Journal of Behavioral Data Science, 2022, *2* (1), 75–105.
DOI: https://doi.org/10.35566/jbds/v2n1/p5

The Impact of Sample Size on Exchangeability in the Bayesian Synthesis Approach to Data Fusion⋆

Katerina M. Marcoulides*[1], Jia Quan[2], and Eric Wright[2]

[1] Department of Psychology, University of Minnesota
kmarcoul@umn.edu
[2] University of Florida

Abstract. Data fusion approaches have been adopted to facilitate more complex analyses and produce more accurate results. Bayesian Synthesis is a relatively new approach to data fusion where results from the analysis of one dataset are used as prior information for the analysis of the next dataset. Datasets of interest are sequentially analyzed until a final posterior distribution is created, incorporating information from all candidate datasets, rather than simply combining the datasets into one large dataset and analyzing them simultaneously. One concern with this approach lies in the sequence of datasets being fused. This study examines whether the order of datasets matters when the datasets being fused each have substantially different sample sizes. The performance of Bayesian Synthesis with varied sample sizes is evaluated by examining results from simulated data with known population values under a variety of conditions. Results suggest that the order in which the dataset are fused can have a significant impact on the obtained estimates.

Keywords: Bayesian synthesis · Data fusion · Exchangeability

1 Introduction

Researchers in psychology and in the social and behavioral sciences more broadly, have recently expressed concerns about a "replication crisis" (Maxwell, Lau, & Howard, 2015). These concerns have driven researchers to explore and develop new strategies for analyzing data across multiple studies and summarizing results. In an effort to combat the replication crisis, open-source data repositories have grown substantially (Bhattacharya & Saha, 2015), and this greater access

⋆ Early versions of this simulation study were presented at the International Meeting of the Psychometric Society and the Annual Meeting of the Florida Educational Research Association. This paper extends the simulation conditions in the first author's dissertation submitted in partial fulfillment of the doctoral degree at Arizona State University, Tempe.

to data both enables and requires the development of new strategies for exploring and analyzing this large amount of publicly available data. Data fusion is one method that has been shown to enable more complex and more appropriate models to be fit to fused (i.e., combined) datasets than just a single dataset (Curran & Hussong, 2009; Marcoulides, 2018). Bayesian Synthesis is a recently proposed Bayesian approach to data fusion whereby results from the analysis of one dataset are used as prior information in the subsequent analysis of the next dataset (Du et al., 2020; Marcoulides, 2017b), and those results are in turn used as prior information for the analysis of yet another dataset. Datasets of interest are sequentially analyzed in this manner until a final posterior distribution is created, which incorporates information from all datasets of interest.

Conducting data fusion using the Bayesian Synthesis approach thus relies on the sequential updating of estimates as new information from each additional dataset becomes available (for complete technical details on the approach and various recent empirical applications see for example, Fujimoto, Gordon, Peng, & Hofer, 2018; Johnson & Guttmannova, 2019; Marcoulides, 2017a, 2017b, 2018; Preston et al., 2018; Saris & Satorra, 2018). This synthesis notion is expressed using Bayes theorem as

$$\begin{aligned} P(Unknowns|Data) &= \frac{P(Data|Unknowns)P(Unknowns)}{P(Data)} \\ &\propto P(Data|Unknowns)P(Unknowns), \end{aligned}$$

where $P(Unknowns)$ is the prior probability distribution for the unknown parameters, $P(Data|Unknowns)$ is the conditional probability of the data given the unknown parameters, and $P(Unknowns|Data)$ is the posterior probability distribution for the unknown parameters given our data. Thus when two datasets are fused, the prior information about the unknown parameters can be considered equivalent to a data set that, when merged with the current data, supports the following Bayesian inference $P\left(Unknowns|\, Data_1, Data_2\right) \propto P\left(Data_2|\, Unknowns\right) P\left(Unknowns|\, Data_1\right)$. Here, $P(Unknowns|Data_1)$ is the posterior distribution that resulted from the first analysis where information from $Data_1$ was incorporated with $P(Unknowns)$ and then serves as the prior distribution for the present analysis that incorporates the data in $Data_2$. When k datasets are to be fused, the process can be denoted in a general form as $P\left(Unknowns|\, Data_1, \ldots, Data_{k+1}\right) \propto P(Data_{k+1}|\, Unknowns) P\left(Unknowns|\, Data_1, \ldots, Data_k\right)$ with the priors and posterior distribution similarly updated.

A major benefit of the Bayesian Synthesis approach over traditional frequentist approaches to data fusion is the ability to incorporate datasets for which raw data is not available. In this manner, the Bayesian Synthesis approach provides an alternative to the necessity to analyze the raw data and instead uses the estimates and summary statistics from the examined studies to incorporate into the prior information. The approach therefore utilizes point summary estimates of the posterior distributions instead of the actual full posterior distributions as required by a fully Bayesian execution of this Bayesian Synthesis approach. Bayesian Synthesis thereby enables summary information from published (or unpublished) research to be incorporated as prior information (i.e., an informative prior) for the analysis of another dataset. The sequential use of informative priors that are based on the information in past data provides an extra source of information to estimate model parameters and this additional information can effectively aid in the accuracy of parameters estimation and in the interpretation of results. However, one concern with the Bayesian Synthesis approach is that it heavily relies on updating the information as new data summary statistics become available, therefore the order in which the data are sequentially analyzed may have an impact on the results (Marcoulides, 2017b). Theoretically, in the Bayesian Synthesis approach this should not be a concern due to the conventional Bayesian exchangeability assumption (de Finetti, 1972, 1974), however, Bayesian Synthesis utilizes point summary estimates of the posterior distributions instead of the full posterior distribution as required for a fully Bayesian execution of this approach. While this has the potential to introduce some bias, using point summary estimates of the posterior distributions greatly increases the ease of execution and enables researchers to straightforwardly implement the Bayesian Synthesis approach in standard programs like M*plus* (Muthen & Muthen, 2017). The cost of potentially introducing bias may outweigh the difficulties of incorporating the full distributions in the sequential analysis (Marcoulides, 2017b).

