

## **Factored Structural Equation Modeling in Blimp**

Craig K. Enders

University of California, Los Angeles

Brian T. Keller

University of Missouri

Cite as: Enders, C.K., & Keller, B.T. (2025). *Factored structural equation modeling in Blimp*. Manuscript submitted for publication.

### **Author Note**

Craig Enders is in the Department of Psychology at the University of California Los Angeles. <https://orcid.org/0000-0002-7048-8369>. Brian Keller is in the Department of Educational, School, and Counseling Psychology at the University of Missouri. <https://orcid.org/0009-0007-6670-2864>.

This work was supported by Institute of Educational Sciences award R305D190002.

There are no conflicts of interest to report. The ideas in this paper have not been presented elsewhere.

Correspondence concerning this article should be addressed to Craig Enders, UCLA Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095. Email: [cenders@ucla.edu](mailto:cenders@ucla.edu).

## Abstract

This tutorial introduces Factored Structural Equation Modeling (FSEM), an alternative to multivariate SEM, and demonstrates its implementation in the Blimp software and the `rblimp` R package. FSEM reconceptualizes the joint distribution of observed and latent variables as a set of univariate and multivariate submodels—each specified via simple regression equations—and treats latent variables as missing data to be imputed via Bayesian data augmentation. This approach seamlessly accommodates combinations of continuous (normal and nonnormal), binary, ordinal, nominal, count, and two-part variables; interactions; nonlinear effects; heteroscedasticity; and multilevel data structures without violating distributional assumptions. We first outline the theoretical foundations of factored regression specification and its connection to the probability chain rule. We then describe the Markov chain Monte Carlo estimation and missing-data imputation algorithms that underlie FSEM. Next, we present a series of illustrative models—ranging from basic confirmatory factor analysis to complex dynamic, multilevel, and hybrid generalized-linear-SEM applications—providing Blimp syntax excerpts and a real-data example. We conclude by discussing practical considerations, limitations, and directions for future methodological research. This tutorial aims to equip researchers with a flexible, user-friendly framework for modeling complex data in behavioral and social sciences.

## Introduction

Researchers familiar with the structural equation modeling (SEM) framework have likely learned the method through a multivariate lens, as described in numerous popular textbooks (Bollen, 1989; Bollen & Curran, 2005; Kaplan, 2009; Kline, 2023; Mulaik, 2009). A key feature of this framework is that the fitted model parameters combine to produce predictions for the mean vector and covariance matrix of a multivariate normal distribution. This paper introduces a factored regression approach to SEM, implemented in the Blimp software (Keller, 2024b; Keller & Enders, 2021), that recasts a multivariate distribution as a set of simpler submodels, each with its own distribution. For example, a factored specification represents a confirmatory factor analysis as a series of simple regression equations whose dependent variables potentially follow different conditional distributions. This equation-wise representation allows different variable types and complex nonlinear effects to coexist seamlessly without violating multivariate distributional assumptions.

At its core, Blimp's factored regression framework is a missing data methodology that incorporates latent variables, treating them as unobserved data to be imputed rather than integrated out of the likelihood (i.e., data augmentation; Schafer, 1997; Tanner & Wong, 1987). The theory for factorization traces to Ibrahim and colleagues (Ibrahim, 1990; Ibrahim et al., 2002; Lipsitz & Ibrahim, 1996), who developed the method to address incomplete covariates in generalized linear models. Factored specifications received considerable attention in the missing data literature, where they have been applied to interactive and nonlinear effects (Keller, in press; Kim et al., 2018; Kim et al., 2015; Lüdtke et al., 2020b; Zhang & Wang, 2017), discrete and nonnormal continuous variables (Lee & Mitra, 2016; Lüdtke et al., 2020b), models for missing not at random processes (Du et al., 2021; Huang et al., 2005; Ibrahim et al., 1999), auxiliary variable models (Daniels et al., 2014; Lüdtke et al., 2020b), multilevel models (Enders et al., 2020a; Erler et al., 2019b; Erler et al., 2016; Goldstein et al., 2014; Grund et al., 2021; Keller & Enders, 2023), and scale scores with missing item responses (Alacam et al., 2023).

Blimp generalizes these ideas to a wide range of structural equation models. The software supports both linear and generalized linear models for data structures with up to three levels (four if the lowest level is a wide-format latent curve model for repeated measurements). Models

can include virtually any combination of normal continuous (manifest and latent), nonnormal continuous (manifest and latent), binary, ordinal, multicategorical nominal, count, and two-part variables. Individual regression equations can include two-way and higher interactions, nonlinear effects, heteroscedastic variation, and functions of predictor variables (e.g., deviation scores, sum scores, residualized scores). Missing data architecture is integrated throughout the framework, allowing researchers to fit the same models to incomplete data that they would specify for complete data with no additional effort. The purpose of this tutorial is to provide a conceptual overview of this factored SEM framework and to demonstrate its implementation in Blimp and rblimp.

We refer to Blimp’s modeling framework as a factored structural equation model (FSEM) for simplicity of exposition. We do not intend to suggest that FSEM represents a new form of SEM. Rather, the acronym emphasizes that FSEM handles distributional assumptions differently from multivariate SEM. Blimp builds on earlier work in Bayesian SEM (Lee, 2007; Lee & Shi, 2000; Lee et al., 2007; Lüdtke et al., 2020b; Merkle & Rosseel, 2018; Palomo et al., 2007; Song & Lee, 2012), and its underlying factorization strategy also relates to likelihood-based formulations that employ analogous conditional distribution representations (Boker et al., 2023; Pritikin et al., 2018; Rabe-Hesketh et al., 2004; Rockwood, 2020).

### **Getting Started With Blimp and rblimp**

Blimp and rblimp are free software tools available at [www.appliedmissingdata.com/blimp](http://www.appliedmissingdata.com/blimp). A detailed user guide with dozens of analysis examples is available at the same URL. The native graphical interface, Blimp Studio, runs on macOS and Windows, with command-line versions available for Linux<sup>1</sup>. The rblimp R package allows researchers to interact with Blimp through R, providing access to R’s scripting and visualization capabilities, including plotting functions not available in the graphical interface. Because rblimp uses the Blimp computational engine, the native application must be installed before installing the R package. The online materials

---

<sup>1</sup> The Blimp computational engine is written in C++ for efficiency. Its performance is comparable to, or faster than, other competing programs. The online materials include a detailed speed comparison of the MCMC estimators implemented in Blimp, Mplus, and blavaan (<https://github.com/blimp-stats/fsem/tree/main/benchmark>).

include detailed installation instructions, and Section 1.3 of the software's user guide provides additional details about `rblimp`.

Model specification uses a simple scripting language that requires no prior knowledge of factored regression or Bayesian estimation. To illustrate, the Appendix shows the Blimp Studio and `rblimp` scripts for the basic SEM in Figure 1a, with inline comments (`#`) explaining the purpose of each line. A typical Blimp script consists of several major commands (ending with a colon), each with its own subcommands or keywords (ending with a semicolon). Broadly speaking, most Blimp commands fall into three categories: those that describe variables and their attributes (e.g., `DATA`, `VARIABLES`, `MISSING`, `ORDINAL`, `NOMINAL`, `COUNT`, `LATENT`), define the analysis model (e.g., `CENTER`, `MODEL`, `SIMPLE`, `PARAMETERS`), or control the MCMC algorithm (e.g., `BURN`, `ITER`, `SEED`). With the exception of variable attributes stored in the R data frame, the commands in a Blimp Studio script correspond to input arguments in the `rblimp()` function. For brevity, the syntax excerpts in this paper focus primarily on the `MODEL` section. The online materials include a quick start guide summarizing key commands, annotated working scripts for all examples in the paper, and an R Markdown file describing the syntax and output for a real-data analysis example. Finally, several YouTube training videos are available at [www.appliedmissingdata.com/videos](http://www.appliedmissingdata.com/videos).

### **Factored Structural Equation Models (FSEM)**

Blimp's FSEM framework uses combinations of three distribution types to construct models: univariate conditional distributions for outcome variables, multivariate distributions for manifest exogenous variables, and multivariate distributions for outcome variables. Table 1 summarizes the permissible variable types for each of the three submodels. Univariate submodels provide options that span most behavioral science applications: normal and nonnormal continuous variables (manifest or latent), common discrete metrics, and two-part models for floor and ceiling effects. These models can include myriad interactive and nonlinear effects with no special adjustments or specification tricks. The two types of multivariate submodels also allow various combinations of continuous and discrete variables. As explained later, binary and ordinal variables appear as normally distributed latent response variables in these multivariate functions. The univariate outcome options in Table 1 include functionality for nonnormal variables, but

they otherwise overlap with *Mplus*. Distinctions between Blimp and other SEM software are more evident in multivariate models, where Blimp provides a wider array of options, particularly with missing data. We highlight some of these differences throughout the paper.

To illustrate a basic factored regression specification, consider the model in Figure 1a. To establish connections with the multivariate SEM, we temporarily assume that all variables are continuous and multivariate normal. FSEM represents Figure 1a as six submodels, each with its own univariate distribution assumption. We use shaded rectangles to enclose variables that belong to the same submodel (and thus share a distribution). To reduce visual clutter, we use black dots to represent intercepts and means.

To begin, the exogenous latent variable is the outcome in an empty regression model with just a mean and disturbance term.

$$\begin{aligned}\eta_i &= \beta_{0\eta} + \varepsilon_{\eta,i} = \bar{\eta}_i + \varepsilon_{\eta,i} \\ f(\eta) &\rightarrow \mathcal{N}(\bar{\eta}_i, \sigma_{\varepsilon_\eta}^2)\end{aligned}\tag{1}$$

Our notation warrants brief clarification. First, to emphasize that an FSEM consists of a collection of regression models, we use  $\beta$ s to denote coefficients and attach the dependent variable's name as an index (e.g.,  $\beta_{0\eta}$  is the intercept in  $\eta$ 's equation). Second, the overbar indicates the model-implied or systematic part of the variable for person  $i$  (i.e., a conditional mean or predicted score). Third,  $f(\cdot)$  is generic shorthand for the model-implied distribution associated with a regression equation. In this case,  $f(\eta)$  in the second line represents a univariate distribution involving only the latent variable (i.e., a marginal distribution). We do not use a tilde ( $\sim$ ) because this function alone does not define the predictive distribution of the latent variable (i.e., the distribution that generates its imputations). Instead,  $f(\eta)$  serves as shorthand for one of the density functions that collectively defines the model structure. Finally, the arrow ( $\rightarrow$ ) assigns a specific functional form to the generic distribution function. In this case,  $f(\eta)$  corresponds to a normal distribution with a conditional mean (predicted value) and variance as its parameters.

Next, the measurement model consists of four univariate submodels, one per indicator. The first line below shows indicator  $p$ 's measurement model, and the second line is its model-implied conditional distribution.

$$\begin{aligned}
 X_{p,i} &= \beta_{0X_p} + \beta_{1X_p}(\eta_i) + \varepsilon_{X_p,i} = \bar{X}_{p,i} + \varepsilon_{X_p,i} \\
 f(X_p|\eta) &\rightarrow \mathcal{N}(\bar{X}_{p,i}, \sigma_{\varepsilon_{X_p}}^2)
 \end{aligned}
 \tag{2}$$

The  $f(X_p|\eta)$  term can be read as “the distribution of  $X_p$  given  $\eta$ ”, with a vertical pipe separating the outcome on the left from its predictors on the right. To reiterate, we use  $\beta$ s for the measurement model parameters to deemphasize the link to multivariate SEM. In this example, the indicator submodels are also conditional normal distributions, with predicted values (conditional means) and residual variances defining the center and spread, respectively. Finally, the structural regression and its conditional distribution are as follows.

$$\begin{aligned}
 Y_i &= \beta_{0Y} + \beta_{1Y}(\eta_i) + \varepsilon_{Y,i} = \bar{Y}_i + \varepsilon_{Y,i} \\
 f(Y|\eta) &\rightarrow \mathcal{N}(\bar{Y}_i, \sigma_{\varepsilon_Y}^2)
 \end{aligned}
 \tag{3}$$

The key feature of FSEM is that it recasts a multivariate distribution into a sequence of simpler submodels, each with its own conditional distribution. The statistical theory behind this approach uses the probability chain rule to express a multivariate distribution as the product of conditional distributions that exactly reproduce the original joint function (Huang et al., 2005; Ibrahim et al., 2002; Ibrahim et al., 1999; Lipsitz & Ibrahim, 1996). Throughout the paper, we use the following shorthand to describe a model's underlying factorization.

$$f(Y, X_4, X_3, X_2, X_1, \eta) = \tag{4}$$

$$f(Y|X_4, X_3, X_2, X_1, \eta) \cdot f(X_4|X_3, X_2, X_1, \eta) \cdot f(X_3|X_2, X_1, \eta) \cdot f(X_2|X_1, \eta) \cdot f(X_1|\eta) \cdot f(\eta) =$$

$$f(Y|\eta) \cdot f(X_4|\eta) \cdot f(X_3|\eta) \cdot f(X_2|\eta) \cdot f(X_1|\eta) \cdot f(\eta)$$

The first line is the multivariate or joint distribution. The linear SEM defines  $f(\cdot)$  as a multivariate normal distribution. Beyond acknowledging that the multivariate distribution exists, FSEM makes no assumptions about its shape or form. Rather, it works with the submodels and distributions on the second and third lines. As a reminder, each  $f(\cdot | \cdot)$  represents the model-implied distribution associated with a univariate regression model, where the vertical pipe separates an outcome on the left from its predictors on the right. Applying the probability chain rule to the multivariate distribution produces the saturated factorization on the second line where each variable links to all others. Figure 1b shows a path diagram of this model, which includes all possible associations among the six variables. Reading from right to left, the order of terms on the second line follows the flow of the directional arrows in the path diagram. Finally, the third line eliminates the  $X$ s from the right side of the vertical pipes to match the omitted paths in Figure 1a. In practical terms, the equation implies that FSEM defines the (unspecified) multivariate distribution as the product of six conditional submodels whose dependent variables potentially follow different distributions<sup>2</sup>.

Although its underlying assumptions and estimation machinery differ from those of multivariate SEM, model specification in Blimp closely parallels that of other SEM software. To illustrate, consider the SEM in Figure 1a. The Appendix shows the complete Blimp and rblimp scripts for this example. The syntax excerpt below illustrates how to specify every model parameter explicitly.

```
LATENT: eta;
```

---

<sup>2</sup> Factorization assumes a directed ordering of variables in which earlier variables appear on the right-hand side (conditioning set) of later ones. Nonrecursive models with feedback loops are not compatible with this formulation because the conditional specifications  $f(X_1|X_2)$  and  $f(X_2|X_1)$  cannot, in general, simultaneously define a coherent joint distribution.

```

MODEL:
# regression equations
eta ~ intercept@0;
x1 ~ intercept eta@1;
x2 ~ intercept eta;
x3 ~ intercept eta;
x4 ~ intercept eta;
y ~ intercept eta;
# variances and residual variances
eta ~~ eta; x1 ~~ x1; x2 ~~ x2; x3 ~~ x3; x4 ~~ x4; y ~~ y;

```

The LATENT command defines a new latent variable called eta. The MODEL command lists six regression equations, each with an outcome variable to the left of a tilde and its predictors to the right. For identification, the first regression equation fixes the latent variable's mean to zero (intercept@0), and the second fixes the first indicator's loading to 1 (eta@1). The final line specifies variances and residual variances.

The model can be specified more succinctly by relying on the following software defaults: latent variable means are fixed at 0, manifest variable intercepts are estimated, the first threshold (for categorical variables) is fixed at 0 while the remaining thresholds are estimated, and all variances and residual variances are freely estimated. The simplified syntax excerpt below leverages these defaults.

```

LATENT: eta;
MODEL:
x1 ~ eta@1;
x2 ~ eta;
x3 ~ eta;
x4 ~ eta;
y ~ eta;

```

An even more compact specification uses the right-arrow shortcut to assign the latent variable on the left as a predictor of the outcomes to the right.

```
# model 1 excerpt
LATENT: eta;
MODEL:
eta -> x1:x4;
y ~ eta;
```

When applying this approach to measurement models, Blimp automatically fixes the first loading to 1 for identification.

