Electronic Journal of Research in
Educational Psychology
E-ISSN: 1696-2095
jfuente@ual.es
Universidad de Almería
España

Wolgast, Anett; Schwinger, Malte; Hahnel, Carolin; Stiensmeier-Pelster, Joachim
Handling Missing Data in Structural Equation Models in R. A Replication Study for Applied Researchers
Electronic Journal of Research in Educational Psychology, vol. 15, núm. 1, enero-abril, 2017, pp. 5-45
Universidad de Almería
Almeria, España

Electronic Journal of Research in Educational Psychology
ejrep
EDUCATION & PSYCHOLOGY I+D+i
Research, Innovation and Solutions on-line
Ilustre Colegio Oficial de Psicología de Andalucía Oriental

Handling Missing Data in Structural Equation Models in R. A Replication Study for Applied Researchers

Anett Wolgast[1], Malte Schwinger[2], Carolin Hahnel[3], Joachim Stiensmeier-Pelster[4]

[1] Department of Educational Psychology, Martin-Luther-University, Halle (Saale)
[2] Department of Educational Psychology, Philipps-University Marburg, Marburg
[3] Centre for International Student Assessment (ZIB)
German Institute for International Educational Research (DIPF), Frankfurt am Main
[4] Department of Educational Psychology, Justus-Liebig-University Giessen, Giessen

Germany

Correspondence: Anett Wolgast. Franckepl. 1, Haus 5, 06110 Halle (Saale), Germany,
e-mail: anett.wolgast@gmail.com

Abstract

Introduction. Multiple imputation (MI) is one of the most highly recommended methods for replacing missing values in research data. The scope of this paper is to demonstrate missing data handling in SEM by analyzing two modified data examples from educational psychology, and to give practical recommendations for applied researchers.

Method. We provide two examples ($N = 589$ and $N = 621$, respectively) based on previous studies of students' self-concepts, mastery goals and performance avoidance goals, and a 7-step tutorial. Then, we produced 20% and 40% missing data under three missing mechanisms by these complete, genuine data sets. The resulting datasets were then analyzed by (1) listwise deletion and structural equation models (SEM), (2) full information maximum likelihood (FIML) with SEM, and (3) MI combined with SEM and pooling. Thus, the results stem from $2 \times 3 \times 3$ conditions.

Results. Previous research was replicated by illustrating a practical way to combine MI with SEM and pooling. The assumed factor structure was depicted in both examples with multiply imputed values applied.

Discussion. We suggest adding variables to clarify the missing data mechanism, especially for dependent variables as motivation. Such variables might indicate whether missing values in dependent variables are correlated with independent variables (e.g., interest) or the dependent variable itself (e.g. lack of motivation independently of interest).

Keywords: missing data, multiple imputation in practice, self-concepts, goals

Reception: 16.11.26 *Initial acceptance: 16.11.30* *Final acceptance: 17.02.07*

Resumen

Introducción. La imputación multiple (IM) es uno de los métodos más recomendados para sustituir valores perdidos en datos de investigación. Este artículo se dedica al manejo de los valores perdidos en MES, analizando dos bases de datos de Psicología Educativa y a recomendaciones para investigadores orientados a las aplicaciones.

Método. Presentamos dos muestras de estudiantes (N = 589 y N = 621, respectivamente) de estudios anteriores que se dedicaron al autoconcepto, a las metas de aprendizaje y de evitación, y al rendimiento en un tutorial de siete pasos. En los datos de las dos muestras producimos artificilamente un 20 y un 40 por ciento de valores perdidos. Luego analizamos estos datos utilizando (1) eliminación de los casos (listwise) y modelos de ecuaciones estructurales (MES), (2) maxima verosimilitud con informacion completa (MVIC) con MES, y (3) IM con MES e agrupamiento de datos (Pooling). Por lo tanto los resultados proceden de un diseño de 2 x 3 x 3 condiciones.

Resultados. Replicamos investigaciones anteriores para ilustrar una manera práctica de combinar IM con MES e Pooling. Imputando valores multiples en las dos muestras, podemos confirmar la estructura supesta de la MES.

Discusión. Recomendamos anadir variables para aclarar el mecanismo de datos perdidos, sobre todo para variables dependentes que se refieren a la motivación. Estos tipos de variables podrían indicar que los valores perdidos en variables dependentes están en correlación con variables independentes (por ejemplo interés) o con la variable dependente en sí (por ejemplo falta de motivación, con independicia de interés).

Palabras claves: valores perdidos, imputación mutiple en páctica, autoconcepto, metas

Recibido: 26.11.16 *Aceptación Inicial: 30.11.16* *Aceptación final: 07.02.16*

Introduction

Results from many simulation studies have indicated that structural equation models (SEM) and additional variables employed in a multiple imputation (MI) model lead to precise results comparable to simulated 'true' values (e.g., Collins, Schafer, & Kam, 2001; Grund, Lüdtke, & Robitzsch, 2015; Merkle, 2011; Si & Reiter, 2013; Sinharay, Stern, & Russell, 2001; Steele, Wang, & Raftery, 2010; van Buuren, Boshuizen, & Knook, 1999). Among statisticians, MI is thus an accepted method of replacing missing values in survey data (Myers, 2011; Rubin, 1996; Schafer & Graham, 2002; Schlomer, Bauman, & Card, 2010). Other researchers, though, remain skeptical of the comparability of results from simulation studies based on full responses and studies based on data including MI, although authors have compared results from simulated data and real world data combined with practical recommendations (e.g., Wiggins & Sacker, 2002).

A further discussion has focused on the advantages and disadvantages of data including multiple imputations and subsequently structural equations, compared to structural equations by FIML (e.g., Collins et al., 2001; Enders, 2010; Graham, 2009). A number of studies have compared different missing data methods using real data, but mostly from the field of medical research (e.g., Kang, Little, & Kaciroti, 2015; Sterne et al., 2009) and rarely from the field of educational psychology. Two prominent educational conceptualizations are academic self-concepts and academic goals. Using SEM to specify predictors of individuals' academic self-concepts (e.g., Craven & Yeung, 2008) and their goals (e.g., Elliot & Murayama, 2008) is well established. Both conceptualizations, however, have rarely been analyzed with regard to different missing data levels and mechanisms.

The scope of this paper is to replicate missing data handling in SEM, to analyze two examples from educational psychology, and to give practical recommendations for applied researchers. We provide a 7-step tutorial on handling MI under the MAR or MNAR assumption on SEM using the R packages *lavaan* and *semTools*. We aimed at replicating findings on the equivalence of FIML and MI, as indicated in the large body of previous research, by extending research on handling missing data to applied research in educational psychology. The two examples are based on previous studies in educational psychology, and we aimed at demonstrating similar results using real responses that there modified to include multiply imputed values (0%, approximately 20%, or 40%). In Example 1, we analyzed teacher education

http://dx.doi.org/10.14204/ejrep.41.16125

students' (TES) academic self-concept, mastery goals and performance avoidance goals. In Example 2, we examined school students' domain-specific self-concepts, mastery and performance avoidance goals (in mathematics and in language arts, respectively). The definition of missing mechanisms and the MI approach are outlined in the following sections.

Missing Mechanisms

Rubin (1976) distinguished between three kinds of missing mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). These mechanisms have been cited, described, and utilized in numerous studies (e.g., Baraldi & Enders, 2010; Graham, 2009; Jolani, van Buuren, & Frank, 2013; Myers, 2011; Rubin, 1996; Schafer & Graham, 2002). MCAR concerns cases where missing values can be described as a random sample. The occurrence of missing values depends neither on the value of the variable itself nor on the value of other variables in the data set (Rubin, 1976, 1996). Nonresponse to an item regarding goals, for instance, depends neither on the amount of goals itself nor on the age of subjects or other characteristics. With MAR, the occurrence of missing values depends neither on the values of the variable itself nor on the expression of other variables in the data set after controlling for additionally observed variables (Rubin, 1976, 1996). When missing values in subjects' age or other subject variables are controlled for, the nonresponse to goals-related items does not depend on the rating of these motivational items themselves. The MNAR mechanism applies when the occurrence of missing values depends on the occurrence of the variable itself even after controlling for responses in additional variables (Rubin, 1976, 1996). Even after controlling for age and other variables, nonresponse to goal-related items depends on the values of the items themselves. Some authors (e.g., Carpenter & Kenward, 2012) prefer the term not missing at random (NMAR) instead of MNAR.

Incorrect assumptions regarding the missing mechanisms can cause varying degrees of bias in research results and misinterpretations of data (Schafer & Graham, 2002). Some authors have proposed that violations of the MCAR assumption can be tested statistically by covariance-based tests (Enders, 2010; Little, 1988), that raise a number of problems, though (Enders, 2010; Kim & Bentler, 2002). MCAR and MAR are not testable themselves. Particularly for dealing with MNAR, the literature shows divergent views, e.g., some argue that MNAR requires a special imputation model to avoid estimation bias (Di Nuovo, 2011; Sinharay et al., 2001; van Buuren, 2012). Other authors argue that MNAR does not require a special imputation model, even for a specific analysis model (e.g., covariates in a regression model can be MNAR, but listwise deletion does not lead to biased estimates; Carpenter & Kenward, 2012). An advantage of MI relative to listwise deletion is that data with imputations have more statistical power than the same data with missing values (Graham, 2009). The mechanism of MAR has been well investigated, mostly in simulation studies (Sinharay et al., 2001; van Buuren, 2012). A disadvantage of simulation studies, in particular regarding structural equations, is that they rarely match the complexity of real data, e.g., nested data structure (Bandelos & Gagné, 2014). Nevertheless, for handling missing data in SEM under the MAR assumption, FIML seems to be used more often than MI, as stated below.

FIML and MI Advantages

Entering the keywords ‘structural equation (all words) in Full Text’ into search engines provides a huge amount of hits (February 13, 2016). Combining these keywords with FIML (‘structural equation AND full information maximum likelihood’), though, usually generated more results than for keywords including MI (‘structural equation AND multiple imputation’) on different platforms (PubPsych: 3 vs. 5 results; Google Scholar: 566,000 vs. 38,700 results). FIML is a popular method for dealing with missing data in SEM. It is often provided in statistical software, and researchers tend to specify a model using the variables of interest, without predictors for missing values (auxiliary variables), that would be useful in the case of MNAR.