To address the concern about the order of the datasets being analyzed, Marcoulides (2017b) examined the exchangeability assumption and found that the order of analysis did not meaningfully impact the final data fusion results. Similar conclusions regarding exchangeability were also recently suggested by Miocevic, Levy, and Savord (2020). One limitation with these conclusions is that they were based on analyzed datasets that were from similarly-sized large samples. Therefore, it is still unknown whether the order of datasets matters when the datasets being fused each have substantially different sample sizes (as is quite common with empirical data). For example, it may be that beginning the Bayesian Synthesis approach with the analysis of a large dataset produces a substantially biased final posterior distribution when the other sequentially analyzed datasets are much smaller, or vice versa.

In this study, we focus on this unexamined scenario in which there are multiple datasets of both small and large sample sizes. Our main question is whether the order in which datasets are incorporated in the Bayesian Synthesis process will impact the results when one dataset is substantially larger than the rest. To evaluate the performance of Bayesian Synthesis with varied sample sizes, results from simulated data with known population values will be examined under a variety of design study conditions. We conclude with a discussion of the results, implications of the findings, and suggestions for further research.

2 Methods

2.1 Monte Carlo Data Simulation

In order to systematically evaluate the exchangeability of datasets with varying sample sizes in the Bayesian Synthesis approach, simulated data using Monte Carlo techniques were analyzed under a variety of longitudinal data design conditions. All simulated data were generated using R (R Development Core Team, 2010), and analyzed in M*plus* (Muthen & Muthen, 2017) through R using the MplusAutomation package (Hallquist & Wiley, 2014). The analyses were specified to use the Gibbs (PX1) algorithm with a minimum of 50,000 iterations, using the Potential Scale Reduction (PSR) convergence criteria of 1.05 (Gelman et al., 2014), a median summarized posterior, 250 replications, and two chains without thinning.

The simulated data were modeled after the Marcoulides and Grimm (2017) study, which analyzed six longitudinal studies measuring students' mathematics ability. These studied six datasets varied in their sample size, timing, and number of measurement occasions. As Marcoulides and Grimm (2017) showed that mathematics ability increased as children got older, we used the linear growth model to generate the simulated data. In this model, individuals (n) are measured on their math abilities (y) across multiple measurement occasions (t), using the data generation model

$$y_{tn} = \eta_{0n} + \left(\frac{t - k_1}{k_2}\right) \eta_{1n} + e_{tn}, \tag{1}$$

where η_{0n} is an individual's latent intercept when time t equals zero, η_{1n} is an individual's latent slope when time t equals zero, k_1 and k_2 are functional variables to center the intercept and scale the slope, and e_{tn} is the residual at time t for individual n. In this model, we assume the two latent variables follow a multivariate normal distribution, $[\eta_{0n},\ \eta_{1n}]' \sim N(\boldsymbol{\beta},\ \boldsymbol{\Psi})$; and the residuals are assumed to follow a normal distribution, $e_{tn} \sim N(0,\ \sigma_e^2)$, with mean 0 and constant variance.

In all simulation conditions, the latent intercept and slope means were fixed at $\beta_{Intercept}$= -2 and $\beta_{Slope} = 0.4$, and the residual variance σ_e^2 was fixed at 0.10 to reflect small amounts of residuals. Asparouhov and Muthén (2010) and McNeish (2016) indicated that, compared to other parameters of the linear growth model, the choice of prior distributions for the variance-covariance matrix ($\boldsymbol{\Psi}$) in this model is extremely important. Therefore, we additionally varied the $\boldsymbol{\Psi}$ matrix to reflect small, medium, and large magnitudes of their variances, and zero and small magnitudes of their covariances, resulting in the following three variance–covariance matrices: $\boldsymbol{\Psi_1} = \begin{bmatrix} 0.20 & 0.0 \\ 0.0 & 0.01 \end{bmatrix}$, $\boldsymbol{\Psi_2} = \begin{bmatrix} 0.70 & 0.05 \\ 0.05 & 0.10 \end{bmatrix}$, and $\boldsymbol{\Psi_3} = \begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$. A summary of these population parameters is presented in Table 1 below. Figure 1 provides an illustrative display of data score plots for one simulated sample of N = 50 observations based on these three variance-covariance matrices for a simulated data design condition comprised of observations obtained across 3 assessment occasions taken every 5 years starting at age 5.

Table 1: Population matrices and covariance parameters.