We generally prefer a fixed-factor identification strategy (Klopp & Klößner, 2021) that constrains the factor variance or disturbance variance to 1, as it eliminates variance priors that inherently introduce some information into the structural model. Simulation studies indicate that pairing MCMC with this approach has better small-sample properties than the fixed marker strategy (Keller, in press). The syntax excerpt below shows a compact version of this specification.

```
# model 1 excerpt
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta -> x1@1o1 x2:x4;
y ~ eta;
```

Using the @ symbol to attach an arbitrary label to the first loading frees it for estimation, and attaching @1 to the factor variance fixes its scale.

Of course, building realistic models solely from univariate regressions is not feasible. As explained previously, Blimp uses factorizations that incorporate two types of multivariate submodels: one reserved for manifest exogenous variables and another for outcomes. To

illustrate the former, consider the model in Figure 2. The  $\eta$  and  $Y$  models now include two additional predictors, but their conditional distributions are still univariate normal curves.

$$\begin{aligned}
 \eta_i &= \beta_{0\eta} + \beta_{1\eta}(M_{1,i}) + \beta_{2\eta}(M_{2,i}) + \varepsilon_{\eta,i} = \bar{\eta}_i + \varepsilon_{\eta,i} \\
 f(\eta|M_1, M_2) &\rightarrow \mathcal{N}(\bar{\eta}_i, \sigma_{\varepsilon_\eta}^2) \\
 Y_i &= \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(M_{1,i}) + \beta_{3Y}(M_{2,i}) + \varepsilon_{Y,i} = \bar{Y}_i + \varepsilon_{Y,i} \\
 f(Y|\eta, M_1, M_2) &\rightarrow \mathcal{N}(\bar{Y}_i, \sigma_{\varepsilon_Y}^2)
 \end{aligned} \tag{5}$$

The measurement models are the same as Equation 2. The underlying factorization is as follows.

$$\begin{aligned}
 f(Y, X_4, X_3, X_2, X_1, \eta, M_1, M_2) &= \\
 f(Y|X_4, X_3, X_2, X_1, \eta, M_1, M_2) \cdot f(X_4|X_3, X_2, X_1, \eta, M_1, M_2) \cdot \dots \cdot f(\eta|M_1, M_2) \cdot f(M_1, M_2) \\
 f(Y|\eta) \cdot f(X_4|\eta) \cdot f(X_3|\eta) \cdot f(X_2|\eta) \cdot f(X_1|\eta) \cdot f(\eta|M_1, M_2) \cdot f(M_1, M_2)
 \end{aligned} \tag{6}$$

The second line again shows a saturated model factorization, and the third line simplifies the expression to match the constraints (omitted paths) depicted in the figure. Notice that the submodel sequence ends without factoring the bivariate distribution of the manifest predictors,  $f(M_1, M_2)$ . The shaded rectangle enclosing these variables in Figure 2 conveys that they share a distribution. To ensure appropriate missing data handling, Blimp constructs this distribution automatically, so users need to specify this part of the model only when using a variable type restricted to univariate models (see Table 1).

The Blimp syntax excerpt below specifies the model in Figure 2. To reiterate, the script uses a fixed factor scaling that constrains the disturbance variance to 1 (see Klopp & Klößner, 2021, Table 4).

```
# model 2 excerpt
MODEL:
eta ~~ eta@1;
eta ~ m1 m2;
eta -> x1@1o1 x2:x4;
y ~ eta m1 m2;
```

Importantly, the manifest predictors appear only on the right side of the tildes; they do not have their own regression equations. Manifest exogeneous variables like  $M_1$  and  $M_2$  do not require distributional assumptions or an explicit model when their data are complete, but they do when values are missing. Because Blimp was developed with missing data in mind, the software automatically invokes a multivariate normal distribution for these variables with no input from the user, and it omits unnecessary components in situations where it makes sense to do so.

Dependent variables to the left of a tilde can also share a multivariate submodel. To illustrate, consider the model in Figure 3. The two covariances (double-headed curved arrows) require a pair of bivariate submodels: one connecting the latent and manifest predictors and another for the correlated residuals. Again, the shaded rectangles enclosing these variables convey a common distribution. The bivariate normal submodel for the predictors is shown below.

$$\begin{aligned}
 \eta_i &= \beta_{0\eta} + \varepsilon_{\eta,i} = \bar{\eta}_i + \varepsilon_{\eta,i} \\
 M_i &= \beta_{0M} + \varepsilon_{M,i} = \bar{M}_i + \varepsilon_{M,i}
 \end{aligned}
 \tag{7}$$

$$f(\eta, M) \rightarrow \mathcal{N} \left( \begin{pmatrix} \bar{\eta}_i \\ \bar{M}_i \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon_\eta}^2 & \sigma_{\varepsilon_\eta \varepsilon_M} \\ \sigma_{\varepsilon_M \varepsilon_\eta} & \sigma_{\varepsilon_M}^2 \end{pmatrix} \right)$$

Notice that the manifest variable  $M$  does not require a proxy latent variable (a factor with  $M$  as its sole indicator) to correlate with  $\eta$ . The bivariate normal submodel for the correlated indicators has the same form.

$$\begin{aligned}
X_{3,i} &= \beta_{0X_3} + \beta_{1X_3}(\eta_i) + \varepsilon_{X_3,i} = \bar{X}_{3,i} + \varepsilon_{X_3,i} \\
X_{4,i} &= \beta_{0X_4} + \beta_{1X_4}(\eta_i) + \varepsilon_{X_4,i} = \bar{X}_{4,i} + \varepsilon_{X_4,i} \\
f(X_3, X_4 | \eta) &\rightarrow \mathcal{N} \left( \begin{pmatrix} \bar{X}_{3,i} \\ \bar{X}_{4,i} \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon_{X_3}}^2 & \sigma_{\varepsilon_{X_3}\varepsilon_{X_4}} \\ \sigma_{\varepsilon_{X_4}\varepsilon_{X_3}} & \sigma_{\varepsilon_{X_4}}^2 \end{pmatrix} \right)
\end{aligned} \tag{8}$$

Historically, residual correlations like these have been the Achilles heel of factorized specifications, and early implementations disallowed these associations (Lee, 2007). Subsequent work used phantom latent variables to induce these associations (Merkle & Rosseel, 2018; Palomo et al., 2007), but this approach is very slow and too limited for general use. Embedding multivariate distribution functions into the factorization addresses this limitation (see Keller, 2025; Keller, in press).

The underlying FSEM factorization for the model in Figure 3 is as follows.

$$\begin{aligned}
f(Y, X_1, X_2, X_3, X_4, \eta, M) &= \\
f(Y|X_4, X_3, X_2, X_1, \eta, M) \cdot f(X_4, X_3|X_2, X_1, \eta, M) \cdot f(X_2|X_1, \eta, M) \cdot f(X_1|\eta, M) \cdot f(\eta, M) \\
f(Y|\eta, M) \cdot f(X_4, X_3|\eta) \cdot f(X_2|\eta) \cdot f(X_1|\eta) \cdot f(\eta, M)
\end{aligned} \tag{9}$$

The second line again shows the factorization for a saturated model, and the third line simplifies the expression to match the restrictions (omitted paths) in the diagram. Multivariate submodels with correlations are invoked by connecting pairs of variables with a double tilde, as follows.

```

# model 3 excerpt
MODEL:
eta ~~ eta@1;
eta ~~ m;
eta -> x1@1o1 x2:x4;

```

```
x3 ~~ x4;
y ~ eta m;
```

Again, the underlying factorization is completely invisible to the user.

### **MCMC and Missing Data Imputation**

Blimp uses MCMC algorithms developed in the Bayesian paradigm. Bayesian analyses have increasingly gained a strong foothold in social and behavioral science research applications (Andrews & Baguley, 2013; König & van de Schoot, 2018; van de Schoot et al., 2017). There are numerous reasons to use Bayesian methods and different perspectives one can adopt when applying them (Levy & McNeish, 2023). Some find the Bayesian framework appealing on philosophical grounds. Representing uncertainty as a distribution of plausible parameters that could have produced the sample data can be more intuitive than imagining a distribution of estimates from other hypothetical samples. For entirely different reasons, researchers are increasingly turning to the Bayesian framework because its flexible computational algorithms are adept at fitting models for complex data structures. Researchers adopting this approach may even use MCMC-generated parameter summaries for frequentist inference (computational frequentism; Levy & McNeish, 2023). Blimp supports both inferential approaches (see the later section on model fit).

Blimp's MCMC algorithm iterates between two main steps. First, it estimates the parameters of each submodel using a complete data set of imputed observed scores and latent variables. Second, it updates the data by imputing missing scores and latent data, using the latest model parameters to derive the predicted values and spread of the imputations. Repeating this two-step process thousands of times produces posterior distributions of plausible parameter values that could have produced the sample data. The median and standard deviation of the parameters are Bayesian point estimates and "standard errors". The 2.5% and 97.5% quantiles of the posterior distribution define 95% interval limits, which are analogous to frequentist confidence intervals but do not rely on the notion of repeated sampling.

Blimp primarily uses a Gibbs sampling algorithm, although certain estimation steps require Metropolis-within-Gibbs (Hastings, 1970) approximations. Gibbs sampling with data

augmentation is appealing because MCMC estimation steps leverage simple, exact solutions for linear regression models. This is also true for generalized linear models, which represent discrete outcomes as continuous latent variables. Blimp's estimation steps are described in various resources, online supplements, and technical appendices (Du et al., 2021; Enders, 2022; Enders et al., 2020b; Keller, 2025, in press; Keller & Enders, 2023). By default, the software adopts common noninformative (or minimally informative, in the case of variances and covariances) prior distributions described in the literature. Our preference to fix the factor variance or disturbance variance to 1 is largely driven by the desire to eliminate or mitigate the impact of variance priors on the structural model. Keller and Enders (2021, Section 1.7) summarize Blimp's default prior distributions, and advanced users can also specify custom priors.

After updating the model parameters, MCMC uses computer simulation to “sample” imputations from distributions based on those quantities. The software treats three quantities as missing data to be imputed: incomplete manifest variables, latent factors, and latent response variables underlying categorical and discrete measures. Computationally, latent factors behave like other continuous manifest variables (normal or skewed) except that they are fully missing. The generalized linear models described in the next section use a link function to map discrete responses onto a continuous scale. This transformed variable, which we denote  $Y^*$ , becomes the outcome in a linear regression. Blimp's MCMC algorithm uses a data augmentation strategy that pairs each person's discrete response with one or more continuous latent variables representing the transformed metric. MCMC imputes these latent response scores at each computational cycle, after which it uses them to estimate the generalized linear model parameters. Augmenting estimation with latent data is computationally appealing because MCMC can utilize simple Gibbs sampling steps for linear regression models. Data augmentation has a long history with probit regression (Albert & Chib, 1993; Cowles, 1996), and more recent work has extended this strategy to logit and negative binomial link functions (Asparouhov & Muthén, 2021; Neelon, 2019; Pillow & Scott, 2012; Polson et al., 2013; Windle et al., 2014). The latent response framework is perhaps the most widely used approach for imputing incomplete discrete variables in the missing data literature, and numerous studies have examined its performance (Carpenter et al., 2023; Carpenter et al., 2011; Enders, 2022; Enders et al., 2018; Goldstein et al., 2009; Quartagno & Carpenter, 2019).

At a high level, an imputed score can be viewed as the sum of a predicted value and a computer-generated random noise term (typically drawn from a normal distribution). In a factored specification, a variable's imputation-generating distribution (its posterior predictive distribution) depends on every submodel in which it appears. To illustrate, reconsider the path diagram in Figure 1a. Each manifest variable serves as the outcome in a simple regression model and has no outgoing arrows. In these cases, imputations depend only on that single model. For example, MCMC imputes the dependent variable by substituting the current parameters and latent data into the regression equation to obtain a person's predicted score,  $\bar{Y}_i$ . It then adds a synthetic  $\varepsilon_Y$  term, simulated from a normal distribution with a mean of 0 and a variance equal to the current value of  $\sigma_{\varepsilon_Y}^2$ . This process aligns with the conditional distribution shown on the bottom line of Equation 3, where  $Y$  scores are normally distributed around points on the regression line. MCMC imputes each  $X_p$  using the same process.

Imputing latent variables is the secret sauce that simplifies complex estimation problems. MCMC also creates latent variable imputations as the sum of a predicted value and noise term. However, these inputs now depend on multiple models. In Figure 1a, the latent variable appears as the outcome in one model (Equation 1) and as a predictor in five others (Equations 2 and 3). Applying probability rules gives the symbolic expression for  $\eta$ 's conditional distribution given all other variables (i.e., its imputation-generating distribution).

$$f(\eta|Y, X_4, X_3, X_2, X_1) \propto f(Y|\eta) \cdot f(X_4|\eta) \cdot f(X_3|\eta) \cdot f(X_2|\eta) \cdot f(X_1|\eta) \cdot f(\eta) \quad (10)$$

In practical terms, the equation states that every model including  $\eta$  contains information about its imputations. In this example, the posterior predictive distribution on the left side of Equation 10 is obtained by multiplying six normal distribution functions and then performing algebraic manipulations to express the result as a function of  $\eta$ . This distribution is also a normal curve, but the predicted values (conditional means) and conditional variance are complex composites of the submodel parameters. Interested readers can find a derivation of  $\eta$ 's posterior predictive distribution in the online materials (<https://github.com/blimp-stats/fsem>).

Beyond generating multiply imputed factor scores, there is no compelling advantage to imputing latent variables in simple models like the one in Figure 1a. However, composite distributions such as the one in Equation 10 adapt to nonlinear features in the submodels. For example, we later present an application where  $\eta$  interacts with a multicategorical nominal variable represented by two dummy codes. In this situation,  $\eta$ 's imputations originate from a mixture of heteroscedastic normal distributions in which the conditional variance depends on a simple slope and a person's dummy codes. In some cases, these distributions can be solved algorithmically (Keller, 2025); in others, their complexity requires a Metropolis-within-Gibbs approximation. This occurs, for example, when  $\eta$  (or another predictor) is involved in a curvilinear effect with polynomial terms (Lüdtke et al., 2020b). For additional information on these composite distributions and worked examples of the Metropolis-within-Gibbs step, see Enders (2022).

### **Generalized Linear Models for Discrete Variables**

One compelling advantage of FSEM is that any given model can include myriad combinations of linear and generalized linear regression equations. Returning to Table 1, Blimp currently offers generalized linear models for binary (probit or logit link), ordinal (probit link), multicategorical nominal (logit link), and overdispersed count outcomes (negative binomial link) that take nonnegative integer values indicating the number of times an event occurs. Blimp places minimal restrictions on generalized linear models within an FSEM. Discrete variables can serve as exogenous predictors, intermediate variables (e.g., mediators), or terminal outcomes. Further, discrete variables can interact with any other variable in a model. For instance, an FSEM might include a latent-by-manifest variable interaction involving a multicategorical moderator and a latent factor. Finally, Blimp incorporates multivariate submodels that enable correlations among binary and ordinal variables via their latent response scores. Among other things, this functionality supports residual correlations between discrete indicators and associations involving discrete and continuous predictors.

Currently, the only restriction on generalized linear models is that an incomplete count variable modeled with negative binomial regression cannot have an outgoing arrow; all other variable types can. One of the unique features of Blimp's factorization is that a variable can exist

in two states. For example, a binary predictor functions as a latent response variable in the exogenous submodel but as a dummy code in the structural model. In essence, variables with both incoming and outgoing arrows can occupy different metrics depending on the direction of the arrow. This flexibility contrasts with other SEM programs, which impose stricter limits on variable types with outgoing arrows, especially with incomplete missing data (e.g., multivariate normality is often assumed, with limited or no functionality for categorical predictors). Several examples in this section cannot be estimated with incomplete data in other software packages.