FIML methods estimate parameters and standard errors using raw data, instead of a covariance matrix, as well as an algorithm for mostly multiple regressions considering missing values (Enders, 2010; Graham, 2009). If auxiliary variables are included in the MI model in addition to variables of interest, all subsequent analyses benefit from accordingly precise imputed values as an advantage of MI relative to FIML. A practical advantage of FIML relative to MI is that only one command is necessary in addition to the SEM code (e.g., by using the R package *lavaan*). However, possible additional predictors in the researcher’s SEM scope are ignored. Rosseel (2012) has mentioned that *lavaan* would apply case-wise (or ‘full information’) maximum likelihood if the missing mechanism is MCAR or MAR in SEM. Thus, MI procedures are obviously necessary to handle MNAR. The R package *mice* (van Buuren & Groothuis-Oudshoorn, 2011) even provides a function for automatically including auxiliary variables.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Using MI procedures, missing values are replaced on the basis of the distribution of different predictors. The selection of predictors should consider all relevant information in the responses and will depend on their theoretical and statistical relevance in computing the values to impute (van Buuren & Groothuis-Oudshoorn, 2011). The unknown values are regarded as a source of random variation (Collins et al., 2001). For each missing value (*mis*) for a person, replacements are drawn from a predictive distribution. Based on what is known about that person, an unknown dependent variable is identified from this predictive distribution. When sex, age, grade point average, or further independent variables are known, their predictive distribution describes the dependent variable 'goals' assessed by several items. As an intermediate result, *mi* is obtained for each imputed value among all complete responses that can be incorporated into further analyses, for example, structural equation models (SEM, van Buuren, 2012). Results can be combined into a single value. Data including MIs show moderate error rates (Type I and Type II, Collins et al., 2001) and close confidence intervals (Collins et al., 2001, Rubin, 1996). Unlike approaches in which cases with missing values are deleted (known as listwise deletion), statistical power is retained in MI. Preparation and diagnostic analyses after MI are self-evident (e.g., plausibility tests; cf. van Buuren, 2012). In four simulation studies of missing data procedures using 10 and 20 imputed data sets, it was concluded that amounts of 25% missing values often led to substantial problems with "bias, efficiency, and coverage" (Collins et al., 2001, p. 347). However, adequate results from computations with simulated response values and data including multiply imputed values have been found when the proportion of multiply imputed values remained under 25% (Collins et al., 2001). Studies have also shown that 50% missing values for each variable severely limits performance in univariate and multivariate data scenarios. Thus, the proportion of missing values in data used should be below 50%.

In several studies, multiply imputed data have been used based on different assumptions with regard to missing mechanisms and proportions, as the following examples demonstrate: In a psychiatric study[1] published in 2012, the authors reported multiply imputed values of up to 72% in several dependent variables under the assumption of MAR. In another study[1], an average of 16% of data was missing from more than one measurement point before using MI in self-related variables, while the cross-sectional proportion of missing values was not

[1] We omit the citation of these studies. For information on these studies, contact the first author.

given[1]. The MNAR mechanism is relatively often discussed in clinical trial studies (Kang et al., 2015; Sterne et al., 2009), in particular in relation to addiction analyses (e.g., McPherson, Barbosa-Leiker, Burns, Howell, & Roll, 2012).

In summary, the advantage of MI relative to FIML and listwise deletion is that little information is lost, since all variables are included in the model generating the imputed values, and are tested to determine whether they affect subjects' responses. Furthermore, this method includes standard errors and a multiple (*m*-times) iterative repetition. Information on MI stems predominantly from simulation studies combined with field research. There is little practical evidence based on complete, genuine responses systematically replaced with multiply imputed values.

Practical Contexts: Academic Self-Concepts and Goals

There is a need for example studies in practical contexts, e.g., in the field of educational psychology (van Buuren & Groothuis-Oudshoorn, 2011). As a result, we focused on academic self-concepts and goals, two conceptualizations that have frequently been examined in psychological research on education. Academic self-concepts are defined as a set of cognitive representations of an individual's own abilities in academic achievement situations in terms of talent, intelligence, learning ability, and mastery of tasks and requirements (Schöne, Dickhäuser, Spinath, & Stiensmeier-Pelster, 2012). Previous research has replicated results which indicate domain-specific self-concepts, e.g., mathematical self-concept or language-related self-concept (Marsh et al., 2015). Conceptual dimensions of goals include mastery goals and performance avoidance goals. Students endorsing mastery goals aim to improve their competencies (Elliot & Murayama, 2008: for a meta-analysis see Chiungjung, 2012). Performance avoidance goals reflect the intention to prevent failure and avoid normative incompetence (Elliot & Murayama, 2008; for a review, see Senko, Hulleman, & Harackiewicz, 2011).

Empirical results support the reciprocal relationship among academic self-concepts, mastery goals or performance avoidance goals, and academic performance, with mastery goals and high academic self-concepts having been shown to be most adaptive in this regard (e.g., Marsh et al., 2015; Mone, Baker, & Jeffries, 1995; Valentine, DuBois, & Cooper, 2004; Wigfield & Karpathian, 1991). This pattern of findings is evident among teacher education students as well as school students (Craven & Yeung, 2008). Furthermore, empirical results

http://dx.doi.org/10.14204/ejrep.41.16125

have indicated that sex and age can explain variance in academic self-concepts (Corker, Donnellan, & Bowles, 2013; Hodis, Meyer, McClure, Weir, & Walkey, 2011; Marsh et al., 2005; Pfeifer et al., 2013) or academic goals (Corker, Donnellan, & Bowles, 2013). Studies of students' academic self-concepts and goals typically are affected by non-trivial amounts of missing data (e.g., on average 11% of reading self-concept variables, Retelsdorf, Schwartz, & Asbrock, 2014; around 13% of learning goals variables, Fischer & Theis, 2014).

Overview of the Present Research

We present two example studies to illustrate SEM outcomes on several independent variables under three missing assumptions and three ways of handling missing data. First, structural equation models were specified using full data from two example studies: (1) academic self-concept, mastery goals, and performance avoidance goals as dependent variables (DV) were regressed on sex, age, and grade point average at school as independent variables (IV) for university students; (2) math and language related self-concepts, mastery and performance avoidance goals (DV) were regressed on sex, age, type of school and previous grade point average (IV) secondary school students. In line with related findings (Corker et al., 2013; Hodis et al., 2011; Pfeifer et al., 2013), we expected sex, age, and grade point average to predict students' academic self-concepts, mastery goals and performance avoidance goals. Second, the full data sets were manipulated by replacing data with missing values. We generated six data sets with missing values for each of both example full data sets: two MCAR data sets with 20% and 40% missing values in dependent variables, two MAR data sets with 20% and 40% missing values in dependent variables, and two MNAR data sets with 20% and 40% missing values in dependent variables. We decided to include levels of 20% and 40% missing values because previous research indicated adequate results for proportions below 25% and biases at levels of 50% or higher (Collins et al., 2001; Enders, 2010). Third, we analyzed the generated data sets using (1) listwise deletion and SEM, (2) FIML, and (3) MI, SEM, and pooling. We report variance explained in dependent variables and fit indices because we assume that every applied researcher is theoretically and empirically familiar with the practical relevance of explained variance in a variable, while fit indices are especially relevant for the method itself.

Systematic Missing Data in the Present Study

When an observation fulfilled the conditions of the implemented missing mechanisms, all values for the dependent variables of academic self-concepts, mastery goals and performance avoidance goals were set to nonresponse, that is, full cases were set as missing. For MCAR, cases were randomly selected. To determine which case data had to be replaced with nonresponses, a random permutation was created containing elements of the dichotomous set {0;1} for each case. The probability of drawing 0 was .20 in one condition and .40 in the other condition. Individuals' values for academic self-concepts and goals were set to be missing when 0 was drawn. Table A1 (see Appendix A) contains the missing value amounts for the full response data in each of the data sets generated for Example 1 and Example 2.

For MAR, missing values were generated on the basis of values for an auxiliary variable regarding participants' interest. Example 1 included the auxiliary variable 'I am interested in the intermediate results of the study'. Individuals responded to the auxiliary variable with a confirmation, non-confirmation, or nonresponse. Example 2 involved the auxiliary variables 'sum of interest in mathematics scores' and 'sum of interest in language scores' from four items each, e.g., 'I am interested in mathematics' (4-point scale, 1 = *strongly* and 4 = *strongly agree*).

In Example 1, the missing values for the academic self-concept and goals variables were set to depend on individuals' responses to the variable 'I am interested in the intermediate results of the study,' since omitted or non-confirmed responses in this auxiliary variable would indicate non-participation in a re-test (Enders, 2010). In the real response data set of Example 1, $n = 119$ (20% of $N = 589$) individuals omitted the response and $n = 236$ (40% of $N = 589$) individuals did not confirm an interest in the results of the study. For the 20% missing data rate condition, cases were replaced with $mis = 119$ missing data points when individuals declined to respond to the auxiliary variable. For the 40% missing data rate condition, cases were replaced with $mis = 236$ missing data points when individuals did not confirm interest. In Example 2, nonresponses were set according to the lower 20% and 40% of sum scores regarding interest in mathematics or language; the proportions are depicted in Table A1 in Appendix A.

For MNAR, Example 1 and Example 2 cases were replaced with nonresponses depending on sum scores on the dependent variables academic self-concept and goals (Rubin, 1976, 1996). Nonresponses were set according to the lower 20% and 40% of the sum scores. Before imputation, we analyzed and prepared the data sets for Example 1 and Example 2 as recommended by van Buuren (2012).

Method and Results

Example Studies and Procedure

Example 1. A total of $N = 589$ teacher education students (TES) at one university completed an online questionnaire via an internal learning platform (female: $n = 339$, sex coded as 1 = *female* and 2 = *male*). The participants' mean age was 22 years ($M = 21.61$, $SD = 3.15$), their high school grade point average (*Abiturnote*) was $M = 2.58$ ($SD = 0.79$, 1 = *highest grade* to 4 = *lowest grade*), and they had finished their first year of study. The TES had chosen one of five different degree programs: $n = 25$ for teaching at elementary schools, $n = 146$ for teaching at secondary schools, $n = 262$ for teaching at high schools involving academically challenging education, $n = 25$ for teaching special education, and $n = 21$ for teaching vocational skills.

In their classes, the students were invited to complete the questionnaire at home and given two e-mail reminders (after one week, and after a further two weeks). The questionnaire was accessible to students for another three weeks and the survey period lasted six weeks. All scales were presented on the computer screen. The TES chose the order in which they responded to the items. Participation in the survey was voluntary and anonymous. The TES received information in class that future teacher education course curricula would refer to the survey content. A raffle of 10 vouchers worth 30 euros each was offered as an incentive to participate. Double participation was prevented by controlling access with personal codes. Responding to the following additional item 'I am interested in the intermediate results of the study' was voluntary (coded as 1 = *yes* and 2 = *no*).

Example 2. This sample consisted of 621 students (321 female, sex coded as 0 = *male* and 1 = *female*) in their sixth ($n = 24$), seventh ($n = 132$), eighth ($n = 206$), ninth ($n = 193$), or tenth ($n = 66$) grade at academic-oriented (553 students) or academic and occupationallly oriented secondary schools in Germany. The mean age of the students was about 14 years (range 11–18 years). Only students who provided parental consent forms on the day of testing were allowed to participate. Students completed a questionnaire assessing their self-concepts, mastery goals or performance avoidance goals with respect to mathematics and language as well as the demographic variables sex and age. Class teachers administered the questionnaire during regular lessons. The hierarchical data structure of individual students within *classes at schools* was taken into account in SEM.

Instruments

DVs. Academic self-concepts were assessed with five items from a standardized instrument (SESSKO, Schöne et al., 2012). Each item was rated on a 5-point scale ranging from 1 = *strongly disagree,* 2 = *disagree,* 3 = *neither agree nor disagree,* 4 = *agree* to 5 = *strongly agree*. In Example 1, the items captured the TES' academic self-concept in general (α = .70; e.g., 'at university, I know little/a lot'). In Example 2, the academic self-concept scale (SESSKO; Schöne et al., 2012) was adapted to measure students' ability self-concept in mathematics and language (α = .78; e.g., 'in German, I know little/a lot').