$\boldsymbol{\Psi}$	$\boldsymbol{\beta_{Intercept}}$	$\boldsymbol{\beta_{Slope}}$	σ_e^2
$\begin{bmatrix} 0.20 & 0.0 \\ 0.0 & 0.01 \end{bmatrix}$	-2.00	0.40	0.10
$\begin{bmatrix} 0.70 & 0.05 \\ 0.05 & 0.10 \end{bmatrix}$	-2.00	0.40	0.10
$\begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$	-2.00	0.40	0.10

For each variance-covariance matrix condition considered, six datasets were simulated. Because the impact of sample size on exchangeability may also depend on how well the particular sample matches the population of interest, the six datasets were varied with respect to the number of assessments, years between assessments, the age of participants' first assessment, and the sample size. The data patterns for these six datasets are presented in Table 2 and are meant to reflect the full growth trajectory (i.e., across the full age range of interest) with early and late age ranges, as well as large and small numbers of observations across different numbers of assessment occasions. As indicated, the sample sizes for these six datasets were also varied across each $\boldsymbol{\Psi}$ matrix condition, such that each of these datasets were simulated to have a sample size of 1000 and incorporated into the Bayesian synthesis approach as the first dataset, randomly varying the order of the remaining 5 datasets each with sample size of 50, and then again as the last dataset, randomly varying the order of the preceding 5 datasets. These different sample sizes of 50 and 1,000 were selected to reflect small and large sample studies that are commonly encountered in longitudinal data analyses (McNeish, 2016; Paxton, Curran, Bollen, Kirby, & Chen, 2001).

The different specifications for the simulated data presented in detail in Table 2 resulted in a total of 36 different simulated data conditions (3 $\boldsymbol{\Psi}$ matrices, 6 data patterns, and 2 fusing sequences). For example, for the first $\boldsymbol{\Psi_1}$ matrix condition, if dataset 1 (measured 3 times with five years between assessments, starting at age 5) was simulated to have 1000 individuals, the other 5 datasets were then simulated to have a sample size of 50 observations. Thus, the Bayesian Synthesis approach begins with the analysis of dataset 1 with 1,000 observations and produces a posterior point estimate that is then used as the informative prior for the analysis of say dataset 2 with 50 individuals. This process contin-

$$\Psi_1 = \begin{bmatrix} 0.20 & 0.0 \\ 0.0 & 0.01 \end{bmatrix}$$

$$\Psi_2 = \begin{bmatrix} 0.70 & 0.05 \\ 0.05 & 0.10 \end{bmatrix}$$

$$\Psi_3 = \begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$$

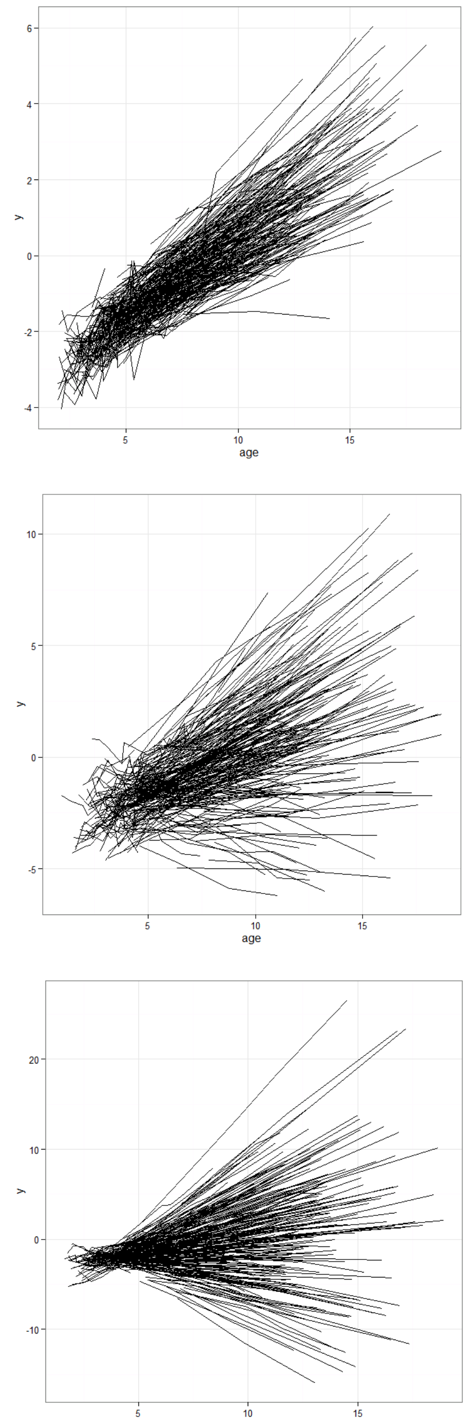

Figure 1: Illustrative data score plots for N = 50 for the first data design.

Table 2: List of the patterns of measurement occasions to be used in the simulated data

Dataset	Number of Assessments	Years between Assessments	StaringAge
1	3	5	5
2	10	1	5
3	10	0.5	2.5
4	10	0.5	10
5	3	0.5	4
6	3	1	11

ues accordingly for each of the remaining datasets until a final posterior point estimate is produced that incorporates information from all 6 datasets. As noted above, this same process is also conducted again but instead with the 1000 observations utilized as the last dataset while randomly varying the order of the preceding 5 datasets each with 50 observations (see additional details below).