### Probit Regression

Blimp defaults to probit regression method for binary and ordered categorical outcomes, and it also uses a multinomial probit model for multicategorical exogenous variables. The probit link for binary and ordered categorical outcomes introduces a normally distributed latent response variable representing a continuous dimension of the measured characteristic. To illustrate, consider the model in Figure 4a. This is a modified version of Figure 3, where  $M$  is a binary dummy code, and the factor indicators are five-point rating scales. The four measurement models feature latent response dependent variables, as does the empty model for the binary predictor. The regression equations below use an asterisk superscript to denote the underlying latent variables.

$$\begin{aligned} X_{p,i}^* &= \beta_{0X_p^*} + \beta_{1X_p^*}(\eta_i) + \varepsilon_{X_p^*,i} = \bar{X}_{p,i}^* + \varepsilon_{X_p^*,i} \\ M_i^* &= \beta_{0M^*} + \varepsilon_{M^*,i} = \bar{M}_i^* + \varepsilon_{M^*,i} \end{aligned} \tag{11}$$

The figure uses ellipses to represent latent responses and rectangles to denote discrete variables. For identification, both residuals are defined as standard normal variables (i.e., latent response variables are standardized within subgroups defined by the same predicted value).

The probit model also introduces threshold parameters that segment the latent response distribution. The areas under the normal curve between these thresholds define the discrete category probabilities. The binary covariate requires a single threshold, and five-point indicators

require four cutpoints each. The following expression gives the relationship between a latent response variable  $V^*$  and a discrete variable  $V$  with response options numbered  $c = 1$  to  $C$ .

$$V_i = c \text{ if } \tau_{c-1} < V_i^* \leq \tau_c \quad (12)$$

In words, the equation says that the latent response scores for discrete category  $c$  fall between two adjacent cutpoints in the normal curve,  $\tau_{c-1}$  and  $\tau_c$ . The lowest category  $c = 1$  is bounded by an imaginary threshold at negative infinity (i.e.,  $\tau_0 = -\infty$ ), and the highest category  $c = C$  is similarly bounded at positive infinity (i.e.,  $\tau_C = +\infty$ ). To identify the mean structure, Blimp fixes  $\tau_1$  to 0. Thresholds are estimated following the procedure in Cowles (1996).

The factorization shown in the bottom line of Equation 9 also applies to the models in Figure 4a. The bivariate normal distribution in the rightmost term now reflects the association between the latent factor and the latent response variable for  $M$ . The indicator distributions all describe the underlying latent response variables, with the residual correlation in the bivariate distribution expressed on the latent-response metric. The Blimp syntax excerpt below shows the model specification. Listing the categorical variables on the ORDINAL command line invokes probit regression. To simplify specification, the software automatically adds the threshold parameters and introduces the necessary identification constraints. All other aspects of the script are the same as before.

```
# model 4 excerpt
ORDINAL: m x1:x4;
MODEL:
eta ~~ eta@1;
eta ~~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
y ~ eta m;
```

It is important to reiterate that distributional assumptions enter via the submodels. Different types of continuous and categorical variables can coexist within the same model because the diverse set of univariate and bivariate submodels are guaranteed to reproduce *some* multivariate distribution. The advantage of factored specifications is that we do not need to specify or assume anything about the form of the joint distribution of the analysis variables.

In Figure 4a, the binary covariate correlates with  $\eta$  via its latent response scores, but the dummy variable  $M$  acts as the predictor in the structural regression. Software packages like *blavaan* and *Mplus* cannot fit this exact model, although they can estimate a model where the underlying latent response variable serves as the predictor. Figure 4b shows this alternative model, which Blimp also supports. The structural regression equation below corresponds to a modified path diagram with a directed arrow from  $M^*$  to  $Y$ .

$$Y_i = \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(M_i^*) + \varepsilon_{Y,i} \quad (13)$$

Because  $M^*$  is standardized,  $\beta_2$  reflects the expected difference in the outcome between two people with the same  $\eta$  value, but who differ by one standard deviation on the latent response variable. Adding `.latent` to  $M$ 's variable name in the structural regression equation invokes this specification (e.g., `y ~ eta m.latent;`).

The final application of probit regression involves multicategorical exogenous variables. Blimp uses a multinomial model that assigns a latent response variable to each discrete category (Aitchison & Bennett, 1970). These so-called indicants or utilities represent a continuous propensity for category membership. Following the logic of dummy coding, the full set of utilities is replaced by one fewer latent difference scores. In the interest of space, we provide a high-level discussion here and point readers to the online materials for a description of the multinomial probit model (<https://github.com/blimp-stats/fsem>). To illustrate the multinomial model, consider the model in Figure 5. This is a modified version of Figure 2, where the factor indicators are ordinal rating scales, and  $M_1$  and  $M_2$  are dummy codes representing three nominal categories, coded 0, 1, and 2. The factorization in the bottom line of Equation 6 also applies to

this model. The term  $f(M_1, M_2)$  now references the multivariate submodel for the incomplete dummy codes (or equivalently, the multicategorical variable,  $M$ ). The multinomial probit model includes two latent difference scores that contrast an individual's propensities for belonging to groups  $M_1 = 1$  and  $M_2 = 1$  relative to the reference group with  $M_1 = 0$  and  $M_2 = 0$ .

The Blimp syntax excerpt below specifies the model in Figure 5. To simplify specification, the software creates a set of dummy codes for binary and multicategorical variables listed on the NOMINAL command line, defining the lowest numeric code as the reference group. Listing the categorical variable as a predictor in the MODEL section introduces these dummy codes into the regression equation.

```
# model 6 excerpt
NOMINAL: m;
ORDINAL: x1:x4;
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta ~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
y ~ eta m;
```

The dummy codes can also be manually entered into an equation by appending their corresponding numeric codes to the predictor's name (e.g.,  $\eta \sim m.1 \ m.2$ ). To facilitate model-wide missing data handling, the software automatically constructs the multinomial probit model without user input, and it iteratively updates dummy codes to reflect the most current imputations. To our knowledge, other SEM programs do not currently support incomplete nominal predictors, although the MCMC routines in *Mplus* and *blavaan* can fit this model if the dummy codes are complete.

## Logistic Regression

Blimp also provides logistic regression models for binary and multicategorical outcomes. MCMC estimation uses a data augmentation strategy like that of the probit model, but the composition of the latent response scores is more complex (Asparouhov & Muthén, 2021; Polson et al., 2013; Windle et al., 2014), and they do not have an intuitive interpretation like they do in the probit model. The online materials provide additional technical details. To illustrate binary logistic regression, consider a modified version of the model in Figure 5 with a dichotomous outcome coded  $Y = 0$  or  $Y = 1$ . The logistic regression equation below shows the predicted log-odds as linear function of the regressors.

$$\ln\left(\frac{P(Y_i = 1)}{P(Y_i = 0)}\right) = \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(M_{1,i}) + \beta_{3Y}(M_{2,i}) \quad (14)$$

$$f(Y|\eta, M) \rightarrow \text{Bernoulli}(P(Y_i = 1))$$

The factorization in the bottom line of Equation 6 also applies here, but the model-implied conditional distribution of  $Y$  simply changes from a normal curve to a Bernoulli distribution.

The Blimp syntax excerpt below specifies the logistic structural model. Listing an outcome variable on the NOMINAL command line triggers logistic regression. All other aspects of the script remain unchanged.

```
# model 7 excerpt
NOMINAL: m y;
ORDINAL: x1:x4;
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta ~~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
```

$$y \sim \text{eta } m;$$

To reiterate, the underlying factored regression specification, which now integrates logit, probit, and normal submodels, is entirely hidden from view.

Next, consider a multicategorical outcome variable with  $K$  groups, numbered  $k = 1$  to  $K$ . The first equation below shows the predicted log-odds of group  $k$  (for  $k > 1$ ) versus the reference group ( $k = 1$ ) as a linear function of the regressors.

$$\ln\left(\frac{P(Y_i = k | k \in 2, \dots, K)}{P(Y_i = 1)}\right) = \beta_{0Y_k} + \beta_{1Y_k}(\eta_i) + \beta_{2Y_k}(M_{1,i}) + \beta_{3Y_k}(M_{2,i}) \quad (15)$$

$$f(Y|\eta, M) \rightarrow \text{Multinomial}(P(Y_i = 1), P(Y_i = 2), \dots, P(Y_i = K))$$

The factorization in the bottom line of Equation 6 also applies to this model. In this case, the model-implied conditional distribution of  $Y$  simply changes from a normal curve to a multinomial distribution. The previous Blimp script remains unchanged; when more than two outcome categories are detected, the software automatically applies multinomial logistic regression, treating the lowest numeric code as the reference group.

### Negative Binomial Regression

Blimp also offers negative binomial regression models for count outcomes—dependent variables that take nonnegative integer values representing the number of times an event occurs (e.g., the number of heavy drinking days per month or aggressive acts during a play period). Researchers frequently use Poisson regression for this type of data. The Poisson distribution is defined by a single parameter,  $\mu$ , which represents both the mean and variance of the counts. In practice, however, count data often display greater variability than the Poisson model allows (a phenomenon known as overdispersion). Overdispersion can arise from factors such as model misspecification or dependent events within individuals (Coxe et al., 2009). Negative binomial regression provides a more flexible alternative by adding a dispersion parameter that captures

heterogeneity among individuals with the same predicted count. When the dispersion parameter equals 0, the negative binomial model reduces to the Poisson model.

To illustrate negative binomial regression, consider a modified version of the model in Figure 5 where  $Y$  is a count outcome. The negative binomial regression equation below shows the logarithm of the predicted count ( $\mu_i$ ) as linear function of the regressors.

$$\ln(\mu_i) = \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(M_{1,i}) + \beta_{3Y}(M_{2,i}) = \bar{Y}_i^* \quad (16)$$

$$f(Y|\eta, M) \rightarrow \text{NB}(\bar{Y}_i^*, \phi)$$

The factorization in the bottom line of Equation 6 also applies here, but the model-implied conditional distribution of  $Y$  changes from a normal curve to a negative binomial distribution with the predicted log-count and dispersion as its parameters ( $\bar{Y}_i^*$  and  $\phi$ , respectively). Blimp once again uses data augmentation with an auxiliary latent variable to simplify estimation of the regression coefficients (Asparouhov & Muthén, 2021; Neelon, 2019; Pillow & Scott, 2012).

The Blimp syntax excerpt below specifies a structural model with a count outcome. Listing an outcome variable on the COUNT command line triggers negative binomial regression. All other aspects of the script remain unchanged.

```
# model 8 excerpt
NOMINAL: m;
ORDINAL: x1:x4;
COUNT: y;
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta ~~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
y ~ eta m;
```

As mentioned previously, the only current limitation with negative binomial regression is that incomplete count variables cannot influence downstream variables. Other generalized linear models do not share this limitation.

### **Nonnormal Continuous Variables**

Blimp offers three strategies for modeling nonnormal continuous data: normalizing transformations, heterogeneous variation models, and two-part models for floor and ceiling effects. Because it is novel among competing software programs, we focus on the Yeo–Johnson normalizing transformation in this section and refer readers to the online materials for additional information on heterogeneous variation and two-part models (<https://github.com/blimp-stats/fsem>).

#### **Yeo–Johnson Normalizing Transformation**

Researchers routinely use transformations to address nonnormal outcomes in univariate regression models. This strategy can be challenging to implement because the data’s shape dictates the optimal transformation. Missing data further complicate this issue because systematic missingness can alter the observed-data distribution. Blimp implements the Yeo–Johnson transformation (Yeo & Johnson, 2000), which attempts to normalize a variable using an iteratively estimated shape parameter. The Yeo–Johnson method is a generalized version of the Box–Cox transformation (Box & Cox, 1964) that accommodates both positive and negative values as well as positively and negatively skewed distributions. Blimp can apply the Yeo–Johnson transformation to univariate and multivariate submodels involving manifest or latent variables. The application of Yeo–Johnson to nonnormal latent variables is analogous to procedures for modeling nonnormal trait scores in the item response theory framework (Reise et al., 2018). We focus on nonnormal manifest variables in this paper and refer interested readers to Keller (2024a) for information about nonnormal latent variables. Additional discussion and empirical evaluations of the Yeo–Johnson transformation can be found in the broader statistical literature (Atkinson et al., 2020; Keller & Enders, 2023; Kim, 2024; Lüdtke et al., 2020a, 2020b; Riani et al., 2023; Woller & Enders, 2024).

To illustrate the Yeo–Johnson procedure, consider a variable  $X$  that is continuous and nonnormal. The Yeo–Johnson transformation is analogous to a generalized linear model in the sense that a function links the nonnormal variable to a normalized version of itself. For notational consistency, we use  $X^*$  to represent the normalized variable, even though it is a deterministic function of  $X$ . The transformation function has four regimes that depend on a shape parameter  $\kappa$  and whether  $X$  is positive or negative.

$$X_i^* = \begin{cases} ((X_i + 1)^\kappa - 1)/\kappa & \text{if } X_i \geq 0 \text{ and } \kappa \neq 0 \\ \ln(X_i + 1) & \text{if } X_i \geq 0 \text{ and } \kappa = 0 \\ -((-X_i + 1)^{2-\kappa} - 1)/(2 - \kappa) & \text{if } X_i < 0 \text{ and } \kappa \neq 2 \\ -\ln(-X_i + 1) & \text{if } X_i < 0 \text{ and } \kappa = 2 \end{cases} \quad (17)$$

The transformation modifies a variable's distribution by raising a shifted version of it to an exponent controlled by  $\kappa$ , an iteratively-estimated shape parameter that governs how strongly the variable is stretched or compressed to reduce skewness and make the distribution more symmetric. When  $\kappa = 1$ , there is no change. When  $\kappa < 1$ , the transformation compresses high values and stretches low values, counteracting right skewness and a long right tail. When  $\kappa > 1$ , it does the opposite, stretching high values and compressing low values to counteract left skewness. The logarithmic cases at  $\kappa = 0$  and  $\kappa = 2$  maintain continuity where the power function would otherwise be undefined (the denominator of the transformation would equal 0). The transformed variable  $X^*$  preserves the sign and rank ordering of the original scores while reshaping their distribution toward normality.

To apply the Yeo–Johnson procedure within an SEM, reconsider the model in Figure 4, this time assuming that the indicators are continuous and nonnormal. In this context, the ellipses represent the normalized variables, and the rectangles denote their nonnormal manifest counterparts. The broken arrows indicate the function linking the raw and normalized scores (the inverse of Equation 17). The path diagram highlights that the latent factor influences the

normalized versions of the indicators. The corresponding measurement model equations are as follows:

$$X_{p,i}^* = \beta_{0X_p^*} + \beta_{1X_p^*}(\eta_i) + \varepsilon_{X_p^*,i} = \bar{X}_{p,i}^* + \varepsilon_{X_p^*,i}$$

$$f(X_p|\eta) \rightarrow \mathcal{N}_{YJ}(\bar{X}_{pi}^*, \sigma_{\varepsilon_{X_p^*}}^2, \kappa_p)$$
(18)

The factorization from the bottom line of Equation 9 also applies here, but the model-implied conditional distribution of each  $X_p$  is now the Yeo–Johnson normal distribution function (Lüdtke et al., 2020b) in the bottom line of the previous equation.

The Blimp syntax excerpt below applies the Yeo–Johnson transformation to the four factor indicators. The transformation is applied by listing the indicator’s name inside the `yjt()` function. All other aspects of the model specification are the same.

```
# model 9 excerpt
MODEL:
eta ~~ eta@1;
eta ~~ m;
eta -> yjt(x1)@1o1 yjt(x2:x4);
x3 ~~ x4;
y ~ eta m;
```

To reiterate, the measurement models reflect associations between the latent variable and normalized indicators, and the residual correlation between  $X_3$  and  $X_4$  is likewise defined on the normalized metric.

### Heterogeneous Variation

Nonnormal data can arise for many reasons. One common source is heterogeneous residual variation that occurs when the spread of a variable differs systematically across individuals or groups. Even if residuals are conditionally normal, combining observations with different

variances produces an overall distribution that departs from normality, often appearing skewed or heavy-tailed. Rather than transforming variables, an alternative strategy is to model this heterogeneity directly by allowing the residual variance to depend on covariates. This idea has a long history in the statistical literature on variance modeling (Cepeda & Gamerman, 2000; Goldstein, 2005; Harvey, 1976) and has recently gained attention in the SEM framework as a tool for capturing measurement or structural heterogeneity, most notably in moderated nonlinear factor analysis (MNLFA) applications (Bauer, 2017, 2023). The online materials illustrate an example in which the factor variance from Figure 5 differs as a function of a covariate, and Enders et al. (2024) provide a more detailed application of this procedure in the context of MNLFA. We return to variance modeling later in the paper in the section on multilevel SEM.

