later), which indicates how much information would have been contained in the data if there were no missing values. Thus, estimates of both J_X and J_Y can be computed from the data and their relative size evaluated. Global measures of missing information exist that take into account the entire structure of the two information matrices (Savalei & Rhemtulla, 2011); for example, the largest eigenvalue of the matrix product $J_{X|Y} J_X^{-1} = I - J_Y J_X^{-1}$ has been termed *the largest fraction of missing information*[3] (Rubin, 1987). However, global measures are difficult to interpret, and methodologists have instead advised practitioners to report estimates of the fraction of missing information for individual parameters (Bodner, 2008; Davey et al., 2005; Enders, 2010; Schafer, 1997), defined next.

The diagonal elements of J_X^{-1} and J_Y^{-1} contain parameter estimates' variances when these estimates are obtained from complete and incomplete data, respectively. The definition of *fraction of missing information* for a given parameter θ_j is

$$\lambda_j = 1 - \frac{\{J_X^{-1}\}_{jj}}{\{J_Y^{-1}\}_{jj}} = 1 - \frac{SE_{j,C}^2}{SE_{j,O}^2}, \tag{4}$$

where $SE_{j,C}$ is the asymptotic standard error of $\hat{\theta}_{j,ML}$, the ML estimate of θ_j based on complete data, and $SE_{j,O}$ is the asymptotic standard error of the FIML estimate $\hat{\theta}_{j,FIML}$ based on incomplete data. Thus, λ_j represents the relative loss of efficiency in the estimate of θ_j due to incomplete data.

When θ contains a single parameter, Equation 4 can be written directly as the ratio of incomplete to complete information: $\lambda = 1 - \frac{J_Y}{J_X} = 1 - \frac{SE_C^2}{SE_O^2}$, because in this case J_Y and J_X are scalars and their inverses are just reciprocals. This equation is given in some textbooks (e.g., Longford, 2005, p. 55), but one should be careful not to generalize it to problems involving multiple parameters. When J_Y and J_X are matrices, $\lambda_j \neq 1 - \frac{\{J_Y\}_{jj}}{\{J_X\}_{jj}}$, in general, unless the parameter estimates are uncorrelated. Equation 4 must be used to define fractions of missing information per parameter in the multiparameter case, which requires that the information matrices be inverted prior to taking the ratio of the diagonal elements.

ESTIMATING λ_j VIA MULTIPLE IMPUTATION

The previous section defined fraction of missing information using concepts from likelihood theory. However, estimates of λ_js are typically obtained as by-products of MI. The connection between this computation and the theoretical definition in Equation 4 might not be obvious.

---

[3]This number actually captures the worst λ_j for any linear combination of the parameters (e.g., Fraley, 1999).

Before we proceed to define FIML-based estimates of λ_j, we briefly review the MI computation and point out that it also estimates Equation 4.

In MI, M complete data sets are created from the incomplete data and the results are averaged across the data sets to obtain parameter estimates and standard errors (Rubin, 1987; Schafer, 1997). The subsequent analysis typically relies on Bayesian arguments, but under the conditions listed in the introduction, the results obtained from MI become increasingly similar to FIML as the number of imputations increases (Allison, 2002; Graham et al., 2007). Furthermore, even when informative priors are used, the MI and FIML approaches become equivalent when, additionally, the sample size $N \to \infty$, by the well-known principle that the data overwhelm the prior.

The MI estimate of θ_j is obtained by pooling complete data estimates across M imputations: $\hat{\theta}_{MI,M} = \frac{1}{M} \sum_{m=1}^{M} \hat{\theta}_m$, where $\hat{\theta}_m$ is the parameter estimate obtained from the mth imputed data set (parameter subscript j is omitted here for readability). The standard error for $\hat{\theta}_{MI,M}$ is obtained by combining within- and between-imputation variance. The within-imputation variance is the average of the squared complete data standard error estimates $\hat{s}_m$ across the M imputations: $\hat{s}_{MI,M}^2 = \frac{1}{M} \sum_{m=1}^{M} \hat{s}_m^2$. Intuitively, $\hat{s}_{MI,M}^2$ is an estimate of what the variance of the parameter estimate would be if there were no missing data. The between-imputation variance is the variance of the parameter estimates $\hat{\theta}_m$ across imputations: $\hat{B}_{MI,M} = \frac{1}{M-1} \sum_{m=1}^{M} (\hat{\theta}_m - \hat{\theta}_{MI,M})^2$. This quantity will be larger to the extent that more information about the parameter is lost due to missing data. Then, the estimated total variance of $\hat{\theta}_{MI,M}$ is a combination of the within- and between-imputation variance: $\widehat{SE}^2(\hat{\theta}_{MI,M}) = \hat{T}_{MI,M} = \hat{s}_{MI,M}^2 + \frac{M+1}{M} \hat{B}_{MI,M} \approx \hat{s}_{MI,M}^2 + \hat{B}_{MI,M}$, where the latter approximation holds when M is large (Rubin, 1987; Schafer, 1997). At the limit (when $M = \infty$), $\hat{T}_{MI,\infty} = \hat{s}_{MI,\infty}^2 + \hat{B}_{MI,\infty}$.