Mastery goals and performance avoidance goals were each assessed by four items adapted from Spinath, Stiensmeier-Pelster, Schöne, and Dickhäuser (2012). The items were ranked on scales ranging from 1 = *strongly disagree* to 5 = *strongly agree*. In Example 1, the TES' mastery goals (α = .83) were assessed with, e.g., 'I strive to learn as much as possible at university' and performance avoidance goals (α = .86) with, e.g., 'At university, I strive not to make a fool of myself by asking stupid questions'. In Example 2, school students' mastery goals (α = .75) were assessed using, e.g., 'In mathematics, I strive to learn as much as possible' and performance avoidance goals (α = .81) by, e.g., 'In German, I strive not to make a fool of myself by asking stupid questions'. Students responded to all items as described above.

A confirmatory factor analysis with Example 1 data indicated three factors: academic self-concept, mastery goals, and performance avoidance goals (maximum likelihood (ML) estimation; $\chi^2(62) = 190.230$; root mean structure error of approximation (RMSEA) = .059, CI [.050, .069]; comparative fit index (CFI) = .954; standardized root mean square residual (SRMR) = .042; Yuan, 2005). In Example 2, the three factors academic self-concepts, mastery goals, and performance avoidance goals represented the data acceptably (simultaneously computed ML estimation for mathematics and language; $\chi^2(62) = 235.510$, RMSEA = .067, CI [.058, .076], CFI = .957, SRMR = .048).

IVs. Participants in both examples reported their sex, age, and previous grade point average at school. The grade point average for TES ranged from 1 (*highest grade*) to 4 (*lowest grade*) in Example 1. For Example 2, students' grades ranged from 1 (*very good*) to 6 (*insufficient*), with lower scores indicating better performance. The type of school (1 = *aca-*

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

demic oriented school, 'Gymnasium'; 2 = *academic and occupationally oriented school, 'Gesamtschule'*) was considered as an additional variable in Example 2.

Analyses in the Present Study

The analyses were conducted with R 3.1.1 (R Core Team, 2015), as well as the R packages *psych* (Revelle, 2015) and *lavaan* (Rosseel, 2012) for the most part. The *runMI* function from the R package *semTools* (semTools Contributors, 2014; Li, Meng, Raghunathan, & Rubin, 1991) combines SEM with pooling that creates single point estimates of the *m* values considering Rubin's (1987) rules (for more information, see van Buuren, & Groothuis-Oudshoorn, 2011, and the R script (see Appendix B3). Imputed data were analyzed according to plausibility (van Buuren, 2012).

Specifically, we used the *lavaan* and *semTools* packages in R to run MI, and we report the results that are part of the output of these packages. The *lavaan* package uses the likelihood function which is derived from a multivariate normal distribution or from the multivariate equivalent of the chi-squared distribution (also known as Wishart distribution). *Lavaan* makes listwise deletions of cases containing missing values if missing values are defined by the researcher. If a *lavaan* script includes the FIML command, an unrestricted model assuming differences between the specified structure and the provided data will automatically be estimated using the Estimation Maximization (EM) algorithm. The EM algorithm assigns expected values from model specifications to data while also adapting the model specifications to the data. From the EM algorithm, *lavaan* derives absolute and incremental fit indices (Rosseel, 2012; for details on the equations see Li et al., 1991; Rubin, 1987). Absolute fit indices represent the equivalence between the specified model and data; the root means squared residual (SRMR) and the root mean square error of approximation (RMSEA) (Hu & Bentler, 1999) were mostly reported. Incremental fit indices represent results comparing baseline unrestricted models with models that are restricted at different levels, e.g., the comparative fit index (CFI) or the Tucker-Lewis index (TLI) (Hu & Bentler, 1999). FIML and MI perform similarly when the same variables are specified in SEM and the number of imputations is appropriate for the proportion of missing values (Graham, 2009), e.g., the number of imputations should be $m = 20$ when the proportion of missing values is about 20%. Accordingly, we used $m = 20$ imputations for the case with 20% missing values and $m = 40$ imputations for the case with 40% missing values.

The imputation model for MI was specified by full conditional specification, also known as chained equations (van Buuren & Groothuis-Oudshoorn, 2011). The missing predictor variables were known in the MAR condition (non-confirmation of *interest in study results*) and the MNAR condition (low levels of academic self-concepts, mastery goals or performance avoidance goals). The missing predictor variable, the dependent variables academic self-concepts, mastery goals and performance avoidance goals, and the independent variables (Example 1: sex, age, grade point average; Example 2: sex, age, type of school, grade point average) were included in the imputation model. The package *semTools* uses the package *Amelia* to impute the missing data. In the default setting, the SEM is fitted, and the resulting estimates and χ^2 values are aggregated according to the procedure by Meng and Rubin (1992; also called "D3" in Enders, 2010). This procedure is designed to aggregate a series of likelihood-ratio tests obtained from multiply imputed data sets (e.g., the comparison with the saturated model). In the default settings in *semTools*, which attempt to imitate the behavior of the software Mplus, the resulting test statistics proposed by Enders (2010, D3-statistic) are transformed into a single χ^2 -value by means of a large-sample approximation as described in Asparouhov and Muthen (2010). From this aggregated χ^2 -value, the RMSEA is calculated. This background is concurrently recommended as "state-of-the-art" for application in SEM (Enders, 2010; van Buuren, 2012).

SemTools uses an original *lavaan* object which involves, for example, regression coefficients and chi-squares. *SemTools* combines the *lavaan* object with multiple results and pools adjusted fit indices from multiple datasets (according to Rubin, 1987) into a *lavaanStar* object including the original *lavaan* object and adjustment values to the null model gleaned by taking auxiliary variables into account (semTools Contributors, 2014). For example, the SRMR across multiple imputation data sets results from the model related average means and the multiple covariance matrices (semTools Contributors, 2014; Li et al., 1991). Multiply imputed values in data sets are not yet standard practice, although applied researchers would benefit from using them. We give a brief description of the analysis to help illustrate *lavaan's* and *semTools'* behavior and to demonstrate how the fit indices and χ^2 -values presented in this paper were derived.

First, SEMs were conducted using the full data sets (0% missing data) in both example studies. The SEMs specified academic self-concepts, mastery goals, and performance avoidance goals as DV in Example 1 and Example 2. In Example 1, TES' sex, age and grade point

http://dx.doi.org/10.14204/ejrep.41.16125

average were included in SEM (see Figure B1 in Appendix B; Epskamp, 2014). In Example 2, school students' sex, age, type of school and grade point average were included in SEM (see Figure B2 in Appendix B).

The results indicated that the assumed structure of the model was reflected in the structure of the data (see Table A2.1, Table A3.1 and Table A4.1 in Appendix A). Small (8%, Example 1, see Table A2.2) to medium (45%, Example 2, see Table A3.2 and Table 4.2) proportions of variance in academic self-concept were explained. Small (up to 15%, both Examples, see Table A2.2, Table A3.2 and Table A4.2) proportions of variance in mastery goals and performance avoidance goals were explained. The TES' academic self-concept was significantly determined by their grade point average at school (1 = *highest grade,* 4 = *lowest grade;* see Table A2.2*)*. Mastery goals were significantly determined by sex in favor of female TES, but TES' performance avoidance goals were not determined by sex, age, or grade point average (see Table A2.2).

Example 2 data involved the domain-specific DV self-concept, mastery goals and performance avoidance goals related to mathematics (SEM_{ma}) and to language (SEM_{la}). Students' mathematical self-concept and mastery goals were significantly determined in favor of boys and by a higher previous grade point average (see Table A3.2 in Appendix A*)*. Math and language-related mastery goals and performance avoidance goals were significantly determined by the type of school (see Table A3.2 and A4.2) and previous grade point average. Students with higher previous grade point averages showed higher levels of mastery goals and lower levels of performance avoidance goals. Students' language-related self-concept and mastery goals were significantly determined in favor of girls and by a higher previous grade point average. Language-related performance avoidance goals were significantly determined by the type of school (see Table A4.2). Students showed higher levels of performance avoidance goals when they attended the academic and occupationally-oriented type of school ('Gesamtschule') than students who attended the academically-oriented type of school ('Gymnasium').

Equivalent SEM structures were specified in analyzing data sets with missing values manipulated. Listwise deletion and FIML led to the same beta-coefficients if, after listwise deletion, the distribution included in the ML calculation was congruent with the distribution included in FIML; ML and FIML based on equivalent information (Enders, 2001; Myers,

http://dx.doi.org/10.14204/ejrep.41.16125

2011; Schafer & Graham, 2002). When listwise deletion and SEM were applied, the results from the conditions with 20% and 40% missing values were similar: A low proportion of variance in academic self-concept and goals was explained in Example 1 (see Tables A2.2–A2.4) while a medium proportion of variance was explained in Example 2 (see Tables A3.2–A3.4 and A4.2–A4.4). However, the SEM structure regarding mathematical self-concept and goals was not identified under MNAR with 40% missing data using listwise deletion and SEM in Example 2.

When FIML was applied, again, a low proportion of variance in academic self-concept and goals was explained in Example 1 (see Tables A2.2–A2.4) and a medium proportion of variance in Example 2 (see Tables A3.2–A3.4 and A4.2–A4.4). When MI and SEM were applied, the proportions of variance in academic self-concepts and goals explained in both examples were similar to results from FIML (see Tables A2.2–A2.4, A3.2–A3.4 and A4.2–A4.4).

Discussion

The first aim of this research was to reduce skepticism regarding the effectiveness of MI with brief information of research on dealing with missing data. Previous research compared different missing data methods in real data, but mostly in a medical research context (Kang et al., 2015; Sterne et al., 2009; van Buuren et al., 1999). Second, we demonstrated how to handle different missing response mechanisms by applying listwise deletion, FIML, and MI with SEM in two example studies in an educational context. Genuine complete responses were analyzed under MCAR, MAR, and MNAR combined with conditions in which 20% or 40% of values were missing. Listwise deletion and SEM; SEM using FIML; and MI, SEM, and pooling were applied (Robitzsch, 2015; Rubin, 1976; Rubin, 1996; semTools Contributors, 2014; van Buuren, 2012). Both examples involved dependent variables concerning academic self-concepts (Marsh, 1989; Marsh et al., 2015; Schöne et al., 2012) and academic goals (Chiungjung, 2012; Spinath et al., 2012). Listwise deletion and SEM led to fit indices which indicated model divergence, whereas SEM using FIML and SEM including MI resulted in acceptable fit indices close to cut-off criteria as defined by Hu and Bentler (1999).

A relatively new result concerns the stability of the SEM structure in both examples under missing data conditions with multiply imputed values applied. Several authors have suggested avoiding standard MI methods under the MNAR assumption because results may be biased (Horton & Lipsitz, 2001; Sinaray et al., 2001). Other authors have stated that data

http://dx.doi.org/10.14204/ejrep.41.16125

both within and outside of special imputation models (e.g., covariates in a regression model; see Carpenter & Kenward, 2012) can be MNAR without leading to biased estimates. This statement was true for our specific analysis model.