### **Two-Part Models**

Two-part models are a strategy for nonnormal data resulting from floor or ceiling effects (often called semicontinuous data). For example, a factor indicator could have a high proportion of scores at its minimum value, perhaps because it measures an extreme aspect of the construct that isn't relevant to all respondents. As its name implies, a two-part model represents the data with a dichotomous variable coding whether an observation is above the floor and a continuous part that is observed only when the value exceeds the floor. Blimp generally supports numerous variants of two-part models, both for single-level and multilevel regression models (Blozis et al., 2020; Muthén et al., 2024; Neelon & O'Malley, 2019; Olsen & Schafer, 2001; Smith et al., 2017). In the interest of space, we refer interested readers to the online materials for information about two-part modeling and the corresponding Blimp scripts (<https://github.com/blimp-stats/fsem>).

### **Interactions And Nonlinear Effects**

Factored regression specifications for incomplete interaction and nonlinear effects are an important development that has received considerable attention in the missing data literature (Bartlett & Morris, 2015; Enders et al., 2020a; Erler et al., 2019b; Erler et al., 2016; Keller & Enders, 2023; Kim et al., 2018; Kim et al., 2015; Lüdtke et al., 2020b; Zhang & Wang, 2017). Because our FSEM framework treats latent variables as missing data to be imputed, it uses the same estimation machinery to model higher-order latent effects. Parallel developments in the likelihood framework (Boker et al., 2023; Kelava & Brandt, 2023; Klein & Moosbrugger, 2000;

Klein & Muthén, 2007) address latent variable interactions. When the interacting variables are continuous, these likelihood-based approaches yield results similar to those from factored specifications (Keller, in press). However, they are generally limited to normally distributed variables, whereas factorization accommodates a broader range of variable types. Keller (in press) provides an in-depth explanation of Blimp’s latent variable interaction framework, including latent-by-latent, manifest-by-latent, three-way latent variable interactions, and regression diagnostics with latent variables. Additionally, Keller (2022) discusses latent variable interactions in the context of mediation models, and Enders et al. (2024) apply the methodology to moderated nonlinear factor analysis.

Building on previous examples, we use a scenario where a multicategorical nominal variable moderates the influence of a latent factor. Figure 6 shows a path diagram of the model, with  $M_1$  and  $M_2$  representing dummy codes that contrast the groups with the highest numeric codes against the reference group (the lowest code). The diagram illustrates that there is no interaction latent variable with its own product indicators. Rather, the dummy codes and latent factor combine to produce the interaction effect (denoted as black diamonds). In line with standard procedures from linear regression (Aiken & West, 1991; Cohen et al., 2002), the structural model includes products of the dummy codes and the latent variable.

$$Y_i = \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(M_{1,i}) + \beta_{3Y}(M_{2,i}) + \beta_{4Y}(\eta_i)(M_{1,i}) + \beta_{5Y}(\eta_i)(M_{2,i}) + \varepsilon_{Y,i} = \bar{Y}_i + \varepsilon_{Y,i}$$

$$f(Y|\eta, M) \rightarrow \mathcal{N}(\bar{Y}_i, \sigma_{\varepsilon_Y}^2) \quad (19)$$

The structural model is fundamentally no different from any other regression with a multicategorical moderator;  $\eta$  is simply a continuous predictor with a 100% missing data rate. The lower-order  $\beta_1$  term represents the influence of the latent variable in the reference group, while  $\beta_4$  and  $\beta_5$  are slope differences for the  $M_1 = 1$  and  $M_2 = 1$  groups, respectively. Importantly, the product terms are not unique variables with their own distributional properties. Rather, imputations for  $\eta$  and  $M$  are generated from a composite distribution that incorporates parameters from multiple submodels (see Equation 10). The presence of an interaction term in

the factorization produces nonnormal latent variable imputations, with  $\eta$  values drawn from a mixture of heteroscedastic normal distributions whose conditional variance depends on the simple slope and the person's dummy codes.

When modeling latent predictors, it is important to preserve their associations with other covariates, as this ensures that the structural model's partial slopes are accurate. Blimp currently does not support multivariate distributions with multicategorical variables (see Table 1). The dummy codes and latent variable must now be linked with regressions rather than correlations. The dummy codes can predict the latent variable in a linear regression, or the latent variable can predict the nominal outcome in a multinomial logistic model. Figure 6 uses the former approach, which is consistent with the factorization in the bottom line of Equation 6. The Blimp syntax excerpt below shows how to specify the model.

```
# model 11 excerpt
ORDINAL: x1:x4;
NOMINAL: m;
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta ~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
y ~ eta m eta*m;
SIMPLE: eta | m;
```

Recall that Blimp automatically creates a set of dummy codes for binary and multicategorical variables listed on the NOMINAL command line. Listing the categorical variable as a predictor in the MODEL section introduces the dummy codes into the regression equation. Joining the latent variable and moderator with an asterisk similarly enters the product terms in Equation 20. Finally, the SIMPLE command computes and evaluates the conditional effect of the latent variable within each group (simple slopes). The `rb1imp` package also has built-in functionality for

plotting these conditional effects and determining Johnson–Neyman significance regions. To our knowledge, no other SEM software can fit this model.

The factorization in Equation 6 accommodates a variety of other nonlinear effects involving the latent variable. For instance, suppose that  $\eta$  also exerts a curvilinear influence on the outcome. The structural model below includes group-specific quadratic effects. The conditional distribution of the outcome is unchanged.

$$\begin{aligned}
 Y_i = & \beta_{0Y} + \beta_{1Y}(\eta_i) + \beta_{2Y}(\eta_i^2) + \beta_{3Y}(M_{1,i}) + \beta_{4Y}(M_{2,i}) + \\
 & \beta_{5Y}(\eta_i)(M_{1,i}) + \beta_{6Y}(\eta_i)(M_{2,i}) + \beta_{7Y}(\eta_i^2)(M_{1,i}) + \beta_{8Y}(\eta_i^2)(M_{2,i}) + \varepsilon_{Y,i} = \bar{Y}_i + \varepsilon_{Y,i} \\
 f(Y|\eta, M) \rightarrow & \mathcal{N}(\bar{Y}_i, \sigma_{\varepsilon_Y}^2)
 \end{aligned} \tag{20}$$

The corresponding Blimp syntax excerpt is shown below.

```

# model 12 excerpt
ORDINAL: x1:x4;
NOMINAL: m;
LATENT: eta;
MODEL:
eta ~~ eta@1;
eta ~ m;
eta -> x1@1o1 x2:x4;
x3 ~~ x4;
y ~ eta eta^2 m eta*m (eta^2)*m;

```

### Residual Structural Equation Models

Residual structural equation modeling (RSEM; Asparouhov & Muthén, 2023) builds on the established idea of representing residuals as latent variables by adding directed paths among the latent residuals to capture structured dependencies beyond the primary model’s covariance

structure. Compelling applications include latent growth curve models with an autoregressive error structure (Curran et al., 2014) and random intercept cross-lagged panel models (RI-CLPMs; Hamaker et al., 2015; Mulder & Hamaker, 2021), among others. Blimp can accommodate certain residualized associations, but it does so without residual latent variables. We use a latent curve model to illustrate the approach and refer readers to the online materials for details and a script for the RI-CLPM (<https://github.com/blimp-stats/fsem>).

Continuing with the variables from previous examples, now assume the  $X$ s represent repeated measurements (indexed  $p = 1, 2, 3$ , and 4),  $M$  is a binary dummy code, and  $Y$  is a distal outcome. Figure 7a shows a path diagram of the linear latent curve model in the RSEM framework. Each repeated indicator loads on the intercept and slope latent variables,  $\alpha_i$  and  $\eta_i$ , which represent each person's initial status and linear rate of change. The residuals of the indicators are also represented as latent variables (shown as ovals). Directed arrows among these residuals define an autoregressive structure in which each residual depends on the one preceding it. This secondary residual model captures temporal dependencies that remain after accounting for the primary growth process.

One of the unique features of Blimp's factorization is that a variable can exist in two states. For example, a binary predictor functions as a latent response variable in the exogenous variable submodel but as a dummy code in the structural model (see Figure 4a). In essence, variables with both incoming and outgoing arrows can occupy different metrics depending on the direction of the arrow. Residual associations can be incorporated by exploiting this property. In the latent curve example, the repeated measurements have incoming arrows from the latent variables in which the raw  $X_p$  scores serve as the dependent variables. The first three indicators also have an outgoing arrow to the next indicator in which  $X_p$  is centered at its predicted value from the growth curve. This centering defines the outgoing arrow as the effect of  $X_{p-1}$ 's residual on  $X_p$ . The measurement model with the residualized predictor (occasions  $p > 1$ ) is shown below:

$$X_{p,i} = \alpha_i + t_p(\eta_i) + \beta_{X_p}(X_{p-1,i} - \bar{X}_{p-1,i}) + \varepsilon_{X_p,i} \quad (21)$$

$$\bar{X}_{p-1,i} = \alpha_i + t_{p-1}(\eta_i)$$

In Figure 7b, we represent the residualized predictor as a deterministic manifest variable that does not have its own distribution (i.e., it is not enclosed in a shaded rectangle). The primary distinction between this approach and RSEM lies in the definition of the standardized estimates: RSEM standardizes the residual pathways using the residual latent variable's variance, whereas Blimp standardizes effects using each  $X_p$ 's total variation.

The Blimp syntax excerpt below specifies the latent curve model in Figure 7b.

```
# model 13 excerpt
ORDINAL: m;
LATENT: alpha eta;
MODEL:
  intercept m -> alpha eta;
  alpha ~~ eta;
  x1 ~ intercept@alpha eta@0;
  x2 ~ intercept@alpha eta@1 (x1 - alpha - 0*eta)@rho;
  x3 ~ intercept@alpha eta@2 (x2 - alpha - 1*eta)@rho;
  x4 ~ intercept@alpha eta@3 (x3 - alpha - 2*eta)@rho;
```

Blimp's scripting language is unique because it allows functions to serve as predictors on the right side of regression equations (i.e., equations-within-equations). Among other things, these functions can define deviation scores, product terms, squares, transformations, and sum scores. In the bottom three lines, a residualized predictor is specified by centering the previous indicator at its predicted value on the right side of the equation. Finally, attaching the same label to all three residuals (i.e., @rho) imposes an equality constraint on the autoregressive slope. This constraint can be relaxed, degrees of freedom permitting.

### Sum Scores

In recent years, sum scores have experienced a revival within the methodology literature, as researchers continue to debate their merits and deficiencies relative to latent variable models and factor scores (Liu & Pek, 2024; McNeish, 2024; McNeish & Wolf, 2020; Mislevy, 2024; Rhemtulla & Savalei, 2024; Sijtsma et al., 2024; Widaman & Revelle, 2022, 2024). The previous example highlighted that Blimp’s scripting language allows functions to serve as predictors on the right side of regression equations (i.e., equations-within-equations). Among other things, these inline functions can define deviation scores, product terms, squares, transformations, and sum scores. The primary advantage of this sum score functionality is missing data handling; MCMC incorporates item-level imputation models, and the resulting sum scores update at every iteration to reflect the most recent imputations.

To illustrate, consider a linear regression model with dummy code and sum score predictors. The structural regression model is shown below, with the sum of four ordinal indicators in parentheses.

$$Y_i = \beta_{0Y} + \beta_{1Y}(X_{1,i} + X_{2,i} + X_{3,i} + X_{4,i}) + \beta_{2Y}(M_i) + \varepsilon_{Y,i} = \bar{Y}_i + \varepsilon_{Y,i} \quad (22)$$

$$f(Y|X_1, X_2, X_3, X_4, M) \rightarrow \mathcal{N}(\bar{Y}_i, \sigma_{\varepsilon_Y}^2)$$

The Blimp syntax excerpt below shows how to specify a sum score predictor directly within a regression equation.

```
# model 14 excerpt
ORDINAL: x1:x4;
NOMINAL: m;
MODEL:
y ~ (x1 + x2 + x3 + x4) m;
```

This excerpt leverages the software’s default joint distribution for manifest exogenous variables, represented as  $f(X_1, X_2, X_3, X_4, M)$  in our notation scheme. Because all variables are categorical,

this multivariate normal distribution represents associations among the underlying latent response scores (i.e., the  $X_p^*$  and  $M^*$  variables). This saturated item-level missing data model maximizes power, but it may be challenging to estimate in high-dimensional settings with many incomplete items. Alacam et al. (2023) describe some straightforward strategies for simplifying these background models that work well in small samples. Finally, the inline sum score specification also supports interaction effects, as shown in the excerpt below, in which a sum score interacts with a categorical moderator.

```
# model 15 excerpt
ORDINAL: x1:x4;
NOMINAL: m;
MODEL:
y ~ (x1 + x2 + x3 + x4) m (x1 + x2 + x3 + x4)*m;
```

To reiterate, the primary advantage of this sum-score functionality is that it handles item-level missing data using an appropriate imputation model for categorical variables.

### Multilevel SEMs

Factored regression specifications for multilevel models have received considerable attention in the missing data literature (Enders et al., 2020b; Erler et al., 2019a; Erler et al., 2016; Goldstein et al., 2014; Grund et al., 2018; Grund et al., 2021; Keller & Enders, 2023). Leveraging this work, Blimp can fit multilevel structural equation models (MLSEMs) with up to three levels (four if the lowest level is a wide-format latent curve model for repeated measurements). With few exceptions, MLSEMs use the same variable and distribution types shown in Table 1. Multilevel models can include latent group means (Asparouhov & Muthén, 2019; Hamaker & Grasman, 2015; Lüdtke et al., 2008), random within-cluster variation (Hedeker et al., 2008, 2012; McNeish, 2021), and dynamic effects with lagged variables (Asparouhov et al., 2018; Hamaker et al., 2021; McNeish & Hamaker, 2020), among others. We demonstrate key features in this section and direct readers to the online materials for information about dynamic SEM with lagged predictors (<https://github.com/blimp-stats/fsem>). Kim et al. (2025) provide a tutorial about MLSEM in Blimp.

There are two ways to fit an MLSEM in Blimp. The first is consistent with a mixed model formulation that designates terms as fixed or random effects. The second way to is to define the random effects as explicit level-2 latent variables. We detail this approach because it provides greater flexibility and broadens the scope of accessible models. Conventional MLSEMs decompose lower-level outcome variables into orthogonal within- and between-cluster latent variables (Muthén & Asparouhov, 2011; Preacher et al., 2016; Preacher et al., 2010). In Blimp, there is no under-the-hood decomposition—the “full” dependent variable always serves as the outcome. Following the hierarchical linear modeling tradition, decomposition is achieved by centering predictors.

To illustrate and MLSEM, consider a two-level random slope regression model where a binary level-2 variable,  $M$ , moderates the within- and between-cluster parts of a level-1 predictor,  $X$ . The latent variable specification uses hierarchical regression equations that mimic those from Raudenbush and Bryk (2002). The level-1 regression equation includes just the within-cluster (group mean centered) predictor.

$$Y_{ij} = \alpha_j + \eta_j(X_{ij} - \mu_{X,j}) + \varepsilon_{Y,ij} = \bar{Y}_{ij} + \varepsilon_{Y,ij} \quad (23)$$

$$f(Y|X, \alpha, \eta) \rightarrow \mathcal{N}(\bar{Y}_{ij}, \sigma_{\varepsilon_Y}^2)$$

The  $\alpha_j$  and  $\eta_j$  terms are level-2 latent variables representing cluster  $j$ 's random intercept and slope, respectively. We isolate the within-cluster component of  $X$  by centering each cluster's scores at the level-2 latent group mean,  $\mu_{X,j}$ . This centering defines  $\eta_j$  as the pure within-cluster association for cluster  $j$ .