The definition of fraction of missing information for the jth parameter involving variance components from multiple imputation is as follows:

$$\lambda_j = \frac{B_j}{s_j^2 + B_j} = 1 - \frac{s_j^2}{T_j}, \tag{5}$$

where $T = s^2 + B$ are the limits of the preceding variance components as $N \to \infty$ and $M \to \infty$. Conceptually, it is between-imputation variance over total variance. A sample estimate of this quantity is given by Schafer and Graham (2002):

$$\hat{\lambda}_{j,M} = \frac{\dfrac{M+1}{M} \hat{B}_{MI,M}}{\hat{s}_{MI,M}^2 + \dfrac{M+1}{M} \hat{B}_{MI,M}} \approx \frac{\hat{B}_{MI,M}}{\hat{s}_{MI,M}^2 + \hat{B}_{MI,M}}, \tag{6}$$

where the latter approximation works when the number of imputations is large. Other sample estimates exist, which involve a numerator correction that depends on the degrees of freedom for a t distribution that is the posterior distribution of $\hat{\theta}_{MI,M}$ (Schafer, 1997, Equation 4.30). We do not state this more complicated sample definition here because it becomes equivalent to Equation 6 as $M \to \infty$.

We have used the same notation, λ_j, in the population definitions in Equations 4 and 5 because these quantities are in fact the same, when MI is done under the same assumptions

as FIML (as listed earlier). The total asymptotic variance T_j is the same as the asymptotic variance of $\hat{\theta}_{FIML}$, $T_j = SE_{j,O}^2$. The variance had the data been complete, s_j^2, is the same as the asymptotic variance of $\hat{\theta}_{ML}$, the complete data estimator, $s_j^2 = SE_{j,C}^2$. Finally, the between-imputations variance component B_j is the "amount of missing information." The equivalence of the asymptotic variance components from MI and the corresponding quantities involved in the missing information principle from likelihood theory was pointed out by Rubin (1987).

We showed earlier that the MI and the likelihood-based population definitions in Equations 4 and 5 are equivalent. The population definitions can be interpreted as describing the situation when $M \rightarrow \infty$ and $N \rightarrow \infty$. When N is finite but $M \rightarrow \infty$, the MI estimator based on infinitely many imputations will converge to the FIML estimator, $\hat{\theta}_{MI,\infty} \rightarrow \hat{\theta}_{FIML}$. The total variance $\hat{T}_{MI,\infty}$ also approaches the FIML-estimated variance of $\hat{\theta}_{FIML}$. In fact, when $M = \infty$, the MI estimate of fraction of missing information $\hat{\lambda}_{j,\infty} = \frac{\hat{B}_{MI,\infty}}{\hat{s}_{MI,\infty}^2 + \hat{B}_{MI,\infty}}$ is equivalent to the FIML-based estimate $\hat{\lambda}_{ML,j}$ that we are about to propose. But this FIML estimate is much easier to obtain than the MI-based estimate $\hat{\lambda}_{j,\infty}$ because it does not require running a huge number of MIs.

AN ESTIMATE OF λ_j VIA FIML

The population definition of fraction of missing information given in Equation 4 suggests that an estimate of λ_j for each parameter can be computed from the output of any computer program capable of running complete and incomplete data ML (FIML) estimation. This procedure does not seem to be known in the applied literature. Briefly, it involves running the SEM program twice: once using the FIML routine on the original incomplete data set, and once using the complete-data ML routine on the model-implied covariance matrix and vector of means obtained from the FIML run. One minus the ratio of the corresponding squared standard errors from each output forms an estimate of λ_j defined by Equation 4. We label this estimate $\hat{\lambda}_{j,ML}$. The details of the procedure are now given, specific to three programs: EQS, M*plus,* and the R package *lavaan*.[4] Any other program capable of running ML and FIML can in principle be used.

Step 1

Fit the model to the original incomplete data set using FIML. Observed information should be requested to obtain consistent estimates of information for MAR data (Kenward, Lesaffre, & Molenberghs, 1994; Kenward & Molenberghs, 1998; Savalei, 2010; Schafer & Graham, 2002). The assumption of MCAR is almost never tenable unless missingness is planned (Graham, 2009; Little, 1988). Observed information is the default in M*plus,* EQS,[5] and *lavaan* when FIML is requested. Sample syntax for all three programs is given in the Appendix. For each parameter of interest, obtain its estimated standard error from the output, $\widehat{SE}_{j,O}$.

---

[4]The method for obtaining FMI estimates described in this paper will be implemented as a command in *lavaan* before the end of 2012; for details, please see lavaan.org.

[5]The default is observed information in EQS only when the model is built using EQS Model Builder, and ML is requested to treat missing data. When raw syntax is created, the user must specify SE=OBSERVED or SE=EXACT for numeric or analytic raw second derivatives, respectively. Omitting the SE command will default to Fisher standard errors.