Listwise deletion resulted in similar regression coefficients and proportions of variance explained while also leading to similar conclusions as those generated using genuine complete responses, except in the case of 40% missing values under MNAR, where standard errors could not be computed. These results are in line with findings from simulation studies (e.g., Jolani et al., 2013). SEM using FIML led to similar conclusions as results from genuine complete responses, even though little additional information was included in MI. As expected, the MI approach resulted in similar effects that might increase acceptance of the effectiveness of MI and its frequency of use among applied researchers.

From our view, the interesting aspect of the *runMI* code in *semTools* concerns pooling test statistics and fit statistics in SEM. It is evident in the presented results that fit statistics (RMSEA) tended to indicate greater equivalence when missing value rates were higher. The lower RMSEA values can be explained due to Enders' test statistic "D3" (2010) being transformed into the aggregated χ^2 -value (Asparouhov & Muthen, 2010) upon which the RMSEA was based. This finding is in line with Davey (2005). However, the fit statistics SRMR, CFI, and TLI indicated differences between the specified model and the data for both academic self-concepts and goals variables when the missing rate was about 40%.

In previous research, it was argued that the 'visibility' of imputed data is not actually that important. Whereas some authors prefer that imputations 'look like' observed data, mostly by indicating values to two decimal places (Engels & Diehr, 2003; van Buuren, 2012), other authors have argued that the 'visibility' of all decimals computed is necessary to successfully recover parameter estimates (e.g., the discussion on rounding off imputations based on normal model MI; Horton, Lipsitz, & Parzen, 2003; Schafer, 1997). From a practical perspective, we argue for 'visibility' of imputations in all their decimal places. Imputed values should 'look different' than real data so that researchers can recognize cells with imputed values due to the many decimals recorded and displayed, in contrast to genuine values which are usually recorded and displayed with just two decimals. Keeping imputed values in data sets obvious helps to ensure that researchers remain aware of the different nature of these values, especially in the case of conducting secondary analyses on large-scale data.

Furthermore, we aimed to replicate previous research by illustrating a practical way to combine MI and SEM, while also giving a 7-step-tutorial (see Appendix C). We discussed methods that are not yet standard practice in applied research, e.g., the calculation of fit indices for SEM with multiply imputed data sets. As a practical implication, we recommend the R package *semTools,* mentioned above, and give a 7-step tutorial in Appendix C. Furthermore, van Buuren and Groothuis-Oudshoorn (2011) provided an example of analyzing data under the MNAR assumption. Grund, Lüdtke, and Robitzsch (2015) evaluated MI under specific MNAR conditions. They concluded that MI can work - better than listwise deletion - with MNAR data for missing covariate values in multilevel models with random slopes. Grund et al. (2015) provided a code which can be used to analyze data under the MNAR assumption.

Limitations and Implications

The limitation of this study is that both investigations involved examples with a high number of possible patterns of missing values. The conclusions drawn from these data sets might not be able to be generalized to other missing data mechanisms or other data sets. A simulation study would be appropriate to draw general conclusions from the examples. We presented example studies here because mistrust in stochastic paradigms and their validity for specific real-world data have led to doubts regarding the result of simulation studies. We conducted deterministic analyses to demonstrate how applied researchers can handle missing data.

An important divergence in statistical views regarding MNAR should be bore in mind. We generated MNAR and considered the missing mechanism to be MNAR because we knew which mechanism we generated. If we did not know the nature of our MNAR data sets and were confronted with the data the first time, our statistical diagnosis would be MAR (Collins et al., 2001), due to the inclusion of students' grade point average as a determinant in Example 1 and Example 2. We suggest adding variables to clarify the missing data mechanism, especially for dependent variables relevant to motivation, e.g., (1) a lack of interest and related missing values for dependent variables (MAR), or (2) missing values for dependent variables, such as self-concepts or goals, because of low levels of these values (MNAR). If a student saw no reason to respond to items on a questionnaire related to self-concepts and goals due to his or her low levels of these constructs, this would be MNAR, although given that low school performance predicted the proportion of missing values in self-concepts and academic goals, researchers might assume MAR.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

In summary, research reports should provide all information on the diagnosis of missing values and the MI procedure used, which is necessary for replication. MI is an appropriate method for analyzing self-related variables and missing data therein under the assumptions of MCAR, MAR, and MNAR (Graham, 2009; Rubin, 1996; Schlomer et al., 2010; van Buuren, 2012). The R package *semTools* provides the very helpful and practical function *runMi*, which allows MI and pooling with the *lavaan* model and provides all the advantages of MI relative to listwise deletion or FIML, particularly for MNAR data.

Acknowledgement

The data used in this study were collected for part of a study funded by the *German Federal Ministry of Education and Research*. The authors are also grateful to Nadja Teistler for translation the abstract into Spanish.

References

Asparouhov, T. & Muthen, B. (2010). *Chi-Square Statistics with Multiple Imputation. Technical Report*. Retrieved from http://www.statmodel.com

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48,* 5–37.

Bandelos, D. L., & Gangné, P. (2014). Simulation methods in structural equations. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (p. 93). New York: The Guilford Press.

Carpenter, J., & Kenward, M. (2012). Multiple imputation and its application. New York: John Wiley & Sons.

Chiungjung, H. (2012). Discriminant and criterion-related validity of achievement goals in predicting academic achievement: A meta-analysis. *Journal of Educational Psychology, 104,* 48–73. doi: 10.1037/a0026223

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive missing-data strategies in modern missing-data procedures. *Psychological Methods*, *6*, 330–351.

Corker, K. S., Donnellan, M. B., & Bowles, R. P. (2013). The development of achievement goals throughout college: Modeling stability and change. *Personality and Social Psychology Bulletin*, *39*, 1 401–1 417.

Craven, R. G., & Yeung, A. S. (2008). Why self-concept matters for teacher education: Examples from performance, mathematics and reading, and aboriginal studies research.

Paper presented at the *Australian Association for Research in Education conference, Brisbane.* Retrieved from http://www.aare.edu.au/08pap/cra08833.pdf

Davey, A. (2005). Issues in evaluating model fit with missing data. *Structural Equation Modeling: A Multidisciplinary Journal, 12,* 578–597.

Di Nuovo, A. G. (2011). Missing data analysis with fuzzy C-Means: A study of its application in a psychological scenario. *Expert Systems with Applications, 38*, 6 793–6 797.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: critique, illustration, and application. *Journal of Educational Psychology, 100*, 613–628. doi: 10.1037/0022-0663.100.3.613

Engels, J. M. & Diehr, P. (2003). Imputation of missing longitudinal data: a comparison of methods. *Journal for Clinical Epidemiology, 56*, 968–976.

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

Epskamp, S. (2014). *semPlot: Path diagrams and visual analysis of various SEM packages' output. R package version 1.0.1.* Retrieved from http://CRAN.R-project.org/package=semPlot

Fischer, N., & Theis, D. (2014). Extracurricular Participation and the Development of School Attachment and Learning Goal Orientation: The Impact of School Quality. *Developmental Psychology*. doi: 10.1037/a0036705

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

Grund, S., Lüdtke, O., Robitzsch, A. (2015). Multiple imputation of missing covariate values in multilevel models with random slopes: A cautionary note. *Behavior Research Methods*. Retrieved from http://link.springer.com/article/10.3758/s13428-015-0590-3#

Hodis, F. A., Meyer, L. H., McClure, J., Weir, K. F., & Walkey, F. (2011). A longitudinal investigation of motivation and secondary school achievement using growth mixture modeling. *Journal of Educational Psychology, 103*, 312–323.

Horton, N. J., & Lipsitz, S. R. (2001). Multiple imputation in practice. Comparison of software packages for regression models with missing variables. *The American Statistician, 55*, 244–254.

Horton, N., Lipsitz, S. R., & Parzen, M. (2003). A potential for bias when rounding in multiple imputation. *The American Statistician, 57*, 229–233.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1–55. doi: 10.1080/10705519909540118

Jolani, S., van Buuren, S., & Frank, L. E. (2013). Combining the complete-data and nonresponse models for drawing imputations under MAR. *Journal of Statistical Computation and Simulation, 83*, 868–879.

Kang, S., Little, R. J., & Kaciroti, N. (2015). Missing not at random models for masked clinical trials with dropouts. *Clinical Trials, 12*, 139–148.

Kim, K. H., & Bentler, P. M. (2002). Tests of homogeneity of means and covariance matrices for multivariate incomplete data. *Psychometrika*, *67*, 609–624.

Li, K.H., Meng, X.-L., Raghunathan, T.E., & Rubin, D.B. (1991). Significance levels from repeated p-values with multiply-imputed data. *Statistica Sinica, 1*, 65–92.

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*, 1 198–1 202.

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*, 417–430.

Marsh, H. M., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. S., Abdelfattah, F., & Jansen, M. (2015). Dimensional Comparison Theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction*, *35*, 16–32.

McPherson, S., Barbosa-Leiker, C., Burns, G. L., Howell, D., & Roll, J. (2012). Missing data in substance abuse treatment research. Current methods and modern approaches. *Experimental and Clinical Psychopharmacology, 20*, 243–250.

Meng, X. L., & Rubin, D. B. (1992). Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika*, *79*, 103–111.

Merkle, E. C. (2011). A comparison of imputation methods for bayesian factor analysis models. *Journal of Educational and Behavioral Statistics*, *36*, 257–276.

Mone, M. A., Baker, D. B., & Jeffries, F. (1995). Predictive validity and time dependency of self-efficacy, self-esteem, personal goals, and academic performance. *Educational and Psychological Measurement*, *55*, 716–727.

Myers, T. A. (2011). Goodbye, listwise deletion. *Communication Methods and Measures*, *5*, 297–310.

Pfeifer, J. H., Kahn, L. E., Merchant, J. S., Peake, S. A., Veroude, K., Masten, C. L., Lieberman, M. D., Mazziotta, J. C., & Dapretto, M. (2013). Longitudinal change in the neural bases of adolescent social self-evaluations: Effects of age and pubertal development. *Journal of Neuroscience*, *24*, 415–7 419.

R Core Team (2015). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. Retrieved from: http://www.R-project.org/

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2014). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*. doi: 10.1037/a0037107

Revelle, W. (2015). *psych: Procedures for personality and psychological research. R package version 1.5.1*. Retrieved from http://personality-project.org/r/psych-manual.pdf

Robitzsch, A. (2015). *miceadds: Some Additional Multiple Imputation Functions, Especially for 'mice'. R package version 1.1-1*. Retrieved from http://CRAN.R-project.org/package=miceadds

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

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

Rubin, D. B. (1987) *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley and Sons. (2004) Classic edition.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association, 91*, 473–489.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.

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

Schlomer, G. L., Bauman, S., & Card, N. A. (2010). Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology, 57*, 1–10.

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2012). *Skalen zur Erfassung des schulischen Selbstkonzepts (SESSKO) [Academic self-concept scales]*. Göttingen: Hogrefe. Retrieved from http://www.testzentrale.de/programm/skalen-zur-erfassung-des-schulischen-selbstkonzepts.html

semTools Contributors (2014). *semTools: Useful tools for structural equation modeling.* R package version 0.4-6. Retrived from http://cran.r-project.org/web/packages/semTools/index.html

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist, 46*, 26–47.