At level-2, the random intercept and slope latent variables each have their own regression equation, and they share a bivariate normal conditional distribution.

$$\alpha_j = \beta_{0\alpha} + \beta_{1\alpha}(\mu_{X,j}) + \beta_{2\alpha}(M_j) + \beta_{3\alpha}(\mu_{X,j})(M_j) + \varepsilon_{\alpha,j} = \bar{\alpha}_j + \varepsilon_{\alpha,j} \quad (24)$$

$$\eta_j = \beta_{0\eta} + \beta_{1\eta}(M_j) + \varepsilon_{\eta,j} = \bar{\eta}_j + \varepsilon_{\eta,j}$$

$$f(\alpha, \eta|M) \rightarrow \mathcal{N}\left(\begin{pmatrix} \bar{\alpha}_j \\ \bar{\eta}_j \end{pmatrix}, \begin{pmatrix} \sigma_{\varepsilon_\alpha}^2 & \sigma_{\varepsilon_\alpha, \varepsilon_\eta} \\ \sigma_{\varepsilon_\eta, \varepsilon_\alpha} & \sigma_{\varepsilon_\eta}^2 \end{pmatrix}\right)$$

Finally,  $X$  and  $M$  also require models, both for missing data handling and for estimating the latent group means. We again leverage the software's default joint distribution for manifest exogenous variables, which decomposes level-1 predictors like  $X$  into within-cluster deviations and latent group means (Enders et al., 2020b, pp. 91-92).

The multilevel model uses a three-part factorization that assigns a univariate conditional distribution to the outcome, a bivariate conditional distribution to the level-2 latent variables, and a bivariate distribution to the manifest predictors.

$$f(Y, X, M, \alpha, \eta) = f(Y|X, M, \alpha, \eta) \cdot f(\alpha, \eta|X, M) \cdot f(X, M) \quad (25)$$

The Blimp syntax excerpt below specifies the multilevel latent variable model.

```
# model 16 excerpt
NOMINAL: m;
CLUSTERID: l2id;
LATENT: l2id = alpha eta;
CENTER: groupmean = x;
MODEL:
alpha ~ intercept x.mean m x.mean*m;
eta ~ intercept m;
alpha ~~ eta;
y ~ intercept@alpha x@eta;
```

The CLUSTERID line is a new command identifies the level-2 clustering variable. The LATENT line now defines the level-2 latent variables representing the random intercepts and slopes (listing the level-2 identifier prior to the latent variable names specifies the second level). Next, the CENTER command requests latent group mean centering. The first three lines of the MODEL command specify the level-2 models from Equation 25. Listing the intercept keyword requests latent variable means, which are normally set to 0 by default<sup>3</sup>. Appending .mean to the level-1 predictor's name introduces its latent cluster means. The final line specifies the level-1 regression model from Equation 24. The @ symbol fixes the coefficients to the values of the level-2 random intercept and slope latent variables.

The Blimp syntax excerpt below shows the corresponding mixed-model specification for the same analysis. The entire model is written on a single line, with the fixed component to the left of the vertical pipe and the random component to the right. The excerpt omits the intercept from both parts because these terms are estimated automatically.

```
# model 17 excerpt
NOMINAL: m;
CLUSTERID: l2id;
CENTER: groupmean = x;
MODEL:
y ~ x x*m x.mean m x.mean*m | x;
```

In the last line, joining regressors with an asterisk creates the cross-level and between-cluster interactions, and listing the within-cluster predictor to the right of the vertical pipe introduces a random slope. Although this specification accommodates a narrower range of models, it offers two useful features: it interfaces with the SIMPLE command (for computing simple slopes) and corresponding `rblimp` graphing functions for conditional effects, and the output includes MLM effect sizes described by Rights and Sterba (2019).

---

<sup>3</sup> The exception to latent means being fixed at zero occurs in MLSEMs, where the @ symbol constrains coefficients to the values of the level-2 random intercept and slope latent variables. Although the latent means are estimated in this case, we include the intercept term for clarity.

## Heteroscedastic Within-Cluster Variation

Multilevel regression analyses typically assume that within-cluster residual variation is constant across level-2 units. Blimp offers two methods for modeling heterogeneous variation. The first, described by Kasim and Raudenbush (1998), treats cluster-specific residual variation as a random variable following a chi-square distribution. MCMC uses this distribution to generate cluster-specific residual variance terms, which then propagate to the estimation steps for both the fixed and random effects. Importantly, the Kasim and Raudenbush method does not include a submodel linking variance heterogeneity to covariates. Instead, estimating cluster-specific variation serves as a corrective mechanism, analogous to using heteroscedasticity-robust standard errors in maximum likelihood. This method is implemented by adding the `OPTIONS` command and `hev` keyword to the earlier scripts.

A location–scale model provides a second way to introduce heterogeneous within-cluster variation. These models have a long history in the multilevel literature (Hedeker et al., 2008, 2012; Raudenbush & Bryk, 2002), with more recent applications to MLSEM (Asparouhov et al., 2018; McNeish, 2021). A common use case involves intensive repeated-measures data in which individuals exhibit different levels of intraindividual variability. Following the procedure described earlier for heteroscedastic factor variation, observation-level heterogeneity at level 1 (e.g., days within persons) can vary deterministically as a function of level-1 predictors. In an MLM, the dispersion submodel also incorporates a level-2 (e.g., person-level) latent variable,  $\omega_j$ , that represents within-cluster variation on the logarithmic scale. The first line below shows the level-2 regression for this random intercept, and the second line defines the level-1 dispersion model:

$$\begin{aligned}\omega_j &= \beta_{0\omega} + \beta_{2\omega}(\mu_{X,j}) + \beta_{3\omega}(M_j) + \varepsilon_{\omega,j} \\ \ln(\sigma_{\varepsilon_{Y,ij}}^2) &= \omega_j + \beta_{1\omega}(X_{ij} - \mu_{X,j})\end{aligned}\tag{26}$$

In addition to the overall variation captured by  $\omega_j$ , the observation-level model includes deterministic “shocks” associated with  $X$  that cause temporary deviations from the typical variation regime (e.g., daily observations that depart from an individual’s usual pattern of intraindividual variability). Finally, the additional logarithmic random effect is permitted to correlate with the random intercepts and slopes in the multivariate normal distribution from Equation 25.

The Blimp syntax excerpt below illustrates a location-scale model specification. Following the procedure for single-level regression, log-variance models are introduced by listing the outcome variable’s name inside the `var()` function. The LATENT command now defines an additional level-2 random effect, which has its own regression equation on the third line of the MODEL command.

```
# model 19 excerpt
ORDINAL: m;
CLUSTERID: l2id;
LATENT: l2id = alpha eta omega;
CENTER: groupmean = x;
MODEL:
alpha ~ intercept x.mean m x.mean*m;
eta ~ intercept m;
omega ~ intercept x.mean m;
alpha eta omega ~~ alpha eta omega;
y ~ intercept@alpha x@eta;
var(y) ~ intercept@omega x;
```

The MLSEM functionality in *Mplus* implements a level-2 dispersion model (McNeish, 2021), but to our knowledge, it does not support level-1 predictors of the log-variance.

### Evaluating Model Fit

The defining feature of FSEM is that distributional assumptions are introduced through multiple submodels. The previous applications highlight that these submodels can invoke myriad

combinations of distributions and effects that are incompatible with conventional multivariate functions. Abandoning the multivariate normality assumptions changes how model fit is evaluated, because FSEM parameters do not combine to reproduce a mean vector and variance–covariance matrix as they do in linear SEMs. Consequently, global fit indices such as the chi-square, RMSEA, SRMR, and CFI are not defined. To be fair, the same limitation applies to other likelihood-based and MCMC estimators that accommodate multiple distribution types. Blimp instead provides three tools for evaluating model fit: residual correlations to identify sources of local misfit, Bayesian model comparison indices (DIC and WAIC), and MCMC-based frequentist Wald statistics.

### Correlations Among Residuals

A typical FSEM consists of several regression models. As MCMC iterates, Blimp computes each outcome variable's residuals by subtracting the current scores from their predicted values (the variables denoted with overbars throughout the paper). For binary and ordinal outcomes, predicted values and residuals are on the latent response metric; for nonnormal outcomes, they are on the transformed metric. After computing these deviations, Blimp uses the residual data matrix to estimate a saturated model that includes all possible correlations. Importantly, these residual correlations are not differences between the sample and model-predicted correlations, as in a multivariate SEM. Rather, they reflect remaining sources of linear association that persist in the latent and manifest data after fitting the model. These correlations are analogous to modification indices because they signal localized sources of misfit that manifests as linear dependence. In models involving nonlinear relationships, such as quadratic effects, residual correlations may remain close to 0 even when substantive misfit is present. Nonetheless, linear dependencies are often a salient and interpretable indicator of local misfit, making this diagnostic a useful starting point for evaluating model adequacy. Graphical diagnostics can detect these types of misspecification, however (Keller, in press).

Residual correlations are available for binary and ordinal outcomes with a probit link and any continuous outcome (normal, nonnormal, manifest, latent). To illustrate, reconsider the model in Figure 4a. MCMC would iteratively compute residuals for  $Y$ ,  $X_1^*$ ,  $X_2^*$ ,  $X_3^*$ ,  $X_4^*$ ,  $M^*$ , and  $\eta$ . Their correlations could reveal unmodeled associations among the indicators (e.g.,  $X_p^*$  with  $X_{-p}^*$ ),

omitted direct effects (e.g.,  $X_p^*$  with  $M^*$ ), or an incorrect functional form (e.g.,  $\eta$  with  $Y$  or  $X_p^*$ ). Residual diagnostics are not currently available for outcomes with logit or negative binomial links because the underlying latent response scores are nonnormal. Additionally, residuals are not available for manifest exogenous variables with autogenerated supporting models (variables that appear only to the right of a tilde). For instance, the  $M_1$  with  $M_2$  variables in Figure 2 would not have residuals by default. To obtain residuals for these variables, you must explicitly specify a multivariate submodel by adding their correlation to the MODEL section.

### Bayesian Model Comparison Indices

Blimp provides two Bayesian model comparison indices, the deviance information criterion (DIC; Spiegelhalter et al., 2002) and the Watanabe-Akaike information criterion (WAIC; Watanabe, 2010). These indices are useful for comparing nested or non-nested models. The DIC and WAIC both attempt to produce an unbiased estimate of out-of-sample prediction error, which is the error obtained when applying the fitted model parameters to new data. They do this by first calculating the predictive accuracy of the observed data and then applying a bias correction to account for the fact that the fitted model's predictions will be less accurate when applied to new, unseen data (Gelman et al., 2014). A desirable feature of the WAIC is that it uses parameters from every MCMC iteration to estimate a model's predictive accuracy, whereas the classic version of the DIC uses the average parameters. Following ideas from Celeux et al. (2006), Blimp implements a modified version of the DIC that similarly uses the entire posterior distribution. Both indices accommodate multilevel data and missing data (Du et al., 2024).

Three features of Blimp's DIC and WAIC are worth highlighting. First, like residual correlations, manifest exogenous variables with autogenerated supporting models (variables that appear only to the right of a tilde) do not contribute information. For instance, the supporting models for  $M_1$  and  $M_2$  in Figure 2 would not contribute to the DIC and WAIC. Second, Blimp generally uses a conditional formulation of the DIC and WAIC where the imputed latent data function like observed data. In contrast, a marginal approach would use the mean and variance of the latent data rather than the latent scores themselves. Third, the combined-model specification for MLSEMs offers both marginal and conditional versions of the DIC and WAIC. The former uses variances of the random effects, while the latter uses the random effect

imputations as though they are observed data. Multilevel models with a logit or negative binomial link are restricted to the conditional approach. The optimal definitions for information-based indices and the identification constraints needed to elicit valid comparisons remain an active area of research (Du et al., 2024; Graves & Merkle, 2022; Merkle et al., 2019).

### **Frequentist Significance Tests**

Blimp's output provides a regression summary table for each outcome variable that includes unstandardized coefficients, standardized weights, and variance explained effect sizes (Rights & Sterba, 2019). For each quantity, the summary table reports posterior medians, standard deviations, and credible intervals obtained from MCMC estimation. These quantities are the Bayesian analogs of frequentist point estimates, standard errors, and confidence intervals, but they summarize a distribution of plausible parameter values that could have produced the realized sample data. Of course, these summaries can be used for Bayesian inference, but they can also serve as direct substitutes for frequentist inference—a perspective that Levy and McNeish (2023) describe as “computational frequentism”.

A researcher adopting the computational frequentism perspective may do so because a frequentist estimator is either unavailable or computationally impractical. While we strongly appreciate the philosophical appeal of Bayesian inference, several of the example models are compelling use cases for computational frequentism. For example, likelihood-based estimators do not currently support interactions involving a multicategorical moderator and a latent factor (see Figure 6). Missing data analyses are another compelling application, where likelihood (and even other Bayesian approaches) substantially limit the types of incomplete variables that can be included in a model. To facilitate MCMC-based frequentist inference, Blimp's output includes univariate “Bayesian Wald tests” (Asparouhov & Muthén, 2021a) for most parameters. Custom tests of multiple parameters are also available from the WALDTEST command. Limited research to date suggests that the MCMC-based Wald test provides good frequentist properties, potentially offering more accurate Type I error rates and power estimates than likelihood-based Wald statistics (Asparouhov & Muthén, 2021b; Woller & Enders, 2024).

### Real Data Analysis Example

The real data analysis uses a moderated mediation model that showcases several of the features described in the paper, namely mixtures of linear and generalized linear regressions, interaction effects, and nonnormal variables. The analysis uses data from  $N = 99$  participants in a trial investigating the effect of early life stressors on alcohol use via systemic inflammation as a mediator (McManus et al., In preparation). Figure 8 shows a path diagram of the model with histograms depicting the score distributions and shaded rectangles enclosing submodels with unique distributional assumptions. Interested readers are referred to the original study for detailed information on data collection, study design, measures, and the physiological and behavioral processes represented in the path diagram. The online materials at include an annotated PDF describing the Blimp syntax and output for various modeling steps (<https://github.com/blimp-stats/fsem>).

To begin, the manifest exogenous variables include a binary focal predictor (early life stress; *ELS*), a binary moderator (biological sex; *SEX*), and a numeric covariate (age in years; *AGE*). As depicted in the diagram, the binary predictors appear as dummy codes in the structural model (the rectangles), but their relationships with age and each other utilize latent response variables (ellipses). The broken arrows represent inverse link functions that convert the latent responses to the manifest metric.

Next, an inflammation latent variable—indicated by three plasma biomarkers (TNF- $\alpha$  [*TNF*], IL-6 [*IL6*], and C-reactive protein [*CRP*])—served as the mediator. Two of the inflammation indicators, *CRP* and *IL6*, exhibited substantial positive skew. In the original study, these variables were log-transformed prior to forming a composite mediator. Here, we instead applied a Yeo–Johnson transformation that incorporates an iteratively-estimated shape parameter that optimizes the transformation to the observed data’s distribution. In the diagram, the ellipses represent normalized indicators and rectangles are the manifest (untransformed) variables. The broken arrow corresponds to a function that converts normalized scores to their skewed observed values (the inverse of the transformation in Equation 17).

Finally, the model featured two alcohol consumption outcomes: drinks per drinking day (*DPDD*) and the number of heavy drinking days (*HDD*). In lieu of a log transformation, we again

applied the Yeo–Johnson transformation to the positively skewed *DPDD* variable. The path diagram elements are the same as before. The number of heavy drinking days was modeled as a count outcome using a negative-binomial regression with a log link function. In the path diagram, the ellipse represents the linear predictor, and the rectangle denotes the discrete *HDD* variable. Although the drinking outcomes undoubtedly correlate beyond their common predictors, correlated residuals are unavailable for count outcomes, and we could not form a latent drinking factor with only two indicators. Because the outcomes share the same predictor set, omitting the residual association still yields consistent (i.e., asymptotically unbiased) estimates of each equation’s regression coefficients, although some precision will be lost.

McManus et al. (In preparation) hypothesized that the indirect effect of early life stress on drinking via systemic inflammation would be moderated by sex. To evaluate this prediction, the model incorporated an interaction effect where the influence of *ELS* on inflammation differed by sex. Conditional indirect effects were quantified by computing the product of coefficients estimator separately for males and females. When using count outcomes, indirect effects are further conditional on the values of the exogenous manifest variables. These conditional indirect effects were computed following procedures described in Hayes and Preacher (2010) and Geldhof et al. (2018, Table 2). Consistent with the original study, which used a composite score as a mediator, the conditional indirect effect of *ELS* on *DPDD* was significant for females but not for males, because the association between early life stress and inflammation was stronger among women. Although the count outcome provided evidence of mediation, its 95% asymmetric confidence interval just included 0, which likely reflects the reduced power of the negative-binomial regression model.