2.2 The Bayesian Synthesis Approach

After simulating the datasets according to the conditions described above, we conducted data fusion using the Bayesian Synthesis approach. We began this sequential data integration process with non-informative priors for the analysis of the first dataset because, according to Asparouhov and Muthén (2010), these initial non-informative priors should not introduce bias, even in small sample sizes. The rationale for using non-informative priors also follows recommendations provided by Gelman et al. (2014) to "…let the data speak for themselves, so that inferences are unaffected by information external to the current data (pg. 51)". This is because "the information about model parameters contained in the data will far outweigh any reasonable prior probability specification" (Gelman et al., 2014). Similar recommendations concerning the use of non-informative priors for estimation in growth curve analyses were also provided by Liu, Zhang, and Grimm (2016). So, unless dependable prior information about the range of possible values that model parameters might take is available, beginning the sequential data integration process with non-informative priors appears to be preferable to simply taking a guess at the values for the priors and using informative inaccurate priors, which are known to yield less accurate estimates than non-informative priors (e.g., Shi & Tong, 2017, 2018).

Thus, the Bayesian Synthesis approach begins by using non-informative priors as parameters for the first data set. This implies that for the intercept and slope means a Normal prior of the form N(mean, variance) is used with N(0, 10^{10}), which is the default for a non-informative Normal prior in M*plus* (Muthen & Muthen, 2002). Then, for the parameters in the $\boldsymbol{\Psi}$ matrix the Inverse Wishart prior IW(0,-3) is used, which is also the default non-informative prior in *Mplus* (Muthen & Muthen, 2002). This prior is of the general form IW(S, d), where

d is the pseudo-sample size and S is the scale matrix $\begin{bmatrix} d(\sigma^2_{Intercept}) & d(\sigma_{IS}) \\ d(\sigma_{IS}) & d(\sigma^2_{Slope}) \end{bmatrix}$, with the estimated intercept ($\sigma^2_{Intercept}$) and slope (σ^2_{Slope}) variances and their covariance (σ_{IS}). Finally, the Inverse Gamma IG(-1, 0) for the residual variance, which is also the default non-informative prior in *Mplus* (Muthen & Muthen, 2017). The Inverse Gamma is of the general form IG(α, β), in which $\alpha = \nu_0/2$ and $\beta = \nu_0\sigma_0^2/2$, and where σ_0^2 can be interpreted as the best estimate of the variance and ν_0 can be interpreted as a pseudo-sample size. Upon analyzing the first data set based on the non-informative priors, posterior point summary estimates for the $\boldsymbol{\beta}$, $\boldsymbol{\Psi}$, and σ_e^2 parameters are then sequentially substituted into the respective priors for the next data analysis and the pseudo-sample size of the current data set is then changed by the sample size of the previous data set. This process then makes the priors in the subsequent analyses informative priors and continues sequentially until the sixth and final data set is analyzed and the final posterior distribution and point estimates produced. The simulation results for these analyses are presented in Tables 3a through 8a.

To further investigate the extent to which the use of non-informative versus informative priors at the outset of Bayesian Synthesis might also induce some sort of bias into the obtained results, we also performed all sequential data integration processes using informative priors for the analysis of the first dataset. Given that in practice it may be unlikely for a researcher to have exactly accurate prior information regarding the parameter values, we followed the recommendations of Depaoli (2014) and Finch and Miller (2019) to use "informative" priors as those in which the specified priors correspond to the estimated maximum likelihood growth parameters for each model. Using these estimated values as informative priors, the analyses were then repeated using the 36 different specified simulation data conditions (3 $\boldsymbol{\Psi}$ matrices, 6 data patterns, and 2 fusing sequences) and are presented in Table 3b to Table 8b.

2.3 Parameter Evaluation

The final posterior point summary estimates obtained for each of the 36 different simulated data conditions were evaluated in terms of their raw bias ($B(\hat{y})$), relative bias ($RB(\hat{y})$), accuracy ($RMSE(\hat{y})$), and efficiency ($\mathit{Efficiency}(\hat{y})$. These criteria were selected based on past research on dependable parameter evaluation benchmarks (Bandalos & Gagne, 2012; Bandalos & Leite, 2013). These criteria were computed based on the following formulas:

$$B(\hat{y}) = \frac{\sum_{r=1}^{R} (\hat{y}_r - y)}{R} \tag{2}$$

$$RB(\hat{y}) = \left(\frac{\hat{y} - y}{y}\right) \times 100 \tag{3}$$

$$RMSE\,(\hat{y}) = \sqrt{\frac{\sum_{r=1}^{R}(\hat{y}_r - y)^2}{R-1}} \tag{4}$$

$$Efficiency\,(\hat{y}) = \sqrt{\frac{\sum_{r=1}^{R}(\hat{y}_r - \overline{\hat{y}})^2}{R-1}} \tag{5}$$

where R represents the total number of simulation replications, which is 250 in our study; $\hat{y}$ is the estimated parameter; y is the known population value for our simulation; and $\overline{\hat{y}}$ is the average parameter estimate.