Si, Y., & Reiter, J. P. (2013). Nonparametric bayesian multiple imputation for incomplete categorical variables in large-scale assessment surveys. *Journal of Educational and Behavioral Statistics, 38*, 499–521.

Sinharay, S., Stern, H. S., & Russel, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6*, 317–329.

Spinath, B., Stiensmeier-Pelster, J., Schöne, C., & Dickhäuser, O. (2012). *Skalen zur Erfassung der Lern- und Leistungsmotivation (SELLMO). [Achievement goals scales]* Göttingen: Hogrefe. Retrieved from http://www.testzentrale.de/programm/skalen-zur-erfassung-der-lern-und-leistungsmotivation.html

Steele, R. J., Wang, N., & Raftery, A. E. (2010). Inference from multiple imputation for missing data using mixtures of normals. *Statistical Methodology, 7*, 351–365.

Sterne, J., White, I. R., Carlin, J. B., Spratt, M. P., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal, 338*, 157–160.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relations between self-beliefs and academic achievement: A systematic review. *Educational Psychologist, 39*, 111–133.

van Buuren, S. (2012). *Flexible imputation of missing data*. Boca Raton, FL: Chapman & Hall/CRC Press.

van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in Medicine, 18*, 681–694.

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*, 1–67.

Wiggins, R.D. & Sacker, A. (2002). Strategies for handling missing data in SEM: A user's perspective. In G.A. Marcoulides & I. Moustaki (eds.), Latent variable and latent structure models (pp. 105–120). Mahwah: Erlbaum.

Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children’s self-concepts and motivation in achievement solutions. Educational Psychologist, 26, 233–261.

Yuan, K.-H. (2005). Fit indices versus test statistics. *Multivariate Behavioral Research, 40*, 115–148.

Appendix A

Table A1. *Generated Missing Rates (in %) for the Dependent Variables Academic Self-concepts (SC), Mastery Goals (MG), and Performance Avoidance Goals (PG)*

Condition	*Missing Mechanism*	*Example 1 (N = 589)*			*Example 2 (N = 621)*		
		SC	MG	PG	SC	MG	PG
20%	MCAR	18	20	20	18	18	18
40%		39	39	41	41	41	41
20%	MAR	20	20	20	18	18	18
40%		40	40	40	37	37	37
20%	MNAR	27	17	18	17	20	20
40%		39	36	38	39	42	40

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A2.1. *Fit Statistics for all Example 1 Models*

Missing Mechanism	**C**	**Method**	**RMSEA**	**CI_{RMSEA}**	**SRMR**	**CFI**	**TLI**	**χ^2 (*df* = 92)**
	Full		**.050**	**[.042, .058]**	**.040**	**.953**	**.940**	**226.411**
MCAR	20%	LD	.054	[.044, .064]	.045	.944	.929	197.093
	20%	FIML	.047	[.039, .056]	.041	.948	.934	212.910
	20%	MI	.045	[.037, .054]	.049	.947	.932	203.724
	40%	LD	.059	[.043, .074]	.055	.932	.913	159.820
	40%	FIML	.046	[.037, .054]	.049	.932	.914	205.241
	40%	MI	.037	[.028, .046]	.068	.938	.921	167.363
MAR	20%	LD	.049	[.040, .059]	.044	.951	.938	196.788
	20%	FIML	.044	[.035, .052]	.042	.951	.938	196.788
	20%	MI	.039	[.030, .047]	.057	.939	.923	172.973
	40%	LD	.044	[.032, .056]	.041	.964	.954	155.537
	40%	FIML	.034	[.025, .043]	.041	.964	.954	155.537
	40%	MI	.036	[.027, .045]	.053	.958	.946	163.433
MNAR	20%	LD	.042	[.028, .055]	.052	.930	.911	142.990
	20%	FIML	.036	[.027, .045]	.042	.939	.923	161.728
	20%	MI	.039	[.030, .047]	.057	.939	.923	172.973
	40%	LD	.053	[.030, .073]	.075	.837	.793	130.559
	40%	FIML	.036	[.026, .045]	.057	.874	.840	160.941
	40%	MI	.030	[.020, .040]	.066	.917	.895	141.307

Note. C = condition; RMSEA = root mean square errors of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index; TLI = Tucker-Lewis index. LD = listwise deletion.

Table A2.2. *Example 1: Results from Full Data and Data Manipulated under the Assumption of MCAR*

C	Method	IV	SC			MG			PG		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
Full		**sex**	**.071**	**.042**	**.087**	**-.243**	**.062**	**.000**	**.020**	**.090**	**.822**
		age	**-.010**	**.006**	**.099**	**-.001**	**.009**	**.900**	**-.014**	**.013**	**.270**
		gpa	**-.219**	**.037**	**.000**	**-.052**	**.053**	**.327**	**.017**	**.079**	**.831**
		R^2	**.077**			**.039**			**.002**		
MCAR 20%	LD	sex	.089	.051	.083	-.016	.078	.037	.055	.110	.616
		age	-.005	.008	.558	.002	.012	.899	-.012	.017	.477
		gpa	-.241	.048	.000	-.132	.070	.058	.058	.099	.557
		R^2	.084			.027			.003		
20%	FIML	sex	.085	.046	.068	-.185	.067	.006	.035	.101	.733
		age	-.007	.007	.372	.001	.010	.885	-.010	.014	.505
		gpa	-.228	.043	.000	-.087	.059	.141	.009	.089	.920
		R^2	.078			.026			.001		
20%	MI	sex	.084	.047	.076	.178	.067	.008	.019	.097	.846
		age	-.004	.007	.591	.000	.010	.980	-.010	.014	.482
		gpa	-.230	.043	.000	-.094	.059	.113	.011	.088	.896
		R^2	.075			.026			.001		
40%	LD	sex	-.020	.070	.770	-.264	.097	.006	-.087	.138	.528
		age	-.012	.011	.306	-.030	.016	.052	-.002	.022	.932
		gpa	-.099	.066	.132	.026	.089	.768	-.031	.129	.814
		R^2	.022			.069			.002		
40%	FIML	sex	.022	.052	.676	-.229	.076	.003	-.081	.116	.485
		age	-.021	.008	.007	-.009	.012	.456	-.005	.018	.800
		gpa	-.155	.048	.001	-.025	.067	.704	.081	.101	.427
		R^2	.060			.039			.004		
40%	MI	sex	.015	.050	.755	-.228	.084	.007	-.090	.110	.412
		age	-.019	.008	.015	-.009	.013	.475	-.001	.020	.955
		gpa	-.151	.049	.002	-.020	.064	.752	.036	.097	.707
		R^2	.054			.037			.003		

Notes. C = condition; IV = independent variable; SC = academic self-concept; MG = mastery goals; PG = performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A2.3. *Example 1: Results from Data Manipulated under the Assumption of MAR*

C	**Method**	**IV**	**SC**			**MG**			**PG**		
			β	*SE*	*P*	β	*SE*	*p*	B	*SE*	*p*
20%	LD	sex	.082	.045	.070	-.207	.072	.004	.030	.100	.761
		age	-.011	.006	.071	-.004	.010	.705	-.010	.014	.460
		gpa	-.216	.042	.000	-.091	.064	.151	.030	.089	.736
		R^2	.079			.031			.002		
20%	FIML	sex	.082	.045	.070	-.207	.072	.004	.030	.100	.761
		age	-.011	.006	.071	-.004	.010	.705	-.010	.014	.460
		gpa	-.216	.042	.000	-.091	.064	.151	.030	.089	.736
		R^2	.084			.032			.002		
20%	MI	sex	.086	.042	.040	-.234	.069	.001	.056	.093	.548
		age	-.009	.006	.147	-.002	.009	.865	-.014	.014	.311
		gpa	-.216	.038	.000	-.076	.058	.194	-.015	.083	.859
		R^2	.083			.060			.001		
40%	LD	sex	.120	.055	.028	-.171	.085	.045	-.018	.120	.879
		age	-.005	.009	.534	.007	.013	.617	-.020	.019	.292
		gpa	-.208	.047	.000	-.073	.071	.304	-.042	.101	.675
		R^2	.080			.019			.005		
40%	FIML	sex	.120	.055	.028	-.171	.085	.044	-.018	.120	.880
		age	-.005	.009	.534	.007	.013	.618	-.020	.019	.294
		gpa	-.208	.047	.000	-.073	.071	.306	-.042	.101	.675
		R^2	.120			-.171			-.018		
40%	MI	sex	.073	.046	.112	-.204	.081	.012	-.031	.096	.750
		age	-.015	.007	.041	-.002	.012	.849	-.015	.015	.320
		gpa	-.193	.040	.000	-.041	.067	.541	.031	.085	.716
		R^2	.062			.022			.003		

Note. C = condition; IV = independent variable; SC = academic self-concept; MG = mastery goals; PG = performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Table A2.4. *Example 1: Results from Data Manipulated under the Assumption of MNAR*

C	**Method**	**IV**	**SC**			**MG**			**PG**		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
20%	LD	sex	.038	.031	.221	-.152	.050	.002	-.146	.110	.183
		age	.008	.005	.133	.007	.006	.273	.004	.018	.826
		gpa	-.129	.032	.000	.001	.032	.979	-.021	.094	.820
		R^2	.116			.063			.008		
20%	FIML	sex	.062	.026	.016	-.108	.036	.003	-.025	.090	.778
		age	.009	.004	.039	.007	.005	.133	.010	.015	.507
		gpa	-.115	.025	.000	-.023	.026	.379	.002	.078	.975
		R^2	.125			.046			.002		
20%	MI	sex	.060	.030	.048	-.189	.049	.000	.007	.089	.936
		age	-.007	.005	.160	.007	.007	.305	-.010	.013	.463
		gpa	-.144	.029	.000	-.036	.036	.323	.021	.078	.785
		R^2	.083			.060			.001		
40%	LD	sex	.025	.019	.188	-.065	.053	.220	.017	.141	.902
		age	-.001	.002	.440	.002	.006	.758	.014	.022	.519
		gpa	-.016	.013	.219	.065	.048	.173	-.014	.110	.901
		R^2	.059			.113			.004		
40%	FIML	sex	.029	.017	.089	-.057	.035	.106	.029	.091	.747
		age	.004	.003	.110	.007	.005	.158	-.005	.014	.727
		gpa	-.050	.018	.006	.035	.029	.228	.087	.081	.284
		R^2	.079			.055			.006		
40%	MI	sex	.037	.022	.087	-.084	.058	.149	.046	.085	.590
		age	-.002	.003	.509	.001	.007	.849	-.009	.014	.529
		gpa	-.086	.024	.000	.043	.040	.286	.029	.080	.718
		R^2	.065			.034			.002		

Note. C = condition; IV = independent variable; SC = academic self-concept; MG = mastery goals; PG = performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A3.1. *Fit Statistics for all Example 2 Models Related to Mathematics*