## Discussion

This paper introduced a factored regression approach to SEM, which recasts a multivariate distribution as a set of simpler submodels, each with its own corresponding distribution. At its core, Blimp’s modeling framework is a missing data methodology that treats latent variables as unobserved data to be imputed rather than integrated out of the likelihood. Latent variable imputation with equation-wise distributional assumptions provides substantial flexibility, as the realized latent variable scores naturally adapt to complex data structures that are incompatible

with standard multivariate distributions. Important use cases include analyses with discrete and numeric variables, nonnormal continuous variables, and models involving interaction and nonlinear effects.

When would a researcher adopt FSEM over a multivariate SEM? The FSEM framework offers no advantages when variables are continuous and approximately normal. In such cases, FSEM is equivalent to the multivariate SEM, and the latter is preferable based on its familiarity alone. A primary use case for FSEM in that context is to generate multiply imputed latent variable scores rather than estimated factor scores (Asparouhov & Muthén, 2010; Rhemtulla & Savalei, 2024). FSEM is ideally suited for applications involving nonnormal variables, broadly defined. Factored regression supports myriad combinations of variable types and nonlinear effects that are incompatible with standard multivariate functions. Applications include models with mixed variable types, nonnormal continuous variables, heteroscedastic variation, interactions, nonlinear effects, and dynamic effects, among others. Many combinations of these features are not available in other software programs, particularly when data are incomplete. A final consideration is that Blimp is free. The development team's philosophy is to democratize powerful MCMC modeling for the broadest possible audience.

Although imputing latent variables simplifies usually-complex estimation problems, that flexibility comes with a price, which is slightly larger mean squared errors (MSEs). Intuitively, imputing latent variable scores that are 100% missing adds noise to the system. All things being equal, the multivariate SEM will be more precise (have less sampling error) than its FSEM counterpart. However, simulation evidence and personal experience suggests that MSE differences are often quite subtle and do not translate into noticeable power differences (Keller, *in press*). This is an important avenue for future research. Additional methodological studies are needed to clarify the sample size requirements for FSEM and to determine when and how these requirements differ from those of multivariate SEM counterparts.

FSEM also introduces opportunities for evaluating and developing new fit measures and test statistics. Avoiding multivariate assumptions changes the way we evaluate model fit because FSEM parameters do not combine to predict the mean vector and variance–covariance matrix like they do in conventional SEMs that assume normality. This limitation applies to virtually any

variant of SEM that does not rely on multivariate normality, including popular likelihood-based methods. To address this challenge, we developed novel residual correlations that can be used to identify local sources of misfit. Somewhat anecdotally, we have found these correlations to be quite useful in situations where other fit measures are unavailable (e.g., factor models with ordinal indicators; Enders et al., 2024). Additional methodological studies are needed to evaluate the utility of these diagnostics and adapt new measures of fit for submodel factorizations.

In summary, FSEM offers numerous modeling opportunities for complex data structures that do not align with multivariate distributional assumptions. This encompasses a wide range of multivariate linear SEMs as well as newer models reflecting recent innovations in the literature. The Blimp software allows researchers to access the benefits of FSEM with minimalistic syntax and no deep knowledge about factored regression or Bayesian statistics.

## Appendix

The following script for the Blimp Studio graphical interface fits the structural model in Figure 1a. The # denotes the beginning of a comment.

```
DATA: example.dat;           # raw text file
VARIABLES: y x1 x2 x3 x4;    # variable names
MISSING: 999;                # missing value code
LATENT: eta;                  # define latent variable
MODEL:                        # model equations
eta -> x1@lo1 x2:x4;          # measurement model
eta ~~ eta@1;                 # identification constraint
y ~ eta;                       # structural model
BURN: 10000;                  # warm-up iterations
ITER: 10000;                  # analysis iterations
SEED: 90291;                  # random number seed
```

The code below shows the corresponding rblimp script. The variable names and missing data are encoded in the R data object.

```
mymodel <- rblimp(           # rblimp function call
data = example               # R data frame
latent = 'eta',               # define latent variable
model = '                     # model equations
  eta -> x1@lo1 x2:x4;        # measurement model
  eta ~~ eta@1;               # identification constraint
  y ~ eta;',                  # structural model
burn = 10000,                 # warm-up iterations
iter = 10000,                 # analysis iterations
seed = 90291)                 # random number seed

output(mymodel)               # print output
```

**Table 1**

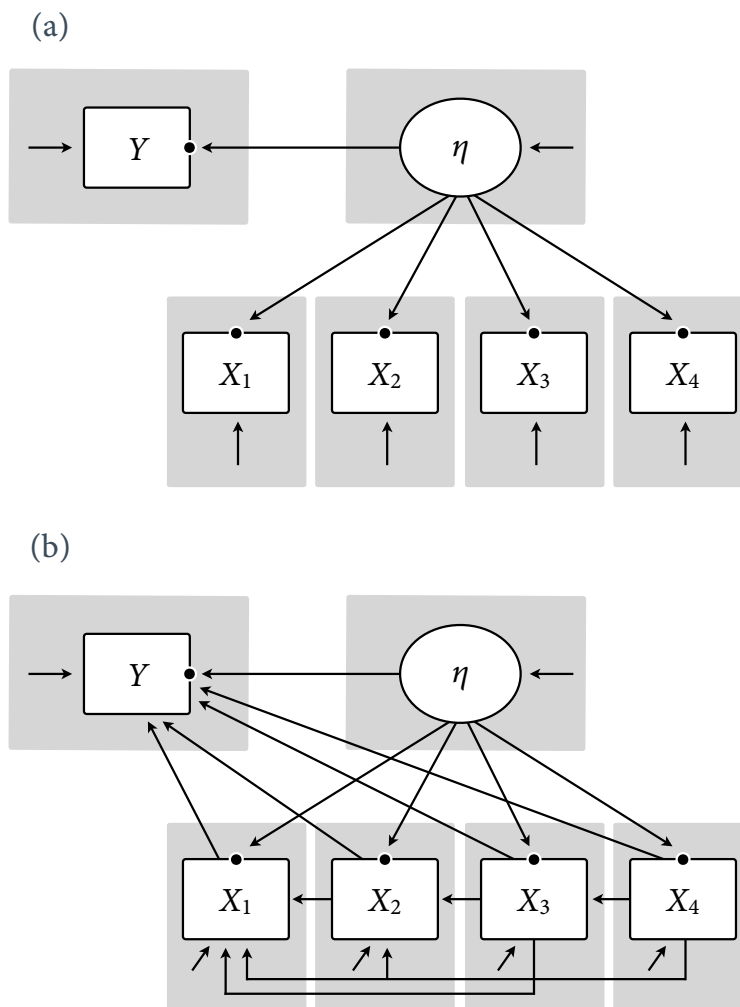
Table 1

*Variable Types and Link Functions for Three Submodels*

| Variable Type      | Submodel                 |                                       |                            |
|--------------------|--------------------------|---------------------------------------|----------------------------|
|                    | Univariate<br>(Outcomes) | Multivariate<br>(Manifest Predictors) | Multivariate<br>(Outcomes) |
| Normal manifest    | ✓                        | ✓                                     | ✓                          |
| Normal latent      | ✓                        |                                       | ✓                          |
| Nonnormal manifest | ✓ (Yeo-Johnson)          |                                       | ✓ (Yeo-Johnson)            |
| Nonnormal latent   | ✓ (Yeo-Johnson)          |                                       | ✓ (Yeo-Johnson)            |
| Binary             | ✓ (logit or probit)      | ✓ (probit)                            | ✓ (probit)                 |
| Ordinal            | ✓ (probit)               | ✓ (probit)                            | ✓ (probit)                 |
| Multicategorical   | ✓ (logit)                | ✓ (probit)                            |                            |
| Count              | ✓ (negative binomial)    |                                       |                            |
| Two-Part           | ✓ (any type)             |                                       | ✓ (any type)               |

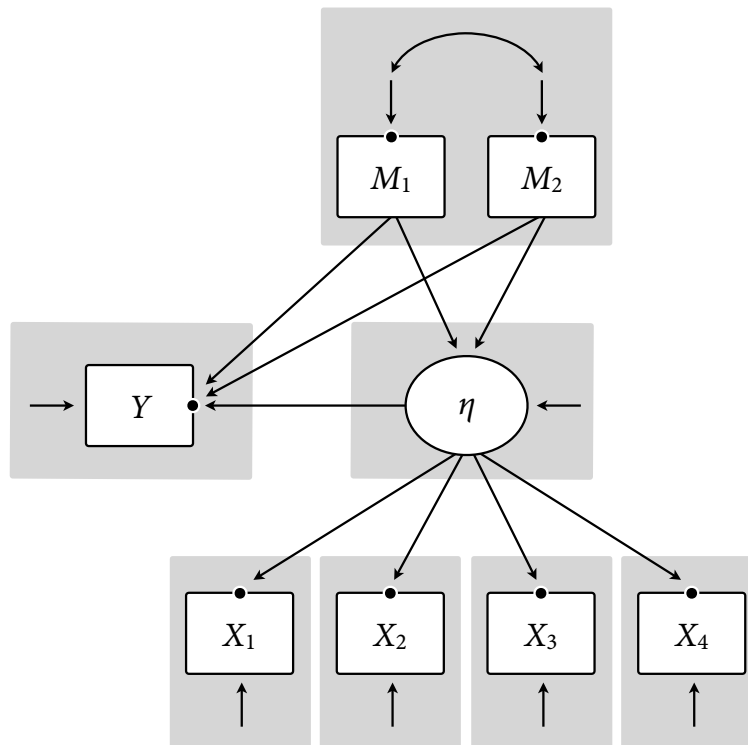
**Figure 1.**

Panel 1a is a path diagram of an FSEM where a normally distributed outcome is regressed on a latent predictor with continuous indicators. The ellipse represents the latent factor, rectangles denote normally distributed manifest variables, and black dots symbolize intercepts and means. The shaded rectangles enclose variables that share a common distribution. Panel 1b is a path model that corresponds to a saturated factorization where every variable links to all others.



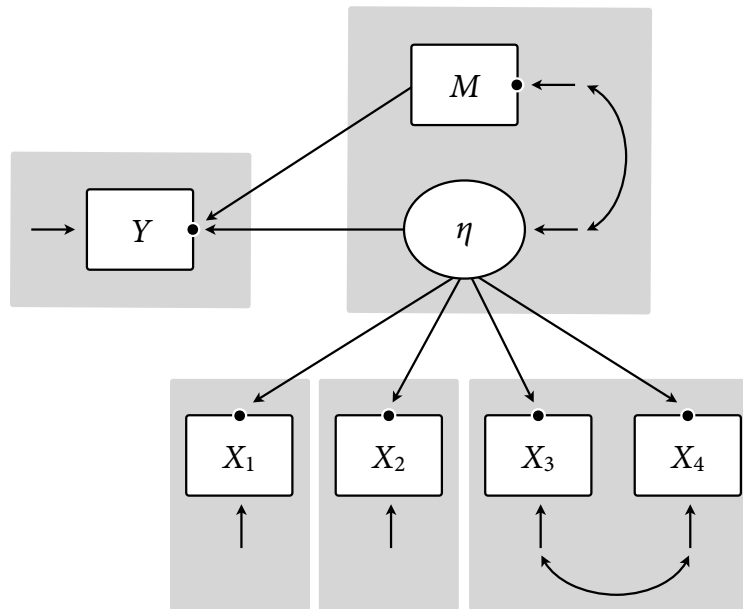
**Figure 2.**

Path diagram of an FSEM where a normally distributed outcome is regressed on a latent predictor with continuous indicators and a pair of normally distributed manifest regressors. The ellipse represents the latent factor, rectangles denote normally distributed manifest variables, and black dots symbolize intercepts and means. The shaded rectangles enclose variables that share a common distribution.



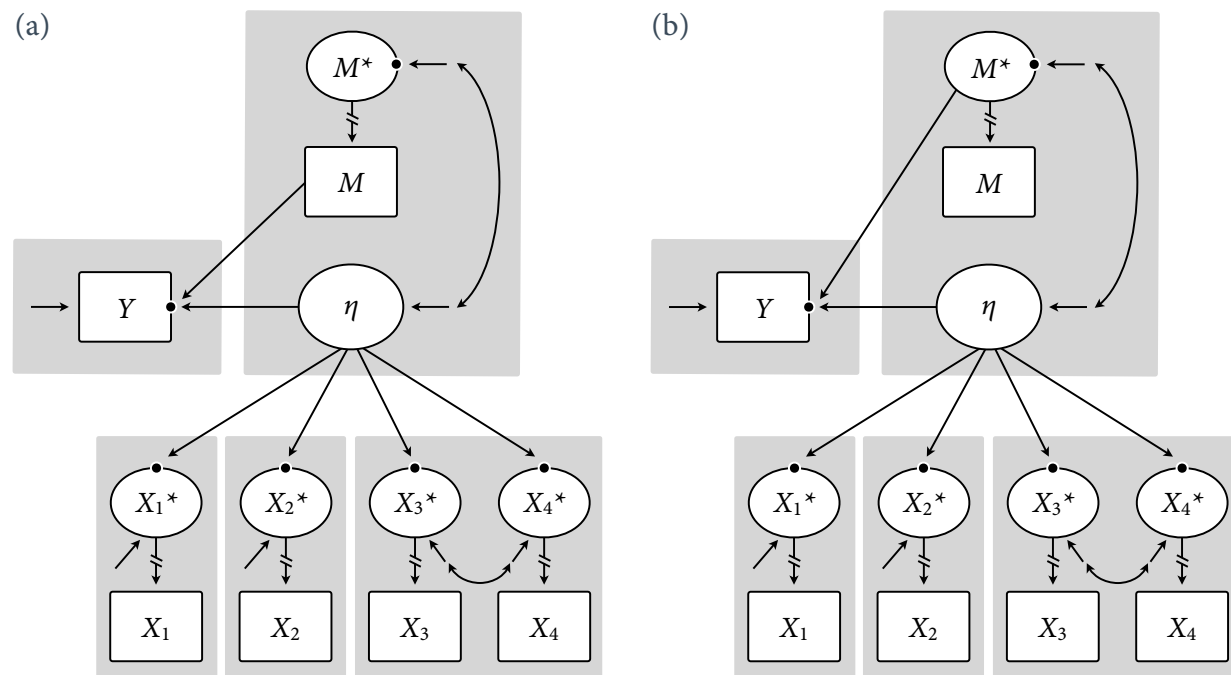
**Figure 3.**

Path diagram of an FSEM where a normally distributed outcome is regressed on normally distributed manifest and latent predictors. The ellipse represents the latent factor, rectangles denote normally distributed manifest variables, and black dots symbolize intercepts and means. The shaded rectangles enclose variables that share a common distribution.



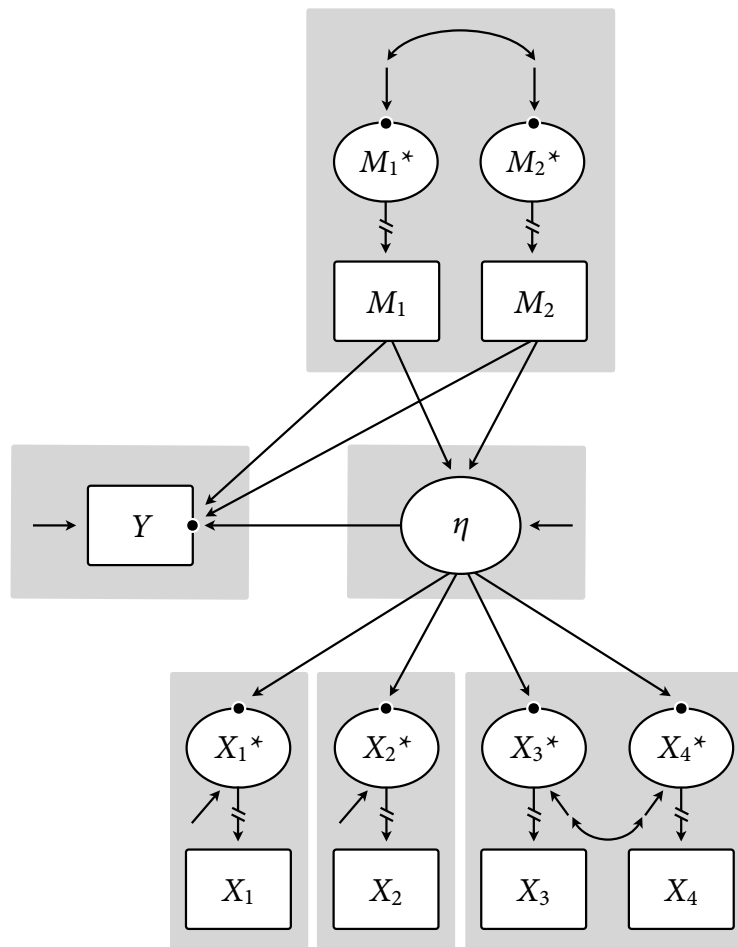
**Figure 4.**

Panel 4a is diagram of an FSEM where a normally distributed outcome is regressed on binary and latent predictors. The factor indicators are ordered categorical variables. Ellipses represent the latent factor and underlying latent response variables, rectangles denote normally distributed manifest variables, black dots symbolize intercepts and means, and broken arrows represent inverse link functions that convert the latent responses back to the discrete metric. The shaded rectangles enclose variables that share a common distribution. Panel 4b is diagram of an FSEM where a normally distributed outcome is regressed on the latent response variable and the latent factor.