Raw bias, accuracy, and efficiency in essence evaluate the average deviation, the square root of the average deviation, and the variability of the final posterior distribution means, respectively. Nonzero positive or negative values of raw bias indicates overestimation or underestimation respectively. Lower values of accuracy correspond to more precise estimates of the parameters, or estimates of parameters that exhibit a smaller range of error (Bandalos & Gagne, 2012). Values closer to zero correspond to more efficient estimates of the parameters. In other words, smaller values correspond to a smaller range of variability, or higher consistency of estimation. In contrast to these criteria, the magnitude of relative bias is expressed on the percentage scale and indicates the percent deviation of the estimate from the population parameter. This measure is ideal for comparisons of the magnitude of bias across different design conditions (Muthen & Muthen, 2002). To evaluate relative bias, it has been suggested that values less than 5% reflect ignorable bias, values between 5% to 10% indicate moderate bias, and values larger than 10% are considered substantial bias (Muthen & Muthen, 2002). Because multiple simulation replications were conducted (in this case 250 replications were analyzed for all linear growth models examined), average values of relative bias for each evaluated parameter are reported across all the replications. In general, the values of raw bias, accuracy, and efficiency are typically much harder to unravel, whereas values of relative bias are much easier to interpret; we therefore pay extra attention to disentangling obtained relative bias values in reviewing the crucial findings in our study.

3 Results

The simulation results for examining the performance of the exchangeability principle applied within the Bayesian Synthesis approach are presented in Tables 3a through 8b. Within each table, the results are organized based on the magnitude of the variance-covariance $\mathbf{\Psi}$ matrix (in order of small, medium, and large magnitudes) and the order of data fusion (i.e., with the large dataset fused first versus last). The computed parameter estimates reported include the latent variable means $\beta_{Intercept}$ and β_{Slope}, their variances and covariances in terms of $\sigma^2_{Intercept}$, σ^2_{Slope}, σ_{IS}, and the residual variance σ^2_e. Each table presents findings based on the designated evaluation criteria (raw bias, relative bias, accuracy, and efficiency) for each of the parameter estimates across the examined data design conditions with both non-informative and informative priors. Except for estimates of $\beta_{Intercept}$, positive values of the criteria indicate positive bias or overestimation and negative values indicate negative bias or underestimation. Because $\beta_{Intercept}$ was fixed in the simulation conditions at a negative value, obtaining a negative bias actually corresponds to overestimation and a positive bias corresponds to underestimation.

The results for the first data design condition with specified covariance matrices $\Psi_1 = \begin{bmatrix} 0.20 & 0.0 \\ 0.0 & 0.01 \end{bmatrix}$, $\Psi_2 = \begin{bmatrix} 0.70 & 0.05 \\ 0.05 & 0.10 \end{bmatrix}$, and $\Psi_3 = \begin{bmatrix} 0.40 & 0.20 \\ 0.20 & 0.40 \end{bmatrix}$ are presented in Table 3a. The first data design condition comprised observations obtained across 3 assessment occasions taken every 5 years starting at age 5. This dataset reflected the feature of large breadth with small numbers of assessments while covering different age ranges. The reported results correspond to those obtained when (i) a sample size of 1000 observations is incorporated as the first dataset into the Bayesian synthesis approach while randomly varying the order of the remaining 5 datasets each with a sample size of 50, and (ii) a sample size of 1000 observations is incorporated into the Bayesian synthesis approach as the last dataset while randomly varying the order of the preceding 5 datasets each with a sample of size 50. To make this distinction clear, the results are labelled in this and all subsequent tables as FIRST and LAST for each reported evaluation criterion.

In general, the obtained values for the raw bias, accuracy, and efficiency criteria are larger when the larger dataset is fused last instead of first in the Bayesian synthesis approach. However, these values are all still very close to zero, indicating precise and efficient estimates of the parameters. These obtained bias values also appear to be similar regardless of which specified variance-covariance matrix is examined. Looking for example at the estimated intercept variance ($\sigma^2_{Intercept}$) for the first dataset condition in Table 3a, some patterns of results can be detected. For instance, the obtained values for the raw bias (.0011, .0048, and .0029), accuracy (.0133, .0331, and .0212), and efficiency (.0133, .0327, .0210) when the large dataset was fused first are negligible. However, when the large dataset is fused last, the estimated intercept variance values increase (though are still rather close to zero) for the raw bias (.1244, .1300, and .1286), the accuracy (.2664, .2828, and .2776), and efficiency (.2355, .2511, .2459). Similar patterns of results are observed when informative priors are used to analyze the large data set first (see Table 3b). As indicated previously, because the values of raw bias, accuracy, and efficiency are typically more challenging to unravel, we instead pay extra attention to disentangling the obtained relative bias values as these are generally easier to interpret.

When examining the relative bias criterion under the three variance-covariance matrixes, some interesting patterns of results are again revealed. Specifically, with initial non-informative priors it can be seen that the relative bias for the estimated intercept variance ($\sigma^2_{Intercept}$) increases from ignorable sizes (0.546%, 0.692%, and 0.740%, respectively) when the large dataset was fused first into substantially biased values (62.208%, 18.572%, and 32.161%, respectively) when the large dataset was fused last. The estimates of the slope variance (σ^2_{Slope}) also showed somewhat similar patterns of results, but different magnitudes when fusing the larger sample dataset last. The variances changed from ignorable biased (0.720%, 0.880%, and 0.802%, respectively) when the large dataset was fused first, to substantially biased values (58.12%) for the small magnitude covariance matrix (Ψ_1) and moderately biased (8.31%, 6.336%) for medium and large magnitude covariance matrixes (Ψ_2 and Ψ_3) when the larger dataset was fused last. Another sizable amount of relative bias was also observed when examining the magnitude of the intercept slope covariance value (σ_{IS}) for the Ψ_2 variance-covariance matrix, shifting from ignorable bias (0.384%) when the large dataset was fused first to substantially biased (-12.984%) when the larger dataset was fused last.