Step 2

From the output in Step 1, take the model-implied covariance matrix, $\hat{\Sigma}$, and mean vector, $\hat{\mu}$. In EQS, this output is requested by specifying COV=YES under the /PRINT command, as illustrated in the sample syntax in the Appendix. Figure 1a shows what this matrix looks like in the output file. For large models, it might be desirable to save these estimates in a file instead; the alternate syntax is also shown in the Appendix. The estimates $\hat{\mu}$ and $\hat{\Sigma}$ can be requested by specifying OUTPUT: residual, as shown in the Appendix. Figure 1b shows what this output looks like in the output file. For large models, estimates $\hat{\Sigma}$ (but not $\hat{\mu}$) can be saved in a data file using the SAVEDATA command, as shown in the Appendix. In *lavaan,* the command fitted.values() produces the model-implied means and covariances, as shown in the Appendix. Most other SEM software packages have options for obtaining the model-implied means and covariance matrix, either as part of the output or in a separate data file.

It is worth noting that when the mean structure is saturated (e.g., if the model is a confirmatory factor model), the model-implied means $\hat{\mu}$ are equal to the "saturated ML" mean estimates, which are typically printed in the output by default. If the entire model is saturated (e.g., if the model is a regression model), $\hat{\mu}$ and $\hat{\Sigma}$ are equal to the "saturated ML" mean and covariance matrix estimates typically printed in the output, also known as the EM means and covariances (Enders & Peugh, 2004; Graham, 2003). In this case, these quantities can be copied directly from the output files.

Step 3

Run the same model as in Step 1 using the FIML estimates of the model-implied means and covariance matrix $\hat{\mu}$ and $\hat{\Sigma}$ from Step 2 as input into the complete data ML routine. The sample size should be equal to the original sample size (i.e., the total number of cases, regardless of whether they have incomplete data). For each parameter, obtain its estimated standard error from the output, $\widehat{SE}_{j,C}$. It is worth noting that some programs (in particular, M*plus* and the default setting in *lavaan*) assume that a user-provided covariance matrix is scaled by $N-1$ rather than N. This difference is of little consequence; however, if the sample size is small (e.g., $N < 100$), one could adjust for it by multiplying the input matrix $\hat{\Sigma}$ by $N/N-1$ before running Step 3. For the best comparison across programs, we apply this adjustment to the computations in M*plus* (the adjusted M*plus* input file is shown in Figure 2) and in *lavaan* (see code in the Appendix).

Step 4

Compute $\hat{\lambda}_{j,ML}$ according to Equation 4, for each parameter of interest.

EXAMPLE

We now provide an empirical example that illustrates this procedure. To verify the asymptotic equivalence of our procedure and the MI procedure, we also compare the estimates $\hat{\lambda}_{j,ML}$ in this example to estimates obtained via MI with 50,000 imputations.
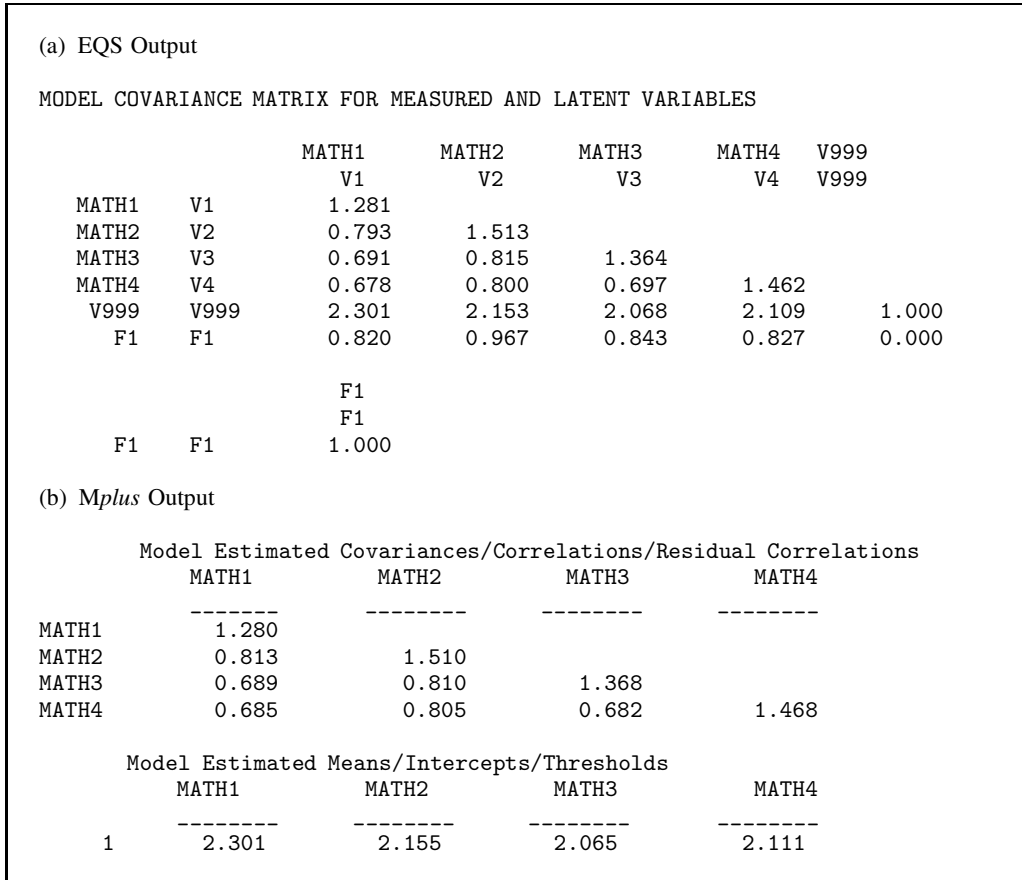
```
(a) EQS Output

MODEL COVARIANCE MATRIX FOR MEASURED AND LATENT VARIABLES

                       MATH1       MATH2       MATH3       MATH4     V999
                       V1          V2          V3          V4        V999
      MATH1    V1       1.281
      MATH2    V2       0.793       1.513
      MATH3    V3       0.691       0.815       1.364
      MATH4    V4       0.678       0.800       0.697       1.462
       V999    V999     2.301       2.153       2.068       2.109     1.000
         F1    F1       0.820       0.967       0.843       0.827     0.000

                       F1
                       F1
         F1    F1       1.000

(b) Mplus Output

        Model Estimated Covariances/Correlations/Residual Correlations
             MATH1          MATH2          MATH3          MATH4
             -------        --------       --------       --------
MATH1          1.280
MATH2          0.813          1.510
MATH3          0.689          0.810          1.368
MATH4          0.685          0.805          0.682          1.468

        Model Estimated Means/Intercepts/Thresholds
             MATH1          MATH2          MATH3          MATH4
             --------       --------       --------       --------
     1         2.301          2.155          2.065          2.111
```