Missing mechanism	**C**	**Method**	**$RMSEA_{ma}$**	**CI_{RMSEA}**	**$SRMR_{ma}$**	**CFI_{ma}**	**TLI_{ma}**	**χ^2 (*df* = 102)**
	Full		**.059**	**[.052, .066]**	**.046**	**.951**	**.937**	**320.319**
MCAR	20%	LD	.064	[.056, .072]	.049	.942	.926	315.620
	20%	FIML	.058	[.051, .065]	.047	.942	.926	315.620
	20%	MI	.056	[.049, .064]	.052	.942	.926	302.490
	40%	LD	.061	[.051, .071]	.052	.946	.931	241.949
	40%	FIML	.047	[.039, .055]	.050	.946	.931	241.949
	40%	MI	.047	[.040, .055]	.070	.945	.930	242.687
MAR	20%	LD	.056	[.048, .064]	.048	.948	.933	265.656
	20%	FIML	.051	[.043, .058]	.046	.948	.933	265.656
	20%	MI	.052	[.045, .060]	.056	.951	.938	274.572
	40%	LD	.056	[.046, .066]	.056	.940	.924	226.010
	40%	FIML	.044	[.036, .052]	.054	.940	.924	226.010
	40%	MI	.048	[.040, .055]	.067	.944	.929	245.047
MNAR	20%	LD	.061	[.050, .071]	.060	.930	.911	230.108
	20%	FIML	.052	[.044, .059]	.052	.939	.923	270.433
	20%	MI	.049	[.041, .056]	.063	.940	.923	251.849
	40%	LD	lavaan messages: "could not compute standard errors! lavaan NOTE: this may be a symptom that the model is not identified."					
			.057	[.034, .078]	.076	.935	.917	145.476
	40%	FIML	.042	[.034, .050]	.052	.949	.935	212.413
	40%	MI	.076	[.069, .083]	.071	.705	.624	468.717

Note. C = condition; $RMSEA_{ma}$ = root mean square errors of approximation; $SRMR_{ma}$ = standardized root mean square residual; CFI_{ma} = comparative fit index; TLI_{ma} = Tucker-Lewis index. LD = listwise deletion.

Table A3.2. *Example 2: Results from Full Data and Data Manipulated under the Assumption of MCAR*

C	**Method**	**IV**	**SC_{ma}** β	*SE*	*p*	**MG_{ma}** β	*SE*	*p*	**PG_{ma}** β	*SE*	*p*
Full		**sex**	**-.409**	**.061**	**.000**	**.116**	**.053**	**.029**	**.073**	**.063**	**.244**
		age	**.000**	**.023**	**.995**	**-.002**	**.019**	**.918**	**-.055**	**.024**	**.021**
		ts	**-.149**	**.097**	**.125**	**-.160**	**.084**	**.057**	**.262**	**.102**	**.010**
		gpa	**-.569**	**.033**	**.000**	**-.221**	**.032**	**.000**	**.130**	**.034**	**.000**
		R^2	**.451**			**.145**			**.055**		
MCAR 20%	LD	sex	-.440	.070	.000	.111	.059	.061	.068	.070	.330
		age	-.012	.025	.614	-.007	.021	.755	-.029	.025	.252
		ts	-.108	.111	.332	-.154	.095	.105	.319	.117	.006
		gpa	-.547	.037	.000	-.212	.035	.000	.123	.037	.001
		R^2	.434			.136			.056		
20%	FIML	sex	-.440	.070	.000	.111	.059	.060	.068	.071	.338
		age	-.012	.025	.614	-.007	.021	.755	-.029	.025	.253
		ts	-.108	.111	.332	-.154	.095	.105	.319	.116	.006
		gpa	-.547	.037	.000	-.212	.037	.000	.123	.037	.001
		R^2	.428			.133			.056		
20%	MI	sex	-.437	.067	.000	.106	.061	.080	.079	.070	.260
		age	-.017	.024	.469	-.010	.021	.632	-.028	.025	.257
		ts	-.118	.107	.270	-.141	.091	.122	.321	.118	.006
		gpa	-.545	.035	.000	-.210	.035	.000	.120	.038	.001
		R^2	.430			.132			.055		
40%	LD	sex	-.405	.080	.000	.061	.071	.394	.166	.093	.073
		age	-.007	.027	.798	-.018	.024	.452	-.084	.032	.009
		ts	-.140	.130	.280	-.277	.120	.021	.287	.152	.060
		gpa	-.575	.045	.000	-.215	.043	.000	.155	.051	.002
		R^2	.421			.138			.071		
40%	FIML	sex	-.405	.080	.000	.061	.072	.395	.166	.095	.079
		age	-.007	.027	.798	-.018	.024	.452	-.084	.032	.009
		ts	-.140	.130	.280	-.277	.121	.022	.287	.152	.060
		gpa	-.575	.045	.000	-.215	.046	.000	.155	.050	.002
		R^2	.452			.147			.078		
40%	MI	sex	-.433	.084	.000	.056	.071	.431	.165	.092	.073
		age	-.006	.028	.822	-.015	.026	.549	-.090	.035	.010
		ts	-.151	.143	.292	-.299	.121	.013	.307	.180	.087
		gpa	-.580	.046	.000	-.221	.048	.000	.155	.048	.001
		R^2	.452			.147			.078		

Note. C = condition; IV = independent variable; SC_{ma} = mathematics-related self-concept; MG_{ma} = mathematics-related mastery goals; PG_{ma} = mathematics-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

 Electronic Journal of Research in Educational Psychology, *15*(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A3.3. *Example 2: Results From Data Manipulated under the Assumption of MAR*

C	**Method**	**IV**	**SC_{ma}**			**MG_{ma}**			**PG_{ma}**		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
20%	LD	sex	-.342	.058	.000	.078	.050	.117	.031	.066	.639
		age	-.008	.018	.633	-.001	.015	.946	-.048	.021	.022
		ts	-.026	.095	.788	-.014	.081	.863	.196	.111	.077
		gpa	-.478	.032	.000	-.167	.031	.000	.139	.036	.000
		R^2	.436			.116			.057		
20%	FIML	sex	-.342	.059	.000	.078	.049	.114	.031	.067	.644
		age	-.008	.018	.633	-.001	.015	.946	-.048	.021	.023
		ts	-.026	.095	.788	-.014	.081	.863	.196	.110	.076
		gpa	-.478	.033	.000	-.167	.034	.000	.139	.036	.000
		R^2	.442			.117			.057		
20%	MI	sex	-.378	.061	.000	.089	.065	.169	.031	.063	.244
		age	-.017	.019	.359	-.008	.020	.682	-.059	.020	.021
		ts	-.054	.103	.602	-.108	.105	.302	.176	.104	.010
		gpa	-.537	.035	.000	-.294	.045	.000	.142	.034	.000
		R^2	.452			.177			.062		
40%	LD	sex	-.273	.057	.000	.074	.047	.118	-.030	.073	.680
		age	-.005	.016	.738	-.007	.013	.628	-.057	.022	.011
		ts	-.057	.095	.550	-.014	.078	.859	.263	.127	.039
		gpa	-.425	.034	.000	-.126	.032	.000	.116	.041	.005
		R^2	.439			.095			.058		
40%	FIML	sex	-.273	.057	.000	.074	.047	.113	-.030	.073	.683
		age	-.005	.016	.738	-.007	.013	.627	-.057	.022	.011
		ts	-.057	.095	.550	.014	.078	.859	.263	.127	.038
		gpa	-.425	.034	.000	-.126	.034	.000	.116	.041	.005
		R^2	.468			.104			.059		
40%	MI	sex	-.340	.065	.000	.077	.085	.363	-.035	.069	.611
		age	-.029	.019	.134	-.011	.025	.657	-.055	.020	.006
		ts	-.077	.104	.460	-.091	.128	.478	.247	.114	.031
		gpa	-.530	.042	.000	-.368	.066	.000	.124	.039	.001
		R^2	.457			.201			.063		

Note. C = condition; IV = independent variable; SC_{ma} = mathematics-related self-concept; MG_{ma} = mathematics-related mastery goals; PG_{ma} = mathematics-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Table A3.4. *Example 2: Results from Data Manipulated under the Assumption of MNAR*

C	Method	IV	SC_{ma}			MG_{ma}			PG_{ma}		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
20%	LD	sex	-.434	.074	.000	-.073	.068	.283	-.004	.027	.892
		age	-.013	.027	.645	.042	.026	.106	-.015	.012	.201
		ts	-.190	.116	.101	-.077	.107	.474	.108	.060	.072
		gpa	-.466	.041	.000	-.220	.044	.000	.042	.022	.053
		R^2	.450			.201			.064		
20%	FIML	sex	-.393	.059	.000	-.025	.049	.606	.019	.024	.422
		age	-.009	.022	.691	-.006	.016	.720	-.012	.009	.195
		ts	-.126	.094	.178	-.077	.074	.294	.114	.050	.024
		gpa	-.528	.033	.000	-.166	.040	.000	.046	.018	.013
		R^2	.498			.198			.073		
20%	MI	sex	-.401	.058	.000	-.007	.051	.885	.014	.022	.519
		age	.000	.021	.996	-.004	.018	.837	-.011	.009	.204
		ts	-.102	.095	.284	-.085	.080	.292	.107	.048	.024
		gpa	-.541	.032	.000	-.195	.042	.000	.045	.018	.010
		R^2	.502			.211			.072		
40%	LD	sex	-.437			.006			.000		
		age	-.063			.017			.000		
		ts	.206			-.021			.000		
		gpa	-.497			-.059			.000		
		R^2	.449			.078			.033		
40%	FIML	sex	-.402	.072	.000	-.081	.081	.317	.013	.015	.379
		age	-.027	.032	.410	-.010	.024	.665	-.002	.004	.726
		ts	.216	.127	.089	.037	.114	.748	.055	.044	.218
		gpa	-.604	.040	.000	-.176	.044	.000	.019	.015	.214
		R^2	.517			.323			.080		
40%	MI	sex	-.406	.067	.000	-.028	.067	.676	.014	.019	.469
		age	-.009	.028	.742	-.006	.021	.778	-.001	.005	.883
		ts	.021	.119	.862	-.019	.104	.852	.054	.043	.211
		gpa	-.566	.039	.000	-.178	.053	.001	.021	.016	.180
		R^2	.482			.368			.057		

Note. C = condition; IV = independent variable; SC_{ma} = mathematics-related self-concept; MG_{ma} = mathematics-related mastery goals; PG_{ma} = mathematics-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A4.1. *Fit Statistics for all Example 2 Models Related to Language*

Missing Mechanism	C	Method	$RMSEA_{la}$	CI_{RMSEA}	$SRMR_{la}$	CFI_{la}	TLI_{la}	χ^2 (*df* = 102)
	Full		**.054**	**[.047, .062]**	**.040**	**.947**	**.932**	**298.636**
MCAR	20%	LD	.057	[.049, .065]	.043	.943	.927	269.769
	20%	FIML	.039	[.031, .047]	.039	.952	.939	269.769
	20%	MI	.049	[.042, .057]	.048	.943	.927	256.744
	40%	LD	.051	[.040, .061]	.041	.952	.939	198.155
	40%	FIML	.051	[.044, .059]	.041	.943	.927	241.949
	40%	MI	.040	[.032, .048]	.054	.947	.933	204.084
MAR	20%	LD	.051	[.043, .060]	.042	.946	.932	237.214
	20%	FIML	.046	[.039, .054]	.040	.946	.932	237.214
	20%	MI	.047	[.039, .055]	.053	.950	.936	242.308
	40%	LD	.051	[.041, .061]	.043	.941	.924	204.485
	40%	FIML	.040	[.032, .048]	.042	.941	.924	204.485
	40%	MI	.042	[.034, .050]	.042	.942	.925	213.648
MNAR	20%	LD	.049	[.037, .060]	.051	.939	.922	183.906
	20%	FIML	.044	[.036, .051]	.046	.938	.921	222.513
	20%	MI	.043	[.036, .051]	.054	.929	.910	221.511
	40%	LD	.060	[.038, .080]	.067	.902	.875	150.159
	40%	FIML	.036	[.027, .044]	.057	.935	.917	182.407
	40%	MI	.060	[.038, .080]	.067	.902	.875	576.540

Note. C = condition; $RMSEA_{la}$ = root mean square errors of approximation; SRM-R_{la} = standardized root mean square residual; CFI_{la} = comparative fit index; TLI_{la} = Tucker-Lewis index. LD = listwise deletion.