In contrast, when initial informative priors are used, the relative bias for the residual variance when the large sample is incorporated into the Bayesian synthesis approach as the last dataset were found to be moderate across the three covariance matrix conditions (7.248 %, 5.244%, and 5.108%). Additionally, the relative bias for the estimated intercept ($\sigma^2_{Intercept}$) and slope (σ^2_{Slope}) variances increased from ignorable sizes when the large dataset was fused first into substantially biased for the first covariance matrix condition (12.01% and 16.52% respectively) when the large dataset was fused last.

The observed relative bias for the variance of the intercept, the slope, and to a lesser degree the intercept-slope covariance in this first data design condition highlight the importance that the order of fusing the datasets can play in the estimation of parameters when implementing Bayesian Synthesis strategies. Interestingly, this finding is not in line with past research that has suggested that the order of data fusion does not meaningfully impact the final posterior distribution results (Marcoulides, 2017b; Miocevic et al., 2020). When the data sets being fused are of differing sizes (50 vs. 1,000), ending with the fusion and analysis of a large dataset can in fact produce a substantially biased final posterior distribution when the other sequentially analyzed datasets are much smaller. However, these results are only discernable when using the measure of relative bias, they do not appear sizeable when examining the values of raw bias, accuracy, or efficiency.

The results for the second simulated data design condition for the variance-covariance matrices Ψ_1, Ψ_2, and Ψ_3 are presented in Table 4a. The second data design condition comprised of observations across 10 assessment occasions, measured every year, starting from age 5. This simulated condition reflected the feature of large breadth of measurement years covering different age ranges. As described previously, the reported results correspond to those obtained when a sample size of 1000 observations is incorporated as the first dataset and again as the last dataset while randomly varying the order of the other 5 datasets each with 50 observations. The results are similarly labelled as FIRST and LAST for each evaluation criterion examined.

In general, the obtained values for the raw bias, accuracy, and efficiency criteria reflect similar results to those observed for the first simulated data design condition. Looking for example at the estimated intercept variance ($\sigma^2_{Intercept}$)

Table 3a: Data Condition 1 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	0.0005[a]	-0.0009	0.4840[a]	-0.9440	0.0033[a]	0.0055	0.0033[a]	0.0054
	0.0006[b]	-0.0002	0.5600[b]	-0.2200	0.0033[b]	0.0042	0.0033[b]	0.0042
	0.0005[c]	-0.0003	0.5000[c]	-0.2880	0.0034[c]	0.0053	0.0033[c]	0.0053
σ_{IS}	0.0001	-0.0154	−−	−−	0.0019	0.0323	0.0019	0.0284
	0.0002	0.0065	0.3840	**-12.984**	0.0083	0.0389	0.0083	0.0384
	0.0000	0.0040	0.3520	0.1960	0.0140	0.0614	0.0140	0.0614
$\beta_{Intercept}$	0.0007	0.0016	-0.0334	-0.0776	0.0163	0.0172	0.0163	0.0171
	0.0011	0.0020	-0.0570	-0.0986	0.0267	0.0267	0.0267	0.0266
	0.0007	0.0015	-0.0354	-0.0740	0.0212	0.0209	0.0212	0.0208
β_{Slope}	-0.0001	-0.0001	-0.0220	-0.0130	0.0034	0.0034	0.0034	0.0034
	0.0001	0.0002	0.0130	0.0450	0.0091	0.0092	0.0091	0.0092
	0.0006	0.0008	0.1430	0.2000	0.0181	0.1790	0.0181	0.0179
$\sigma^2_{Intercept}$	0.0011	0.1244	0.5460	**62.208**	0.0133	0.2664	0.0133	0.2355
	0.0048	0.1300	0.6920	**18.573**	0.0331	0.2828	0.0327	0.2511
	0.0029	0.1286	0.7370	**32.161**	0.0212	0.2776	0.0210	0.2459
σ^2_{Slope}	0.0001	0.0058	0.7200	**58.120**	0.0006	0.0107	0.0006	0.0090
	0.0009	0.0083	0.8800	**8.312**	0.0043	0.0189	0.0042	0.0169
	0.0032	0.0253	0.8020	**6.336**	0.0168	0.0589	0.0165	0.0531

Note: [a] Denotes results for covariance matrix Ψ_1, [b] denotes results for covariance matrix Ψ_2, and [c] denotes results for covariance matrix Ψ_3. The relative bias for estimates of the intercept-slope covariance for Ψ_1 cannot be computed, as the population value was zero. For $\beta_{intercept}$, negative bias corresponds to overestimation and positive bias corresponds to underestimation. For all other parameters values negative bias corresponds to underestimation and positive bias corresponds to overestimation. Bolded values indicate moderate or substantial bias.