FIGURE 1   Model covariance matrix in EQS and M*plus* output for Example 2. Model-estimated means in EQS (a) are the first four values in the row labeled V999.
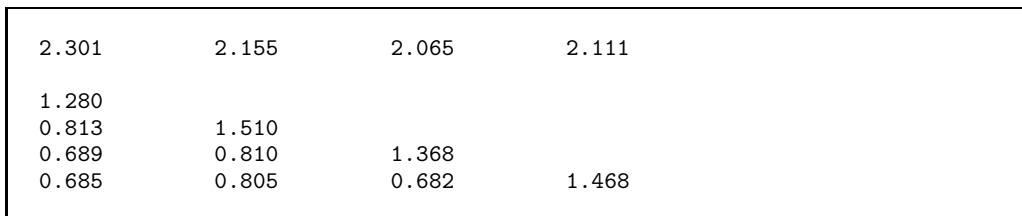
```
2.301          2.155          2.065          2.111

1.280
0.813          1.510
0.689          0.810          1.368
0.685          0.805          0.682          1.468
```

FIGURE 2   Contents of means and covariance matrix file to read in to M*plus* for complete data routine in Example 2, 'math.model.dat.' The first line of numbers is obtained by taking the model-implied means from the M*plus* output file in Step 1; the next four lines are obtained by taking the model-implied covariance matrix from the same output file (see Figure 1b). If the sample size is small ($N < 100$), one can multiply the elements in the last four lines by $(N/N - 1)$ for slightly better estimates. This file should be saved as a .dat file.

The data set is a subsample of $N = 500$ cases from the public use portion of the National Longitudinal Study of Adolescent Health data set (Harris, 2009). This study followed a large sample of adolescents for 4 years, measuring a wide range of psychological, social, behavioral, and contextual variables, with the goal of predicting health outcomes. We used four variables corresponding to math grades in the four consecutive years of the study (we call these *math1* to *math4*) to illustrate the four-step procedure for $\hat{\lambda}_{j,ML}$ calculations. Missingness was present on all four variables, and as is typical with longitudinal data, the amount of missing data increased over time. Six cases had missingness on all four variables; thus, the effective sample size was $N = 494$. In addition, there were about 2% missing data on *math1*, 7% missing on *math2*, 16% missing on *math3*, and 40% missing on *math4*. Forty-five percent of cases had missing data on at least one variable, and there were 10 missing data patterns. We fit a one-factor confirmatory factor analysis (CFA) model to the data. The variance of the factor was fixed to one for identification; the parameters estimated were four intercepts, four factor loadings, and four error variances.

We followed Steps 1 through 4 given earlier in three SEM software programs: EQS, M*plus,* and *lavaan.* The syntax for each of these programs is given in the Appendix. Standard errors from Steps 1 and 3 and the $\hat{\lambda}_j$ estimates for the intercepts and loadings of the first and fourth time points (i.e., *math1* and *math4*) are presented in the first three columns of Table 1. Estimates across software matched within .01. Slight differences are likely due to different computational approaches to obtained observed information matrices with incomplete data, as well as to varying defaults for type of information with complete data.