Table A4.2. *Example 2: Results from Full Data and Data Manipulated under the Assumption of MCAR*

C	Method	IV	SC_{la} β	SC_{la} *SE*	SC_{la} *p*	MG_{la} β	MG_{la} *SE*	MG_{la} *p*	PG_{la} β	PG_{la} *SE*	PG_{la} *p*
Full		**sex**	**.122**	**.053**	**.021**	**.147**	**.058**	**.011**	**.047**	**.063**	**.459**
		age	**.004**	**.020**	**.830**	**.001**	**.021**	**.960**	**-.047**	**.024**	**.045**
		ts	**-.003**	**.085**	**.975**	**.015**	**.091**	**.866**	**.259**	**.103**	**.012**
		gpa	**-.426**	**.034**	**.000**	**-.214**	**.038**	**.000**	**.141**	**.040**	**.001**
		R^2	**.268**			**.098**			**.047**		
MCAR 20%	LD	sex	.105	.059	.074	.152	.062	.015	.038	.068	.574
		age	.011	.021	.584	.003	.022	.894	-.043	.025	.078
		ts	-.029	.096	.763	-.044	.099	.654	.264	.114	.020
		gpa	-.435	.037	.000	-.212	.040	.000	.136	.043	.002
		R^2	.268			.098			.047		
20%	FIML	sex	.060	.059	.074	.142	.062	.015	.166	.069	.575
		age	.015	.021	.584	.007	.022	.895	-.055	.025	.078
		ts	-.062	.096	.763	-.142	.099	.655	.222	.113	.020
		gpa	-.365	.037	.000	-.184	.041	.000	.165	.043	.002
		R^2	.211			.086			.056		
20%	MI	sex	.099	.059	.091	.157	.064	.014	.046	.068	.495
		age	.006	.026	.810	.018	.023	.442	-.041	.023	.078
		ts	-.019	.097	.847	-.033	.101	.742	.265	.112	.018
		gpa	-.430	.036	.000	-.214	.041	.000	.138	.043	.001
		R^2	.267			.105			.051		
40%	LD	sex	.060	.068	.384	.142	.074	.055	.166	.089	.062
		age	.015	.023	.519	.007	.024	.766	-.055	.030	.061
		ts	-.062	.112	.579	-.142	.121	.239	.222	.145	.126
		gpa	-.365	.046	.000	-.184	.049	.000	.165	.058	.004
		R^2	.201			.084			.055		
40%	FIML	sex	.105	.068	.384	.151	.074	.056	.038	.090	.064
		age	.011	.023	.519	.003	.024	.766	-.043	.030	.061
		ts	-.029	.112	.580	-.044	.121	.239	.264	.145	.125
		gpa	-.435	.046	.000	-.212	.050	.000	.136	.058	.004
		R^2	.271			.105			.049		
40%	MI	sex	.071	.071	.215	.137	.070	.051	.163	.092	.076
		age	.015	.023	.504	.013	.023	.569	-.064	.032	.045
		ts	-.070	.126	.579	-.166	.126	.188	.255	.150	.089
		gpa	-.352	.049	.000	-.188	.050	.000	.162	.065	.012
		R^2	.200			.093			.059		

Note. C = condition; IV = independent variable; SC_{la} = language-related self-concept; MG_{la} = language-related mastery goals; PG_{la} = language-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Table A4.3. *Example 2: Results from Data Manipulated under the Assumption of MAR*

C	**Method**	**IV**	**SC_{la}**			**MG_{la}**			**PG_{la}**		
			β	*SE*	*p*	β	*SE*	*p*	β	*SE*	*p*
20%	LD	sex	.121	.052	.020	.109	.052	.035	-.004	.069	.954
		age	.019	.018	.292	.017	.018	.339	-.055	.024	.025
		ts	-.066	.083	.425	-.001	.081	.994	.325	.114	.004
		gpa	-.371	.034	.000	-.171	.036	.000	.139	.045	.002
		R^2	.267			.097			.064		
20%	FIML	sex	.121	.052	.020	.109	.052	.035	-.004	.069	.954
		age	.019	.018	.293	.017	.018	.341	-.055	.024	.025
		ts	-.066	.083	.425	-.001	.081	.994	.325	.113	.004
		gpa	-.371	.035	.000	-.171	.037	.000	.139	.045	.002
		R^2	.272			.100			.062		
20%	MI	sex	.168	.055	.002	.164	.063	.009	-.017	.068	.809
		age	.000	.019	.996	.013	.020	.504	-.040	.023	.084
		ts	.014	.090	.881	.076	.098	.440	.306	.113	.007
		gpa	-.423	.038	.000	-.247	.043	.000	.148	.044	.001
		R^2	.267			.123			.054		
40%	LD	sex	.088	.055	.107	.090	.055	.100	-.053	.082	.522
		age	.006	.019	.740	-.018	.019	.346	-.067	.030	.023
		ts	-.068	.089	.445	-.012	.088	.891	.451	.143	.002
		gpa	-.316	.038	.000	-.127	.038	.001	.126	.055	.022
		R^2	.234			.064			.079		
40%	FIML	sex	.088	.055	.107	.090	.054	.099	-.053	.083	.522
		age	.006	.019	.740	.018	.019	.349	-.067	.030	.023
		ts	-.068	.089	.446	-.012	.088	.891	.451	.143	.002
		gpa	-.316	.039	.000	-.127	.039	.001	.126	.055	.023
		R^2	.251			.070			.078		
40%	MI	sex	.148	.058	.010	.104	.063	.100	-.057	.074	.441
		age	-.016	.020	.420	.028	.024	.232	-.063	.026	.015
		ts	.016	.103	.873	-.011	.113	.925	.450	.141	.001
		gpa	-.397	.045	.000	-.202	.055	.000	.148	.055	.007
		R^2	.273			.090			.083		

Note. C = condition; IV = independent variable; SC_{la} = language-related self-concept; MG_{la} = language-related mastery goals; PG_{la} = language-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school.

Table A4.4. *Example 2: Results from Data Manipulated under the Assumption of MNAR*

C	**Method**	**IV**	**SC_{la}**			**MG_{la}**			**PG_{la}**		
			β	*SE*	*p*	B	*SE*	*p*	β	*SE*	*p*
20%	LD	sex	.205	.062	.001	.076	.054	.157	-.005	.026	.861
		age	.003	.023	.906	-.004	.020	.829	-.011	.011	.323
		ts	.138	.098	.159	.145	.087	.095	.122	.066	.066
		gpa	-.319	.042	.000	-.136	.039	.000	.037	.023	.105
		R^2	.243			.083			.065		
20%	FIML	sex	.136	.052	.009	.075	.040	.060	.004	.030	.907
		age	.014	.019	.473	.000	.014	.996	-.013	.012	.254
		ts	.032	.084	.700	.041	.065	.529	.159	.062	.010
		gpa	-.339	.035	.000	-.112	.031	.000	.060	.024	.011
		R^2	.240			.082			.065		
20%	MI	sex	.139	.053	.005	.081	.058	.078	.003	.068	.912
		age	.005	.020	.778	.004	.021	.830	-.015	.025	.229
		ts	.028	.085	.718	.064	.091	.380	.152	.114	.013
		gpa	-.332	.034	.000	-.145	.038	.000	.058	.043	.013
		R^2	.240			.095			.060		
40%	LD	sex	.132	.089	.138	.055	.068	.420	.082	.078	.293
		age	.013	.040	.741	.011	.030	.722	.088	.049	.073
		ts	.037	.156	.813	.202	.129	.117	.151	.140	.280
		gpa	-.285	.061	.000	-.065	.045	.154	.044	.048	.366
		R^2	.230			.070			.151		
40%	FIML	sex	.047	.059	.425	.092	.040	.022	.002	.014	.862
		age	.003	.026	.903	.004	.011	.717	-.002	.005	.698
		ts	-.004	.105	.969	.047	.058	.416	.066	.054	.221
		gpa	-.347	.042	.000	-.048	.026	.063	.018	.015	.233
		R^2	.238			.076			.055		
40%	MI	sex	.057	.056	.312	.104	.043	.017	.014	.029	.644
		age	.011	.024	.660	.003	.015	.849	-.006	.010	.513
		ts	.001	.103	.991	.069	.078	.375	.086	.071	.226
		gpa	-.333	.043	.000	-.067	.036	.065	.020	.020	.304
		R^2	.218			.078			.032		

Note. C = condition; IV = independent variable; SC_{la} = language-related self-concept; MG_{la} = language-related mastery goals; PG_{la} = language-related performance avoidance goals; LD = listwise deletion; gpa = grade point average in school; R^2 = proportion of variance explained.

 Electronic Journal of Research in Educational Psychology, 15(1), 5-47. ISSN: 1696-2095. 2017. no. 41
http://dx.doi.org/10.14204/ejrep.41.16125