Table 3b: Data Condition 1 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ^2_e	0.0003[a]	0.0072	0.3280[a]	**7.2480**	0.0031[a]	0.0169	0.0031[a]	0.0152
	0.0004[b]	0.0052	0.3640[b]	**5.2440**	0.0032[b]	0.0220	0.0031[b]	0.0214
	0.0003[c]	0.0051	0.3240[c]	**5.1080**	0.0031[c]	0.0172	0.0031[c]	0.0164
σ_{IS}	-0.0001	-0.0050	−−	−−	0.0020	0.0233	0.0020	0.0228
	-0.0003	-0.0025	-0.6880	-4.9920	0.0085	0.0380	0.0085	0.0379
	-0.0006	-0.0098	-0.2820	-4.9060	0.0148	0.0615	0.0148	0.0607
$\beta_{Intercept}$	0.0000	0.0004	-0.0016	-0.0206	0.0156	0.0166	0.0156	0.0166
	-0.0003	0.0009	0.0130	-0.0468	0.0254	0.0254	0.0254	0.0254
	-0.0004	0.0011	0.0200	-0.0528	0.0202	0.0217	0.0202	0.0217
β_{Slope}	-0.0002	0.0000	-0.0420	0.0010	0.0033	0.0036	0.0033	0.0036
	-0.0002	0.0003	-0.0490	0.0700	0.0093	0.0094	0.0093	0.0094
	-0.0004	0.0008	-0.0940	0.1930	0.0185	0.0196	0.0185	0.0196
$\sigma^2_{Intercept}$	0.0007	0.0240	0.3260	**12.0100**	0.0136	0.1740	0.0136	0.1723
	0.0014	-0.0103	0.1994	-1.4657	0.0337	0.2292	0.0337	0.2290
	0.0015	0.0128	0.3690	3.1940	0.0216	0.1890	0.0216	0.1886
σ^2_{Slope}	0.0000	0.0017	-0.3200	**16.520**	0.0006	0.0066	0.0006	0.0064
	0.0002	0.0006	0.1920	0.6080	0.0042	0.0153	0.0042	0.0152
	0.0004	0.0009	0.0960	0.2210	0.0159	0.0416	0.0159	0.0416

Note: Same as Table 3a.

for the first dataset condition in Table 4a, the same patterns of results can be detected. For instance, the obtained values for the raw bias (.0064, .0219, and .0127), accuracy (.0196, .0685, and .0396), and efficiency (.0186, .0649, .0375) when the large dataset was fused first are negligible. However, when the large dataset is fused last, the estimated intercept variance values increase (though are still rather close to zero) for the raw bias (.0688, .0756, and .0721), accuracy (.1444, .1521, and .1474), and efficiency (.1268, .1319, .1285).

Focusing on the measure of relative bias with initial non-informative priors, the biased values are observed when comparing estimates to the true population values when the larger dataset was fused last in the Bayesian synthesis approach and regardless of the specified variance-covariance matrix examined. Specifically, one can see that the relative bias for the estimated intercept variance ($\sigma^2_{Intercept}$) increases from ignorable sizes (3.192%, 3.129%, and 3.185%, respectively) when the large dataset was fused first into substantially biased values (34.414%, 10.804%, and 18.013%, respectively) when the large dataset is fused last. The estimates of the slope variance (σ^2_{Slope}) also showed somewhat similar patterns of results, but with somewhat different magnitudes when fusing the larger sample dataset last. The variances changed from ignorable bias (1.52%, 1.56%, and 1.89%, respectively) when the large dataset was fused first, to substantially biased values (33.92%) for the small magnitude covariance matrix (Ψ_1) to moderately biased (5.86%) for medium magnitude covariance matrix (Ψ_2) to ignorable (4.53%) for the large magnitude covariance matrix (Ψ_3) when the larger dataset was fused last. A sizable amount of relative bias was again observed when examining the magnitude of the intercept slope covariance value (σ_{IS}) for the Ψ_2 variance-covariance matrix, shifting from ignorable bias (1.84%) when the large dataset was fused first to substantially biased (-10.696%) when the larger dataset was fused last. Table 4b displays the results for when initial informative priors are used in the data fusion process. Here, moderate bias is only observed for the estimates of the intercept variance ($\sigma^2_{Intercept}$) and slope variance (σ^2_{Slope}) for the first covariance matrix condition (Ψ_1) when the large dataset is analyzed last (7.706% and 6.96% respectively).

The findings obtained under the second simulated data design condition once again highlight the importance that the order of fusing the datasets plays in the estimation of parameters when implementing Bayesian Synthesis strategies. Fusing data in which a larger dataset is fused last can produce a substantially biased final posterior distribution when the other sequentially analyzed datasets are much smaller.

It is important to note that these same patterns of results were also observed for the both the third (see Table 5a) and the fourth (see Table 6a) simulated data design conditions using both non-informative and informative priors for the initial dataset in the data fusion process (see Tables 5b and 6b). The third data design condition encompassed a longitudinal study with 10 assessment occasions, measured every 6 months, starting from age 2.5. This data design was meant to reflect small breadth studies that start at an early age range but with large numbers of observations. The fourth data design condition also covered