TABLE 1
Fraction of Missing Information Estimates for Selected Parameters for the Example

| Parameter | Estimate | FIML | | | MI |
		EQS	Mplus	lavaan	Mplus
math1	$\widehat{SE}_{j,O}$	.05129	.05123	.05123	.05124
Intercept	$\widehat{SE}_{j,C}$	.05097	.05091	.05091	.05092
	$\hat{\lambda}_j$	**.01260**	**.01241**	**.01241**	**.01268**
math4	$\widehat{SE}_{j,O}$	.06577	.06588	.06588	.06603
Intercept	$\widehat{SE}_{j,C}$	.05446	.05452	.05452	.05449
	$\hat{\lambda}_j$	**.31423**	**.31506**	**.31506**	**.31899**
math1	$\widehat{SE}_{j,O}$	.05108	.05026	.05026	.05099
Loading	$\widehat{SE}_{j,C}$	.04844	.04813	.04813	.04840
	$\hat{\lambda}_j$	**.10080**	**.08321**	**.08321**	**.09908**
math4	$\widehat{SE}_{j,O}$	.06537	.06574	.06574	.06601
Loading	$\widehat{SE}_{j,C}$	.05255	.05261	.05261	.05256
	$\hat{\lambda}_j$	**.35380**	**.35948**	**.35948**	**.36608**

Note. FIML $=$ full information maximum likelihood; MI $=$ multiple imputation based on 50,000 imputations. $\widehat{SE}_{j,O} =$ standard error estimate taking into account missing data; $\widehat{SE}_{j,C} =$ standard error estimate assuming complete data; $\hat{\lambda}_j =$ estimate of fraction of missing information. The model is a four-indicator confirmatory factor analysis with math grade point average (GPA) measured at four time points: *math1* is math GPA at the first time point; *math4* is math GPA at the fourth time point. Values from M*plus* and *lavaan* have adjusted $\widehat{SE}_{j,C}$ to use the multiplier $N-1$, for better comparison across programs and methods.

To confirm that the estimates of fraction of missing information obtained from multiple imputation, $\hat{\lambda}_{j,M}$, are equivalent to the FIML-based estimates $\hat{\lambda}_{j,ML}$ when the number of imputations is large, we used M*plus* to perform multiple imputation with $M = 50,000$ on the same data set. The quantities $\hat{\lambda}_{j,M}$ were computed using the large-M approximation in Equation 6. The results are presented in the last column of Table 1. The MI estimates based on a large M are within about .01 of the FIML estimates.

INTERPRETATION OF λ_j

We now make a few comments about the interpretation of λ_j. In statistics, the relative efficiency of an estimator as compared to some standard, typically to a fully efficient estimator, is defined as the ratio of the respective variances. If this ratio is 1, we would say that the estimator is 100% efficient; if this ratio is .98, we would say that a particular estimator is 98% efficient. For example, the sample median has 64% efficiency relative to the sample mean in large samples. Relative efficiencies are frequently used to compare estimators in simulation studies (e.g., Savalei & Bentler, 2009). Equation 4 reveals that λ_j is one minus the relative efficiency, and thus can be interpreted as the loss of efficiency in the estimation of a particular parameter as a result of missing data.[6] When the data are complete, $\lambda_j = 0$, and the incomplete data estimator is 100% efficient relative to what it would have been had the data been complete. When the data contain zero information about a parameter (e.g., if the parameter is a covariance but the two variables have no jointly observed cases), $\lambda_j = 1$. We can interpret any other value of the fraction of missing information; for example, if $\lambda_j = 0.2$, then the loss of efficiency due to incomplete data is 2%, or equivalently, the incomplete data estimator is 98% efficient relative to what it would have been with complete data.

Another approach to interpreting λ_j is to relate it to the relative inflation of the width of the confidence interval that results from incomplete data. First, most directly, we can define the effective sample size N_j^* as $N_j^* = N(1-\lambda_j)$. This is the sample size that would have achieved the same efficiency for the jth parameter with complete data. For instance, the intercept for *math4* in Table 1 is based on the effective size $N_j^* = 494(1 - .31506) \approx 338$ (using values from M*plus*), which means that its variability is as high as it would have been had it been based on a complete data set with only 338 cases instead of 494. In this sense, λ_j reflects loss of statistical power due to missing data. Equivalently, one could define target sample size as $N_j^* = N/(1 - \lambda_j)$, which communicates how much larger the sample size would have to be to fully account for incomplete data as it affects the jth parameter. In the preceding example, $N_j^* = 494/(1 - .31506) \approx 721$, which means that to achieve the same efficiency in the estimation of the jth parameter as one would have had with complete data, it would require 721 cases with incomplete data!

A second way to relate λ_j to the relative inflation of the width of the confidence interval is to define the *width inflation factor,* as follows: $WIF_j = \frac{1}{\sqrt{1-\lambda_j}} = \frac{SE_{j,O}}{SE_{j,C}}$. The relationship between WIF_j and λ_j is illustrated in Figure 3. When $\lambda_j = .56$, the confidence interval for the jth parameter is 1.5 times wider than what it would have been had the data been complete.