Appendix B

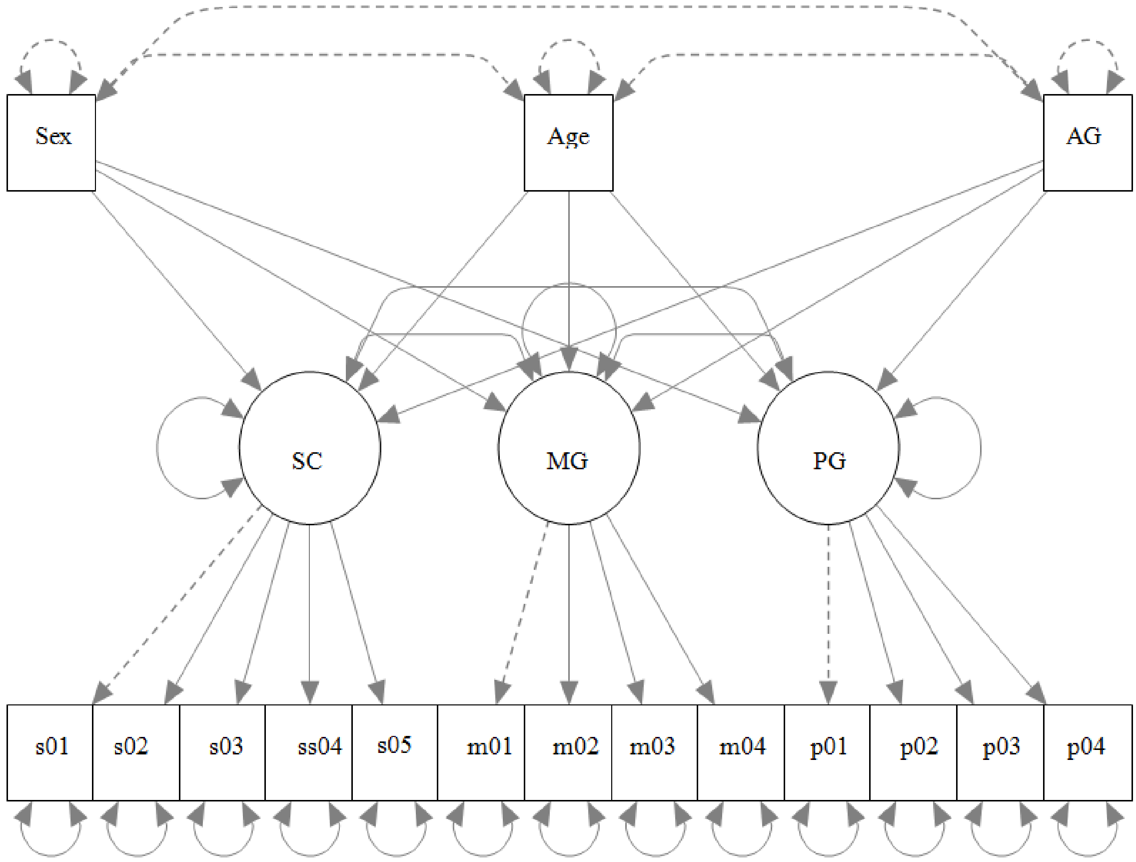

Figure B1. Basic model for SEM in Example 1, grade point average in school (AG), academic self-concept (SC), mastery goals (MG), performance avoidance goals (PG).

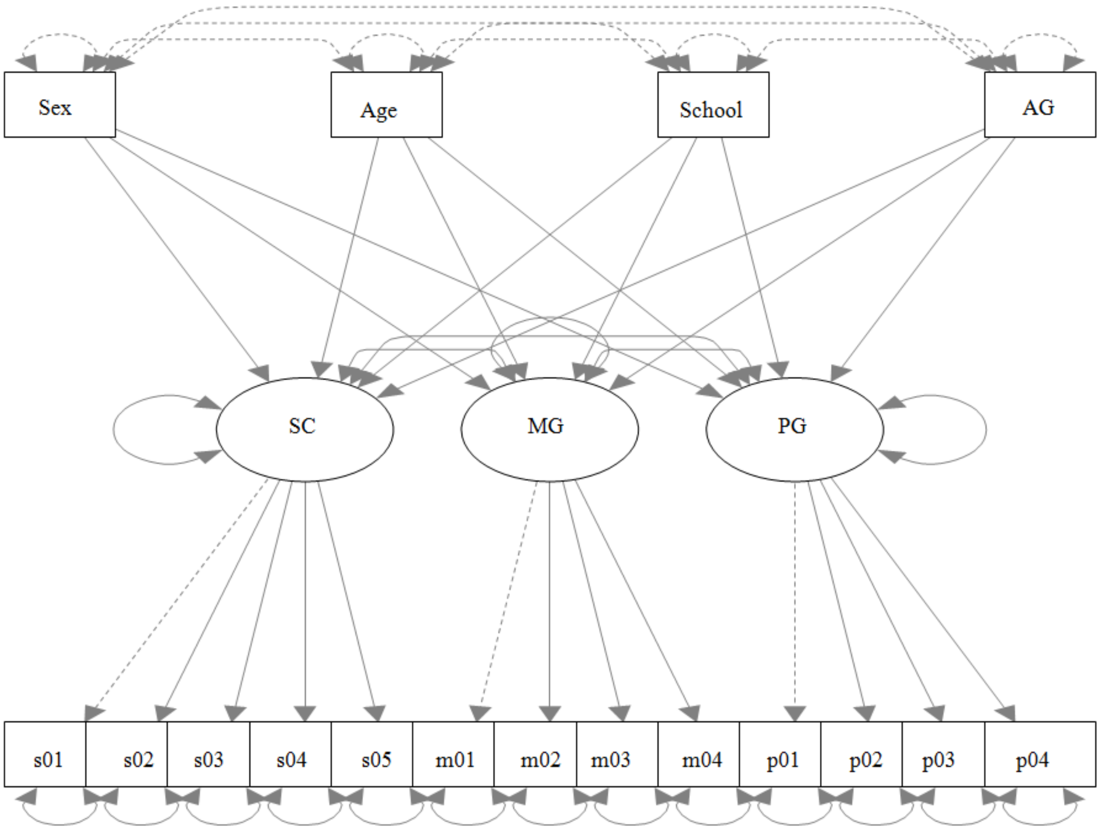

Figure B2. Basic model for SEM in Example 2, grade point average in school (AG), mathematics/language-related self-concept (SC), mastery goals (MG), performance avoidance goals (PG).

Appendix B3. R scripts used for SEM in Example 1 and SEM in Example 2.

```
### Data and variable names in Example 1 used:
# The data set was coded with data = ooomi
# Variables were coded with:
# sesskoabs = academic self-concept (items 1-5, equivalent items were used in Example 2)
# sellmolz = mastery goals (items 4, 5, 7, and 8 because equivalent items were used also in Example 2)
# sellmovl = performance avoidance goals (items 2, 3, 5, and 7 because equivalent items were used also in # Example 2)

# ska = latent factor academic self-concept
# lz = latent factor mastery goals
# vl = latent factor performance avoidance goals
# averagegrade = grade point average

###-----------------------------------------------------------------------------------------------------
### Model and commands for handling missing data in Example 1 used:

s.model <- '
ska  =~ sesskoabs1 +  sesskoabs2 +  sesskoabs3 +  sesskoabs4 +  sesskoabs5
lz  =~ sellmolz4 +  sellmolz5 + sellmolz7 +  sellmolz8
vl  =~ sellmovl2 +  sellmovl3 + sellmovl5 +  sellmovl7

ska ~ sex + age + averagegrade
lz ~ sex + age + averagegrade
vl ~ sex + age + averagegrade
'
#-------
# When listwise deletion used:
sout <- sem(s.model, data=ooomi)

# When FIML used instead of the line above:

# sout <- sem(s.model, data=ooomi, missing = "FIML")

summary(sout)
inspect(sout, "fit")
inspect(sout, "rsquare")

#-------
# When runMI used instead of the "sout"-object above (e.g., m = 40 imputations):
siout <- runMI(s.model, data=ooomi, m = 40, miPackage="Amelia", chi="all", seed=12345, fun="sem", fixed.x=FALSE)

summary(siout)
inspect(siout, "fit")
inspect(siout, "rsquare")
###-------------------------------------------------------------------------------------------------------
###-------------------------------------------------------------------------------------------------------
### Data and variable names in Example 2 used:
# The data set was coded with data = ooomi
# Variables were coded with:
# sk = academic self-concept (items 1-5)
# lz = mastery goals (items 1-4)
# lv = performance avoidance goals (items 1-4)
# skma = latent factor mathematics-related self-concept
# lzma = latent factor mathematics-related mastery goals
# vlma = latent factor mathematics-related performance avoidance goals

# skde = latent factor language-related self-concept
# lzma = latent factor language-related mastery goals
# vlma = latent factor language-related performance avoidance goals
# geschl = sex
# alter = age
# schulform = type of school
# zeug_math = grade point average in school
###-------------------------------------------------------------------------------------------------------
### Models and commands for handling missing data in Example 2 used:

ma.model <- '
skma  =~ sk01_mat +  sk02_mat +  sk03_mat +  sk04_mat +  sk05_mat
lzma  =~ lz01ma +  lz02ma +  lz03ma +  lz04ma
lvma =~ lv01ma +  lv02ma +  lv03ma +  lv04ma

skma ~ geschl + alter + schulform + zeug_math
lzma ~ geschl + alter + schulform + zeug_math
lvma ~ geschl + alter + schulform + zeug_math
'
```

```
de.model <- '
skde  =~ sk01_deu + sk02_deu + sk03_deu + sk04_deu + sk05_deu
lzde  =~ lz01de +  lz02de +  lz03de +  lz04de
lvde =~ lv01de +  lv02de +  lv03de +  lv04de

skde ~ geschl + alter + schulform + zeug_deut
lzde ~ geschl + alter + schulform + zeug_deut
lvde ~ geschl + alter + schulform + zeug_deut
'
# -------
# When listwise deletion used:
maout <- sem(ma.model, data=ooomi)
deout <- sem(de.model, data=ooomi)

# When FIML used instead of the line above:
# maout <- sem(ma.model, data=ooomi, missing = "FIML")
# deout <- sem(de.model, data=ooomi, missing = "FIML")
summary(maout)
inspect(maout, "fit")
inspect(maout, "rsquare")

summary(deout)
inspect(deout, "fit")
inspect(deout, "rsquare")
# -------
# When runMI used instead of the "maout"-object above (e.g., m = 40 imputations):
maiout <- runMI(ma.model, data=ooomi, m = 40,
        miPackage="Amelia", chi="all", seed=12345, fun="sem", fixed.x=FALSE, group.partial = "Klass_ID")
summary(maiout)
inspect(maiout, "fit")
inspect(maiout, "rsquare")

# when runMI used instead of the "deout"-object above (e.g., m = 40 imputations):
deiout <- runMI(de.model, data=ooomi, m = 40,
        miPackage="Amelia", chi="all", seed=12345, fun="sem", fixed.x=FALSE, group.partial = "Klass_ID")

summary(deiout)
inspect(deiout, "fit")
inspect(deiout, "rsquare")
#------
```

Appendix C

Appendix C. A 7-step tutorial for applying multiple imputations to SEM with missing data.

```
# (You should be familiar with R and lavaan before running runMI!)

# Step 1
# Diagnose the missing proportion per variable and patterns for specifying your mechanism assumption.
# Do you need to apply multiple imputation?

# Step 2
# Decide the number of imputations you need according to the largest proportion of missing values per variable
# (for your information see our text or Graham, 2009)

# Step 3
# Install and load the following packages:
install.packages("mice", dependencies=TRUE)
install.packages("mitools", dependencies=TRUE)
install.packages("miceadds", dependencies=TRUE)
install.packages("lavaan", dependencies=TRUE)
install.packages("Amelia", dependencies=TRUE)
install.packages("semTools", dependencies=TRUE)

library(mice)
library(mitools)
library(miceadds)
library(lavaan)
library(Amelia)
library(semTools)

# Step 4
# Have a look at the help page of the runMI() function and at its example.
?runMI
example("runMI")

# and adapt the runMi example to your model and data, e.g.:
out <- runMI(model, data, m=3)

# Step 5
# Run runMI() with 3 imputations (because 3 data sets are faster generated than e.g., 30) and check the plausibility by
summary(out)
inspect(out, "fit")
inspect(out, "impute")
inspect(out, "rsquare")

# Step 6
# If your results are plausible, change the number of imputations as many imputations as you need.

# Step 7
# If your results ar implausible, check your model and make sure that the model works as a simple SEM.
#------------
```


http://dx.doi.org/10.14204/ejrep.41.16125