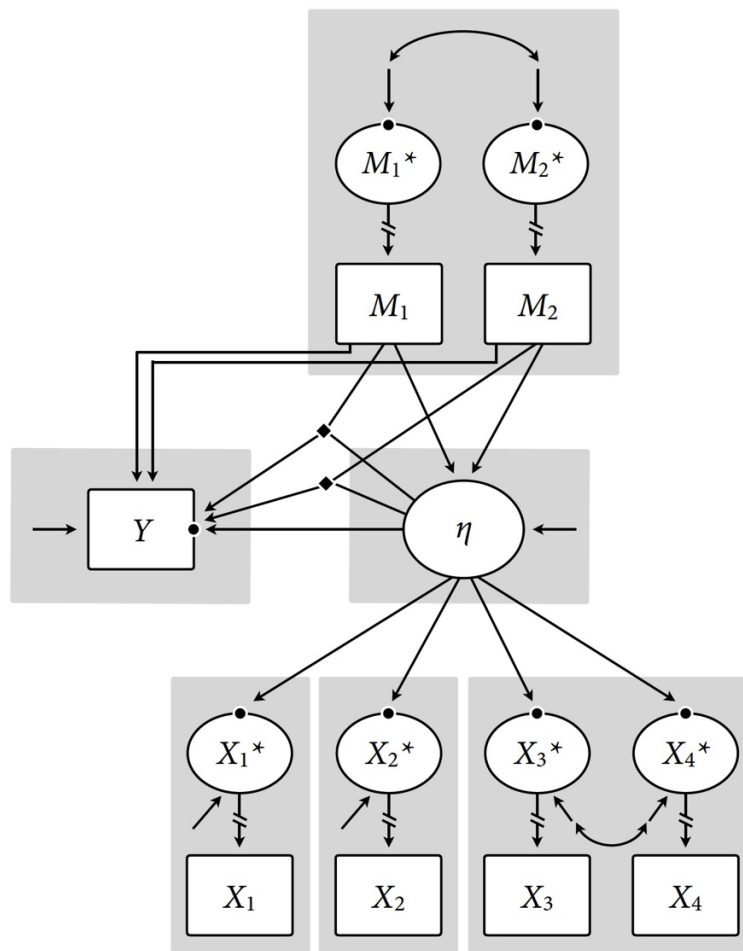
**Figure 5.**

Path diagram of an FSEM where a normally distributed outcome is regressed on a three-category nominal predictor and a latent factor. The multicategorical predictor is represented as two dummy codes. The factor indicators are ordered categorical variables. Ellipses represent the latent factor and underlying latent response variables, rectangles denote normally distributed manifest variables, black dots symbolize intercepts and means, and broken arrows represent inverse link functions that convert the latent responses back to the discrete metric. The shaded rectangles enclose variables that share a common distribution.



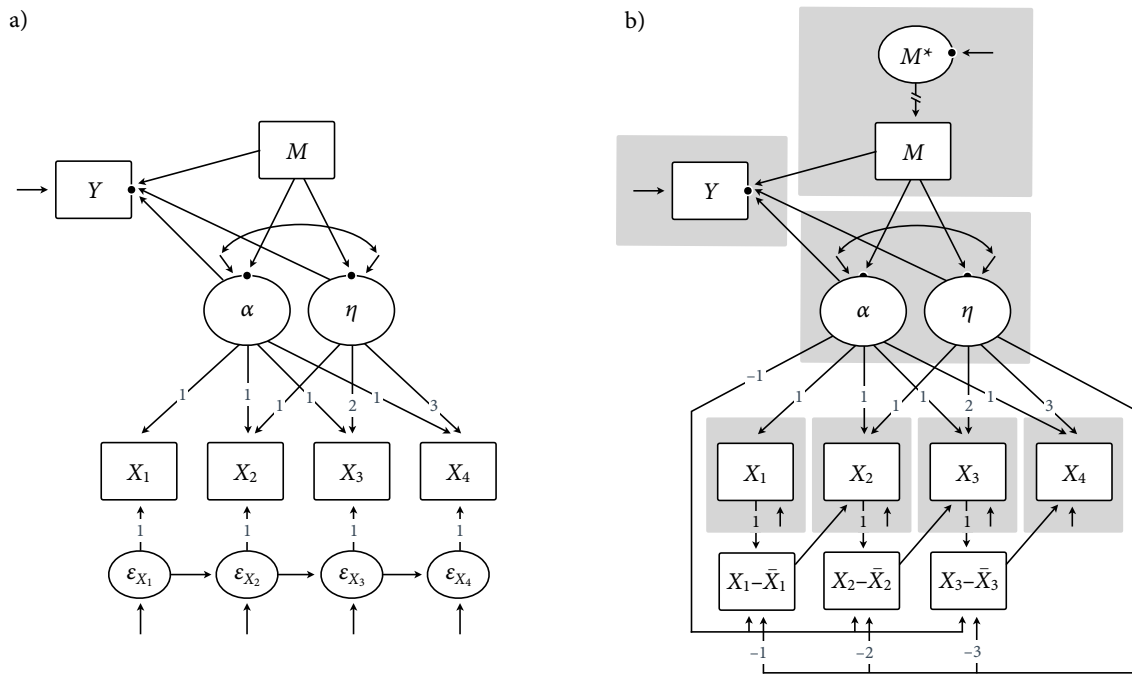
**Figure 6.**

Path diagram of an FSEM where a normally distributed outcome is regressed on a three-category nominal predictor, a latent factor, and their interaction. The factor indicators are ordered categorical variables. The multicategorical predictor is represented as two dummy codes, each of which interacts with the latent factor. Ellipses represent latent variables and underlying latent response variables, rectangles denote normally distributed manifest variables, triangles represent interaction effects, black dots symbolize intercepts and means, and broken arrows denote inverse link functions that convert the latent responses back to the discrete metric. The shaded rectangles enclose variables that share a common distribution.



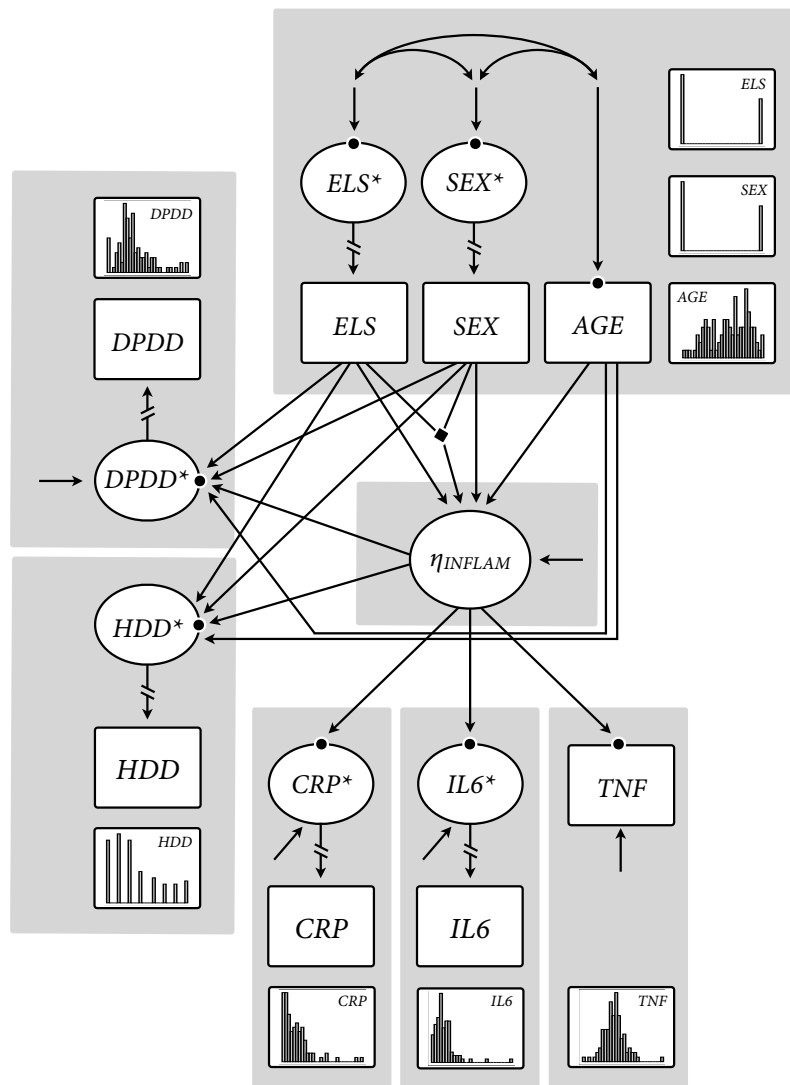
**Figure 7.**

Path diagram of an FSEM linear latent curve model with a binary predictor. Ellipses represent the latent factor and underlying latent response variables, rectangles denote normally distributed manifest variables, black dots symbolize intercepts and means, and broken arrows represent inverse link functions that convert the latent responses back to the discrete metric. The shaded rectangles enclose variables that share a common distribution. In panel a, the directed arrows connecting the residual latent variables in the measurement model capture an AR1 error structure. In panel b, these same effects are captured by centering repeated measurements at their predicted values. The figure shows the residualized predictor as a deterministic manifest variable that does not have its own distribution (i.e., it is not enclosed in a shaded rectangle).



**Figure 8.**

FSEM path diagram of the moderated mediation model from the real data analysis. Two of the predictors (ELS and gender) are binary dummy codes and one is continuous. Two of the factor indicators (CRP and IL6) and the drinks per drinking day (DPDD) outcome are skewed variables that are normalized with a Yeo–Johnson transformation. The number of heavy drinking days (HDD) outcome is a count variable with a negative binomial regression. Ellipses represent the latent factor and underlying latent response variables, rectangles denote manifest variables, black dots symbolize intercepts and means, the triangle represents an interaction effect, and broken arrows represent link functions that convert the latent responses (or normalized variables) to the manifest metric. Shaded rectangles enclose variables that share a common distribution.



## References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- Aitchison, J., & Bennett, J. A. (1970). Polychotomous quantal response by maximum indicant. *Biometrika*, 57(2), 253–262. <https://doi.org/10.2307/2334834>
- Alacam, E., Enders, C. K., Du, H., & Keller, B. T. (2023). A factored regression model for composite scores with item-level missing data. *Psychological Methods*, 30(3), 462–481. <https://doi.org/https://doi.org/10.1037/met0000584>
- Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, 88(422), 669–679. <https://doi.org/10.2307/2290350>
- Andrews, M., & Baguley, T. (2013). Prior approval: The growth of Bayesian methods in psychology. *British Journal of Mathematical and Statistical Psychology*, 66(1), 1–7. <https://doi.org/10.1111/bmsp.12004>
- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 359–388.
- Asparouhov, T., & Muthén, B. (2021). Expanding the Bayesian structural equation, multilevel and mixture models to logit, negative-binomial and nominal variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 28, 622–637. <https://doi.org/doi.org/10.1080/10705511.2021.1878896>
- Asparouhov, T., & Muthén, B. (2023). Residual structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 30(1), 1–31. <https://doi.org/doi.org/10.1080/10705511.2022.2074422>
- Asparouhov, T., & Muthén, B. (2010). Plausible values for latent variables using Mplus.
- Asparouhov, T., & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal*, 26, 119–142. <https://doi.org/10.1080/10705511.2018.1511375>
- Asparouhov, T., & Muthén, B. (2021a). Advances in Bayesian model fit evaluation for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(1), 1–14. <https://doi.org/doi.org/10.1080/10705511.2020.1764360>
- Asparouhov, T., & Muthén, B. (2021b). Advances in bayesian model fit evaluation for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 28, 1–14.

- Atkinson, A. C., Riani, M., & Corbellini, A. (2020). The analysis of transformations for profit-and-loss data. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 69(2), 251–275. <https://doi.org/doi.org/10.1111/rssc.12389>
- Bartlett, J. W., & Morris, T. P. (2015). Multiple imputation of covariates by substantive-model compatible fully conditional specification. *The Stata Journal*, 15, 437–456. <https://doi.org/10.1177/1536867X1501500206>
- Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological Methods*, 22(3), 507.
- Bauer, D. J. (2023). Enhancing measurement validity in diverse populations: Modern approaches to evaluating differential item functioning. *British Journal of Mathematical and Statistical Psychology*, Advance online publication, 1–27. <https://doi.org/10.1111/bmsp.12316>
- Blozis, S. A., McTernan, M., Harring, J. R., & Zheng, Q. (2020). Two-part mixed-effects location scale models. *Behavior Research Methods*, 52, 1836–1847. <https://doi.org/doi.org/10.3758/s13428-020-01359-7>
- Boker, S. M., von Oertzen, T., Pritikin, J. N., Hunter, M. D., Brick, T. R., Brandmaier, A. M., & Neale, M. C. (2023). Products of variables in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 30(5), 708–718. <https://doi.org/doi.org/10.1080/10705511.2022.2141749>
- Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley.
- Bollen, K. A., & Curran, P. J. (2005). *Latent curve models*. Wiley. <https://doi.org/10.1002/0471746096>
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 26(2), 211–243.
- Carpenter, J. R., Bartlett, J. W., Morris, T. P., Wood, A. M., Quartagno, M., & Kenward, M. G. (2023). *Multiple Imputation and its Application (2nd. Ed.)*. Wiley.
- Carpenter, J. R., Goldstein, H., & Kenward, M. G. (2011). REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types. *Journal of Statistical Software*, 45(5), 1–14. <https://doi.org/10.18637/jss.v045.i05>
- Celeux, G., Forbes, F., Robert, C. P., & Titterton, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, 1(4), 651–673.
- Cepeda, E., & Gamerman, D. (2000). Bayesian modeling of variance heterogeneity in normal regression models. *Brazilian Journal of Probability and Statistics*, 207–221.

- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.
- Cowles, M. K. (1996). Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statistics and Computing*, 6(2), 101–111. <https://doi.org/10.1007/Bf00162520>
- Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *J Pers Assess*, 91(2), 121–136. <https://doi.org/10.1080/00223890802634175>
- Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: a latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, 82(5), 879–894. <https://doi.org/10.1037/a0035297>
- Daniels, M. J., Wang, C., & Marcus, B. H. (2014). Fully Bayesian inference under ignorable missingness in the presence of auxiliary covariates. *Biometrics*, 70(1), 62–72. <https://doi.org/10.1111/biom.12121>
- Du, H., Enders, C. K., Keller, B. T., Bradbury, T., & Karney, B. (2021). A Bayesian latent variable selection model for nonignorable missingness. *Multivariate Behavioral Research*, Advance online publication. <https://doi.org/10.1080/00273171.2021.1874259>
- Du, H., Keller, B., Alacam, E., & Enders, C. (2024). Comparing DIC and WAIC for multilevel models with missing data. *Behavior Research Methods*, 56, 2731–2750. <https://doi.org/doi.org/10.3758/s13428-023-02231-0>
- Enders, C. K. (2022). *Applied Missing Data Analysis* (2nd ed.). Guilford Press.
- Enders, C. K., Du, H., & Keller, B. T. (2020a). A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and other nonlinear terms. *Psychological Methods*, 25(1), 88–112. <https://doi.org/10.1037/met0000228>
- Enders, C. K., Du, H., & Keller, B. T. (2020b). A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and other nonlinear terms. *Psychological Methods*, 25, 88–112. <https://doi.org/10.1037/met0000228>
- Enders, C. K., Keller, B. T., & Levy, R. (2018). A fully conditional specification approach to multilevel imputation of categorical and continuous variables. *Psychological Methods*, 23(2), 298–317. <https://doi.org/10.1037/met0000148>

- Enders, C. K., Vera, J. D., Keller, B. T., Lenartowicz, A., & Loo, S. K. (2024). Building a Simpler Moderated Nonlinear Factor Analysis Model With MCMC Estimation. *Psychological Methods, Advanced online publication*. <https://doi.org/10.1037/met0000712>
- Erler, N. S., Rizopoulos, D., Jaddoe, V. W., Franco, O. H., & Lesaffre, E. M. (2019a). Bayesian imputation of time-varying covariates in linear mixed models. *Statistical Methods in Medical Research*, 28, 555–568. <https://doi.org/10.1177/0962280217730851>
- Erler, N. S., Rizopoulos, D., Jaddoe, V. W., Franco, O. H., & Lesaffre, E. M. (2019b). Bayesian imputation of time-varying covariates in linear mixed models. *Statistical Methods in Medical Research*, 28(2), 555–568. <https://doi.org/10.1177/0962280217730851>
- Erler, N. S., Rizopoulos, D., Rosmalen, J., Jaddoe, V. W., Franco, O. H., & Lesaffre, E. M. (2016). Dealing with missing covariates in epidemiologic studies: A comparison between multiple imputation and a full Bayesian approach. *Statistics in Medicine*, 35(17), 2955–2974. <https://doi.org/10.1002/sim.6944>
- Geldhof, G. J., Anthony, K. P., Selig, J. P., & Mendez-Luck, C. A. (2018). Accommodating binary and count variables in mediation: A case for conditional indirect effects. *International Journal of Behavioral Development*, 42(2), 300–308. <https://doi.org/10.1177/0165025417727876>
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24, 997–1016. <https://doi.org/10.1007/s11222-013-9416-2>
- Goldstein, H. (2005). Heteroscedasticity and complex variation. *Encyclopedia of statistics in behavioral science*, 2, 790–795.
- Goldstein, H., Carpenter, J., Kenward, M. G., & Levin, K. A. (2009). Multilevel models with multivariate mixed response types. *Statistical Modelling*, 9(3), 173–197. <https://doi.org/10.1177/1471082x0800900301>
- Goldstein, H., Carpenter, J. R., & Browne, W. J. (2014). Fitting multilevel multivariate models with missing data in responses and covariates that may include interactions and non-linear terms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 177(2), 553–564. <https://doi.org/10.1111/rssa.12022>
- Graves, B., & Merkle, E. C. (2022). A note on identification constraints and information criteria in Bayesian latent variable models. *Behavior Research Methods*, 54(2), 795–804.

- Grund, S., Lüdtke, O., & Robitzsch, A. (2018). Multiple imputation of missing data for multilevel models: Simulations and recommendations. *Organizational Research Methods*, 21(1), 111–149. <https://doi.org/10.1177/1094428117703686>
- Grund, S., Lüdtke, O., & Robitzsch, A. (2021). Multiple imputation of missing data in multilevel models with the R package mdmb: A flexible sequential modeling approach. *Behavior Research Methods*, 53, 2631–2649. <https://doi.org/10.3758/s13428-020-01530-0>
- Hamaker, E., Asparouhov, T., & Muthén, B. (2021). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In R. H. Hoyle (Ed.), *The Handbook of Structural Equation Modeling* (pp. 576–596). Guilford Press.
- Hamaker, E. L., & Grasman, R. P. (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, 5, 1492. <https://doi.org/10.3389/fpsyg.2014.01492>
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116. <https://doi.org/dx.doi.org/10.1037/a0038889>
- Harvey, A. C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica: Journal of the Econometric Society*, 461–465.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97–109. <https://doi.org/10.2307/2334940>
- Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research*, 45(4), 627–660. <https://doi.org/doi.org/10.1080/00273171.2010.498290>
- Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics*, 64(2), 627–634.
- Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2012). Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Statistics in Medicine*, 31(27), 3328–3336. <https://doi.org/10.1002/sim.5338>
- Huang, L., Chen, M. H., & Ibrahim, J. G. (2005). Bayesian analysis for generalized linear models with nonignorably missing covariates. *Biometrics*, 61(3), 767–780. <https://doi.org/10.1111/j.1541-0420.2005.00338.x>

- Ibrahim, J. G. (1990). Incomplete data in generalized linear models. *Journal of the American Statistical Association*, 85(411), 765–769.  
<https://doi.org/10.1080/01621459.1990.10474938>
- Ibrahim, J. G., Chen, M. H., & Lipsitz, S. R. (2002). Bayesian methods for generalized linear models with covariates missing at random. *Canadian Journal of Statistics*, 30(1), 55–78.  
<https://doi.org/10.2307/3315865>
- Ibrahim, J. G., Lipsitz, S. R., & Chen, M. H. (1999). Missing covariates in generalized linear models when the missing data mechanism is non-ignorable. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(1), 173–190.  
<https://doi.org/10.1111/1467-9868.00170>
- Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions* (2nd ed.). Sage.
- Kasim, R. M., & Raudenbush, S. W. (1998). Application of Gibbs sampling to nested variance components models with heterogeneous within-group variance. *Journal of Educational and Behavioral Statistics*, 32, 93–116.
- Kelava, A., & Brandt, H. (2023). Latent Interaction Effects. In R. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 427–446). Guilford.
- Keller, B. T. (2022). An Introduction to Factored Regression Models with Blimp. *Psych*, 4, 10–37.
- Keller, B. T. (2024a). Blimp technical appendix: Nonnormal latent variables via the Yeo–Johnson transformation.
- Keller, B. T. (2024b). *rblimp: Integration of the Blimp Software Into R*. In (Version 0.1.31)  
Retrieved from <https://github.com/blimp-stats/rblimp>.
- Keller, B. T. (2025). *Automatically Derived Full Conditionals for Factored Regression Models with Missing Data and Latent Variables* International Workshop on Psychometric Computing, Ghent University, Belgium.
- Keller, B. T. (in press). A Factored Regression Approach to Modeling Latent Variable Interactions and Nonlinear Effects. *Psychological Methods*.
- Keller, B. T., & Enders, C. K. (2021). Blimp user’s guide (Version 3) [Computer Software].  
<https://www.appliedmissingdata.com/blimp>
- Keller, B. T., & Enders, C. K. (2023). An investigation of factored regression missing data methods for multilevel models with cross-level interactions. *Multivariate Behavioral Research*, 58(52023), 938–963. <https://doi.org/doi.org/10.1080/00273171.2022.2147049>
- Kim, J. (2024). *Factored Regression Approach for Structural Equation Models with Non-Normal Continuous Data* University of Texas].

- Kim, J., Lee, S., Kim, S., & Keller, B. T. (2025). Tutorial for Bayesian Multilevel Structural Equation Modeling Using Blimp. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–13. <https://doi.org/doi.org/10.1080/10705511.2025.2526035>
- Kim, S., Belin, T. R., & Sugar, C. A. (2018). Multiple imputation with non-additively related variables: Joint-modeling and approximations. *Statistical Methods in Medical Research*, 27(6), 1683–1694. <https://doi.org/10.1177/0962280216667763>
- Kim, S., Sugar, C. A., & Belin, T. R. (2015). Evaluating model-based imputation methods for missing covariates in regression models with interactions. *Statistics in Medicine*, 34(11), 1876–1888. <https://doi.org/10.1002/sim.6435>
- Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65(4), 457–474.
- Klein, A. G., & Muthén, B. O. (2007). Quasi-maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, 42(4), 647–673.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
- Klopp, E., & Klößner, S. (2021). The impact of scaling methods on the properties and interpretation of parameter estimates in structural equation models with latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(2), 182–206. <https://doi.org/doi.org/10.1080/10705511.2020.1796673>
- König, C., & van de Schoot, R. (2018). Bayesian statistics in educational research: a look at the current state of affairs. *Educational Review*, 70(4), 486–509. <https://doi.org/doi.org/10.1080/00131911.2017.1350636>
- Lee, M. C., & Mitra, R. (2016). Multiply imputing missing values in data sets with mixed measurement scales using a sequence of generalised linear models. *Computational Statistics & Data Analysis*, 95, 24–38. <https://doi.org/10.1016/j.csda.2015.08.004>
- Lee, S.-Y. (2007). *Structural Equation Modeling: A Bayesian Approach*. Wiley.
- Levy, R., & McNeish, D. (2023). Perspectives on Bayesian inference and their implications for data analysis. *Psychological Methods*, 28(3), 719–739. <https://doi.org/doi.org/10.1037/met0000443>
- Lipsitz, S. R., & Ibrahim, J. G. (1996). A conditional model for incomplete covariates in parametric regression models. *Biometrika*, 83(4), 916–922. <https://doi.org/10.1093/biomet/83.4.916>

- Liu, Y., & Pek, J. (2024). Summed versus estimated factor scores: Considering uncertainties when using observed scores. *Psychological Methods*, Advance online publication. <https://doi.org/dx.doi.org/10.1037/met0000644>
- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 201–229. <https://doi.org/10.1037/a0012869>
- Lüdtke, O., Robitzsch, A., & West, S. G. (2020a). Analysis of interactions and nonlinear effects with missing data: A factored regression modeling approach using maximum likelihood estimation. *Multivariate Behavioral Research*, 55(3), 361–381. <https://doi.org/10.1080/00273171.2019.1640104>
- Lüdtke, O., Robitzsch, A., & West, S. G. (2020b). Regression models involving nonlinear effects with missing data: A sequential modeling approach using Bayesian estimation. *Psychological Methods*, 25, 157–181. <https://doi.org/10.1037/met0000233>
- McManus, K. R., Enders, C., Kirsch, D. E., Grodin, E., & Ray, L. A. (In preparation). Peripheral inflammation mediates the relationship between early life stress and alcohol use.
- McNeish, D. (2021). Specifying location-scale models for heterogeneous variances as multilevel SEMs. *Organizational Research Methods*, 24(3), 630–653.
- McNeish, D. (2024). Practical implications of sum scores being psychometrics' greatest accomplishment. *Psychometrika*, 89(4), 1148–1169. <https://doi.org/doi.org/10.1007/s11336-024-09988-z>
- McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, 25(5), 610–635. <https://doi.org/10.1037/met0000250>
- McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, 52, 2287–2305. <https://doi.org/10.3758/s13428-020-01398-0>
- Merkle, E. C., Furr, D., & Rabe-Hesketh, S. (2019). Bayesian comparison of latent variable models: Conditional versus marginal likelihoods. *Psychometrika*, 84(3), 802–829.
- Merkle, E. C., & Rosseel, Y. (2018). blavaan: Bayesian structural equation models via parameter expansion. *Journal of Statistical Software*, 85(4), 1–30. <https://doi.org/10.18637/jss.v085.i04>

- Mislevy, R. J. (2024). Are sum scores a great accomplishment of psychometrics or intuitive test theory? *Psychometrika*, 89(4), 1170–1174. <https://doi.org/doi.org/10.1007/s11336-024-10003-8>
- Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. Chapman and Hall/CRC.
- Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal*, 28(4), 638–648. <https://doi.org/10.1080/10705511.2020.1784738>
- Muthén, B., & Asparouhov, T. (2011). Beyond multilevel regression modeling: Multilevel analysis in a general latent variable framework. In J. J. Hox & J. K. Roberts (Eds.), *Handbook for advanced multilevel analysis* (pp. 15–40). Routledge.
- Muthén, B., Asparouhov, T., & Shiffman, S. (2024). Dynamic structural equation modeling with floor effects. *Submitted for publication*.
- Neelon, B. (2019). Bayesian zero-inflated negative binomial regression based on pólya—gamma mixtures. *Bayesian Analysis*, 14(3), 829–855.
- Neelon, B., & O'Malley, A. J. (2019). Two-part models for zero-modified count and semicontinuous data. In A. Levy, S. Goring, C. Gatsonis, B. Sobolev, E. v. Ginneken, & R. Busse (Eds.), *Health Services Evaluation* (pp. 695–716). Springer, New York, USA.
- Olsen, M. K., & Schafer, J. L. (2001). A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association*, 96(454), 730–745.
- Palomo, J., Dunson, D. B., & Bollen, K. (2007). Bayesian structural equation modeling. In S.-Y. Lee (Ed.), *Handbook of latent variable and related models* (pp. 163–188). Elsevier.
- Pillow, J., & Scott, J. (2012). Fully Bayesian inference for neural models with negative-binomial spiking. *Advances in neural information processing systems*, 25, 1–9.
- Polson, N. G., Scott, J. G., & Windle, J. (2013). Bayesian inference for logistic models using pólya—gamma latent variables. *Journal of the American Statistical Association*, 108(504), 1339–1349. <https://doi.org/10.1080/01621459.2013.829001>
- Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, 21, 189–205. <https://doi.org/dx.doi.org/10.1037/met0000052>
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15(3), 209–233. <https://doi.org/10.1037/a0020141>

- Quartagno, M., & Carpenter, J. R. (2019). Multiple imputation for discrete data: Evaluation of the joint latent normal model. *Biometrical Journal*, 61, 1003–1019.  
<https://doi.org/10.1002/bimj.201800222>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.
- Reise, S. P., Rodriguez, A., Spritzer, K. L., & Hays, R. D. (2018). Alternative approaches to addressing non-normal distributions in the application of IRT models to personality measures. *Journal of Personality Assessment*, 100(4), 363–374.  
<https://doi.org/doi.org/10.1080/00223891.2017.1381969>
- Rhemtulla, M., & Savalei, V. (2024). Estimated Factor Scores Are Not True Factor Scores. *Multivariate Behavioral Research*, Advanced online publication, 1–22.  
<https://doi.org/doi.org/10.1080/00273171.2024.2444943>
- Riani, M., Atkinson, A. C., & Corbellini, A. (2023). Automatic robust Box–Cox and extended Yeo–Johnson transformations in regression. *Statistical Methods & Applications*, 32(1), 75–102.
- Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*, 24, 309–338. <https://doi.org/10.1037/met0000184>
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.
- Sijtsma, K., Ellis, J. L., & Borsboom, D. (2024). Recognize the Value of the Sum Score, Psychometrics' Greatest Accomplishment. *Psychometrika*, 89(1), 84–117.  
<https://doi.org/doi.org/10.1007/s11336-024-09964-7>
- Smith, V. A., Neelon, B., Maciejewski, M. L., & Preisser, J. S. (2017). Two parts are better than one: modeling marginal means of semicontinuous data. *Health Services and Outcomes Research Methodology*, 17, 198–218. <https://doi.org/10.1007/s10742-017-0169-9>
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583–639. <https://doi.org/doi.org/10.1111/1467-9868.00353>
- Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82, 528–540.
- van de Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: The last 25 years. *Psychological Methods*, 22(2), 217–239. <https://doi.org/10.1037/met0000100>

- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11(12), 3571–3594.
- Widaman, K. F., & Revelle, W. (2022). Thinking thrice about sum scores, and then some more about measurement and analysis. *Behavior Research Methods*, 1–19.  
<https://doi.org/10.3758/s13428-022-01849-w>
- Widaman, K. F., & Revelle, W. (2024). Thinking about sum scores yet again, maybe the last time, we don't know, Oh no...: a comment on. *Educational and Psychological Measurement*, 84(4), 637–659. <https://doi.org/10.1177/00131644231205310>
- Windle, J., Polson, N. G., & Scott, J. G. (2014). Sampling Pólya-gamma random variates: alternate and approximate techniques. *arXiv preprint arXiv:1405.0506*.
- Woller, M. P., & Enders, C. K. (2024). Exploration of the MCMC Wald test with linear regression. *Behavioral Research Methods*. <https://doi.org/doi.org/10.3758/s13428-024-02426-z>
- Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. *Biometrika*, 87(4), 954–959. <https://doi.org/10.1093/biomet/87.4.954>
- Zhang, Q., & Wang, L. (2017). Moderation analysis with missing data in the predictors. *Psychological Methods*, 22(4), 649–666. <https://doi.org/10.1037/met0000104>