---

[6]To be more precise, we are actually talking about asymptotic efficiency here.
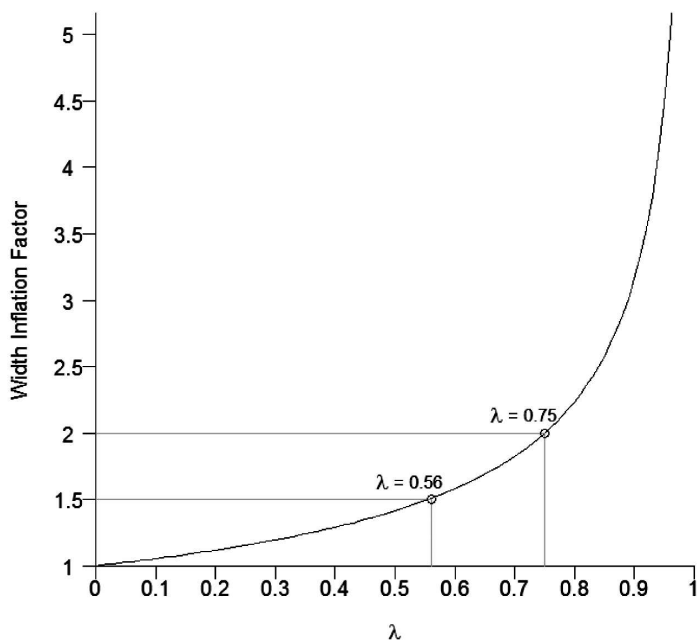
FIGURE 3   The relationship between fraction of missing information λ and the confidence interval width inflation factor (WIF), which is the ratio of the standard error for incomplete data over the standard error for complete data. Values corresponding to 50% inflation (WIF = 1.5) and 100% inflation (WIF = 2) are shown.

When $\lambda_j = .75$, the confidence interval is 2 times wider. In small to moderate sample sizes, $\lambda_j > .55$ for any parameter of interest could present a cause for concern, particularly if the estimated parameter is not significant. Lack of significant results in this case can potentially be explained by missing data. Notice that the function in Figure 3 increases very quickly for higher values of λ_j, but also that for $\lambda_j \leq .2$ the inflation is very small and probably harmless except for very weak effects. To emphasize, this analysis applies to small to medium sample sizes only. If, for example, one has a longitudinal data set based on 5,000 cases and with large amounts of missing data, it is quite likely that FIML estimation will be successful and parameters of interest will be significant, even if $\lambda_j = .8$ for some parameters. In general, because the actual value of λ_j for any given parameter depends on multiple factors, the rough guidelines provided earlier to aid in its interpretation should not be misinterpreted as strict cutoff values. We instead recommend treating λ_j as an effect size measure of the impact on missing data on the estimation of the jth parameter.

Although the preceding discussion centered on parameter estimates, SEM model test statistics will also be adversely affected by missing data and will have reduced power to detect misspecification. Global measures of missing information that can succinctly summarize the extent of this effect are the subject of future research.

Finally, we briefly discuss characteristics of the data and the model affecting λ_j. As the example illustrates (see Table 1), λ_j can be lower than the rate of missing data for the corresponding variable; for instance, $\lambda_j = .32$ for the intercept of *math4*, but *math4* has 40%

missing values. However, λ_j can also be higher than the rate of missing data. With MCAR data, the values of λ_j will be bounded by the largest rates of univariate and pairwise missingness, and will vary based on the model parameterization and intercorrelations among variables. However, with MAR data, λ_j can be significantly higher than the rate of missing data. For instance, if the data set has two variables (X_1 and X_2), and X_2 is missing 20% of its values either randomly (MCAR) or corresponding to the highest values of X_1 (MAR), the corresponding λ_js for intercept and slope when X_2 is predicted from X_1 are both .2 when the data are MCAR, but are .34 and .53, respectively, when the data are MAR (Savalei & Rhemtulla, 2011). In other words, the regression parameters, and particularly the slope, are estimated with much less precision when the data are MAR. The relationship between λ_j and the intercorrelations among the variables is more complicated and depends on the precise parameterization of the model. These relationships will be explored in future research.

DISCUSSION

Fraction of missing information is a useful concept in missing data analysis because it quantifies how much a particular parameter estimate has been affected by missing data. This quantity is quite different from the proportion of missing data per variable, as it also depends on the type of missing data mechanism, the model parameterization, and the degree of interrelationship among the variables. The fraction of missing information can differ from parameter to parameter in the same data set; for example, estimates of a factor loading might be more affected by certain types of missing data than estimates of an intercept. It would be excellent practice for researchers to compute and report fraction of missing data for each parameter (Bodner, 2008; Enders, 2010; Schafer, 1997). Such a practice would allow researchers to begin to gauge the effect that missingness has on their analyses, and could inform future research design. If it is known, for example, that estimation of a particular key parameter is highly affected by missing data, researchers could make efforts to collect more data to inform its estimation.

Bodner (2008) found that the number of imputations in an MI analysis might have to be prohibitively large to achieve accurate estimates of fraction of missing information. To avoid doing hundreds of imputations, Bodner suggested that researchers could obtain a crude conservative estimate of λ_j by simply computing the proportion of cases in the entire sample on which there is missingness on any variable related to the parameter of interest. This conservative estimate could then be used to estimate the number of imputations required to obtain acceptable parameter and standard error estimates (i.e., by consulting his Table 3). Harel (2007) derived a standard error of the MI estimate of λ_j with respect to the number of imputations to better evaluate how variable the MI estimate is given a certain number of imputations.

In this article, we have shown that a single FIML analysis followed by a complete data analysis based on the FIML-estimated model-implied means and covariance matrix results in nearly identical estimates of λ_j to those based on 50,000 imputations. These estimates can be obtained from any SEM program capable of performing ML estimation with complete and incomplete data. Thus, the method presented here directly addresses the concerns raised by Bodner (2008) with regard to the quality of MI-based estimates of fraction of missing information based on small M, and obviates the need for confidence intervals to evaluate the impact of small M. Accurate estimates of λ_j are in fact readily available. Furthermore, it

would be straightforward for the developers of SEM software to automate their computation, so that estimates of fraction of missing information per parameter are printed whenever FIML is requested. This would encourage the use and reporting of these estimates by applied researchers, as well as further study of their properties by methodologists.

ACKNOWLEDGMENTS

REFERENCES

Allison, P. D. (2002). *Missing data.* Thousand Oaks, CA: Sage.

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Erlbaum.

Bentler, P. M. (2008). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.

Bodner, T. E. (2008). What improves with increasing missing data imputations? *Structural Equation Modeling, 15,* 651–675.

Davey, A., Savla, J., & Luo, Z. (2005). Issues in evaluating model fit with missing data. *Structural Equation Modeling, 12,* 578–597.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39,* 1–38.

Enders, C. K. (2010). *Applied missing data analysis.* New York, NY: Guilford.

Enders, C. K., & Peugh, J. L. (2004). Using an EM covariance matrix to estimate structural equation models with missing data: Choosing an adjusted sample size to improve the accuracy of inferences. *Structural Equation Modeling, 11,* 1–19.

Fraley, C. (1999). On computing the largest fraction of missing information for the EM algorithm and the worst linear function for data augmentation. *Computation Statistics & Data Analysis, 31,* 13–26.

Graham, J. W. (2003). Adding missing-data-relevant variables to FIML-based structural equation models. *Structural Equation Modeling, 10,* 80–100.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8,* 206–213.

Harel, O. (2007). Inferences on missing information under multiple imputation and two-stage multiple imputation. *Statistical Methodology, 4,* 75–89.

Harris, K. M. (2009). The National Longitudinal Study of Adolescent Health (Add Health), Waves I & II, 1994–1996; Wave III, 2001–2002; Wave IV, 2007–2009 [Machine readable data file and documentation]. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill.

Jamshidian, M., & Bentler, P. M. (1999). ML estimation of mean and covariance structures with missing data using complete data routines. *Journal of Educational and Behavioral Statistics, 24,* 21–41.

Kenward, M. G., Lesaffre, E., & Molenberghs, G. (1994). An application of maximum-likelihood and generalized estimating equations to the analysis of ordinal data from a longitudinal study with cases missing at random. *Biometrics, 50,* 945–953.

Kenward, M. G., & Molenberghs, G. (1998). Likelihood based frequentist inference when data are missing at random. *Statistical Science, 13,* 236–247.

Lee, S. Y. (2007). *Structural equation modeling: A Bayesian approach.* Chichester, UK: Wiley.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198–1202.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.

Longford, N. T. (2005). *Missing data and small-area estimation: Modern analytical equipment for the survey statistician.* New York, NY: Springer.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction.* New York, NY: Guilford.

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

Muthén, L. K., & Muthén, B. O. (2010). *M*plus *user's guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Orchard, T., & Woodbury, M. A. (1972). A missing information principle: Theory and applications. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1*, 697–715.

Rao, C. R. (2002). *Linear statistical inference and its applications.* New York, NY: Wiley.

Rosseel, Y. (2011). lavaan. Retrieved from lavaan.ugent.be

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48, 3*, 1–36.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* Hoboken, NJ: Wiley.

Ryan, B., & Joiner, B. (1994). *Minitab handbook* (3rd ed.). Belmont, CA: Wadsworth.

Savalei, V. (2010). Expected vs. observed information in SEM with incomplete normal and nonnormal data. *Psychological Methods, 15,* 352–367.

Savalei, V., & Bentler, P. M. (2009). A two-stage approach to missing data: Theory and application to auxiliary variables. *Structural Equation Modeling, 16,* 477–497.

Savalei, V., & Rhemtulla, M. (2011, July). *Some explorations of the local and global measures of missing information.* Paper presented at the 76th annual and the 17th international meeting of the Psychometric Society, Hong Kong.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data.* London: Chapman & Hall.

Schafer, J. L. (2001). Multiple imputation with PAN. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 357–377). Washington, DC: American Psychological Association.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.

Song, X. Y., & Lee, S. Y. (2002). Analysis of structural equation model with ignorable missing continuous and polytomous data. *Psychometrika, 67,* 261–288.

von Hippel, P. T. (2005). How many imputations are needed? A comment on Hershberger and Fisher (2003). *Structural Equation Modeling, 12,* 334–335.

APPENDIX
SYNTAX FOR EXAMPLE 2

EQS Syntax for Step 1 (FIML)

```
/TITLE
Add Health Math Example, FIML
/SPECIFICATIONS
 DATA='math.dat'; VARIABLES = 4; CASES = 500;
 METHOD=ML; ANALYSIS=MOMENT; MISSING=ML; SE=EXACT; MATRIX=RAW;
/LABELS
 V1=MATH1; V2=MATH2; V3=MATH3; V4=MATH4;
/EQUATIONS
 V1 = *V999 + *F1 + E1;
```

```
 V2 = *V999 + *F1 + E2;
 V3 = *V999 + *F1 + E3;
 V4 = *V999 + *F1 + E4;
/VARIANCES
 F1 = 1;
 E1 TO E4 = *;
/PRINT
 COV=YES;  !this command requests the model-implied cov matrix and means
           !to be printed in the output
/OUTPUT
 data="modelcovmeans.ets"; SI; !This command saves Sigma, the model-
  !implied cov matrix (and means) in a separate datafile that can be
  !opened in EQS (this is not necessary unless more significant
  !digits are desired). This file will contain other information in
  !addition to the saved means and covariance matrix, and will thus
  !have to be edited before it can be used as an input data file.
  !Please see EQS manual for details
/END
```

EQS Syntax for Step 3 (Complete Data Routine)

```
/TITLE
Add Health Math Example, complete data routine
/SPECIFICATIONS
VARIABLES= 4; CASES= 494; MATRIX=COV;
!A datafile containing the covariance matrix and means can be specified
!here instead to replace /MEANS and /MATRIX sections below
METHOD=ML; ANALYSIS=MOMENT;
/EQUATIONS
V1 = *V999 + *F1 + E1;
V2 = *V999 + *F1 + E2;
V3 = *V999 + *F1 + E3;
V4 = *V999 + *F1 + E4;
/VARIANCES
F1 = 1;
E1 TO E4 = *;
/MATRIX  !the model-implied cov matrix from the FIML run
1.2806817
0.7926411 1.5128352
0.6910960 0.8153024 1.3636522
0.6780093 0.7998636 0.6973933 1.4623119
/MEANS !the model-implied means from the FIML run
2.3012734 2.1533217 2.0676202 2.1092750
/END
```

M*plus* Syntax for Step 1 (FIML)

```
TITLE: math example, 1-factor, FIML
DATA: FILE IS "mplus.math.dat";
VARIABLE: NAMES ARE MATH1 MATH2 MATH3 MATH4;
MISSING ARE ALL (-9999);
ANALYSIS: ESTIMATOR = ML;
```

```
ALGORITHM = EM;
MODEL:
F1 BY MATH1* MATH2 MATH3 MATH4;
F1@1;

OUTPUT: RESIDUAL;  !this command requests model-implied means and
                   !covariances to be printed in the .out file

!SAVEDATA: COVARIANCE IS 'modelcov.dat';    ! this command saves the
                  !model-implied covariance-matrix; it is redundant with
                  !above command but can be used if more significant
                  !digits are desired
```

M*plus* Syntax for Step 3 (Complete Data Routine)

```
TITLE: math example, 1-factor, read in model-implied cov, treat as complete data.
DATA: FILE IS "math.model.dat";              !see Figure 2
TYPE IS COVARIANCE MEANS;
NOBSERVATIONS = 494;
VARIABLE: NAMES ARE MATH1 MATH2 MATH3 MATH4;
ANALYSIS: ESTIMATOR = ML;
MODEL:
F1 BY MATH1* MATH2 MATH3 MATH4;
F1@1;
```

R Syntax (Using *lavaan* Package) for Steps 1 and 3
(FIML and Complete Data Routine)

```
library(lavaan)
math <- read.table("math.dat", na.strings = "NA")
colnames(math) <- c("MATH1", "MATH2", "MATH3", "MATH4")
#set up model
math.cfa <- 'math =~ MATH1 + MATH2 + MATH3 + MATH4'
step1.cfa <- cfa(math.cfa, data = math, missing = "fiml", std.lv = TRUE)
#save model standard errors
SE.cfa <- parameterEstimates(fit.cfa)$se
#get model-implied covariance matrix and means
cov.cfa <- fitted.values(step1.cfa)\$cov
means.cfa <- fitted.values(step1.cfa)\$mean
#multiply model-implied cov by N/N-1
#note that this is optional and perhaps only worth doing with small N
cov.cfa <- cov.cfa*(494/493)
#run the model using model-implied cov. matrix and means as input
step2.cfa <- cfa(math.cfa, sample.cov = cov.cfa, sample.mean =
                means.cfa, sample.nobs = 494, std.lv = TRUE,
                meanstructure = TRUE, information = "observed")
#get standard errors
SE.step2.cfa <- parameterEstimates(step2.cfa)$se
#compute vector of fraction of missing information estimates
FMI <- 1-(SE.step2.cfa^2/SE.cfa^2)
```