Table 4a: Data Condition 1 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	0.0002[a]	0.0000	-0.208[a]	0.0080	0.0051[a]	0.0016	0.0051[a]	0.0016
	0.0002[b]	0.0001	-0.232[b]	0.1360	0.0051[b]	0.0016	0.0051[b]	0.0016
	0.0002[c]	0.0001	-0.236[c]	0.1160	0.0051[c]	0.0016	0.0051[c]	0.0016
σ_{IS}	0.0001	-0.0093	–	–	0.0029	0.0196	0.0029	0.0173
	0.0009	-0.0053	1.8400	**-10.696**	0.0172	0.0274	0.0171	0.0269
	0.0045	-0.0012	2.274	-0.592	0.0297	0.0416	0.0293	0.0416
$\beta_{Intercept}$	-0.0019	-0.0010	0.0940	0.0506	0.0151	0.0152	0.0150	0.0152
	-0.0034	-0.0024	0.1694	0.1224	0.0260	0.0256	0.0257	0.0255
	-0.0024	-0.0019	0.1194	0.0930	0.0206	0.0202	0.0204	0.0201
β_{Slope}	0.0001	0.0001	0.0250	0.0350	0.0031	0.0032	0.0031	0.0032
	0.0000	0.0003	0.0006	0.0750	0.0089	0.0089	0.0089	0.0089
	0.0004	0.0002	-0.1010	-0.0540	0.0177	0.0171	0.0177	0.0171
$\sigma^2_{Intercept}$	0.0064	0.0688	3.1920	**34.414**	0.0196	0.1444	0.0186	0.1268
	0.0219	0.0756	3.1286	**10.804**	0.0685	0.1521	0.0649	0.1319
	0.0127	0.0721	3.1850	**18.013**	0.0396	0.1474	0.0375	0.1285
σ^2_{Slope}	0.0002	0.0034	1.5200	**33.920**	0.0010	0.0060	0.0010	0.0050
	0.0016	0.0059	1.5600	**5.860**	0.0090	0.0123	0.0089	0.0108
	0.0076	0.0174	1.8920	4.531	0.0363	0.0383	0.0355	0.0342

Note: Same as Table 3a.

Table 4b: Data Condition 2 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	0.0000[a]	0.0010	-0.0240[a]	1.0080	0.0048[a]	0.0036	0.0048[a]	0.0035
	0.0000[b]	0.0004	0.0440[b]	0.3760	0.0049[b]	0.0025	0.0049[b]	0.0025
	0.0001[c]	0.0003	0.0600[c]	0.3200	0.0048[c]	0.0020	0.0048[c]	0.0020
σ_{IS}	0.0000	-0.0016	−−	−−	0.0031	0.0138	0.0031	0.0137
	0.0009	0.0021	1.7600	4.1680	0.0183	0.0278	0.0182	0.0277
	0.0040	0.0030	2.0120	1.5140	0.0307	0.0378	0.0304	0.0376
$\beta_{Intercept}$	-0.0001	0.0000	0.0036	-0.0006	0.0147	0.0157	0.0147	0.0157
	-0.0005	-0.0003	0.0252	0.0144	0.0255	0.0263	0.0255	0.0263
	-0.0004	0.0001	0.0224	-0.0034	0.0198	0.0219	0.0198	0.0219
β_{Slope}	-0.0002	0.0000	-0.0580	-0.0120	0.0036	0.0035	0.0036	0.0035
	-0.0007	-0.0003	-0.1690	-0.0710	0.0100	0.0097	0.0100	0.0097
	-0.0010	-0.0007	-0.2580	-0.1760	0.0194	0.0196	0.0194	0.0196
$\sigma^2_{Intercept}$	0.0063	0.0154	3.1580	**7.7060**	0.0192	0.0889	0.0181	0.0876
	0.0225	0.0129	3.2171	1.8446	0.0672	0.1236	0.0633	0.1229
	0.0133	0.0173	3.3130	4.3160	0.0385	0.1001	0.0361	0.0985
σ^2_{Slope}	0.0001	0.0007	1.1200	**6.9600**	0.0010	0.0043	0.0010	0.0043
	0.0015	0.0001	1.4920	0.0640	0.0089	0.0114	0.0088	0.0114
	0.0067	0.0032	1.6710	0.7950	0.0376	0.0340	0.0370	0.0339

Note: Same as Table 3a.

10 assessment occasions, measured every six months, but starting instead from age 10. This data design represented the feature of small breadth with a large number of observations that covered a late age range. Given the similarity of these observed bias findings, we focus next on examining the results for just the fifth and sixth simulated data design conditions.

Table 5a: Data Condition 3 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
	0.0003[a]	-0.0001	0.324[a]	-0.0064	0.0054[a]	0.0016	0.0054[a]	0.0016
σ_e^2	0.0003[b]	0.0002	0.284[b]	0.1520	0.0055[b]	0.0016	0.0055[b]	0.0016
	0.0002[c]	0.0001	0.244[c]	0.1280	0.0055[c]	0.0016	0.0055[c]	0.0016
	-0.0003	-0.0078	–	–	0.0031	0.0159	0.0031	0.0138
σ_{IS}	0.0002	-0.0050	0.408	**-10.008**	0.0177	0.0229	0.0177	0.0223
	0.0037	-0.0018	1.846	-0.8820	0.0300	0.0331	0.0297	0.0331
	-0.0017	-0.0008	0.0862	0.0406	0.0143	0.0143	0.0142	0.0143
$\beta_{Intercept}$	-0.0031	-0.0025	0.1526	0.1226	0.0253	0.0253	0.0252	0.0252
	-0.0022	-0.0020	0.1080	0.0992	0.0198	0.0198	0.0197	0.0197
	0.0001	0.0001	0.0200	0.0270	0.0034	0.0036	0.0034	0.0036
β_{Slope}	-0.0001	0.0003	-0.0170	0.0780	0.0091	0.0090	0.0091	0.0090
	-0.0006	-0.0001	-0.1530	-0.0270	0.0178	0.0173	0.0178	0.0173
	0.0071	0.0530	3.5620	**26.518**	0.0195	0.1078	0.0182	0.0938
$\sigma^2_{Intercept}$	0.0212	0.0600	3.0354	**8.5703**	0.0673	0.1161	0.0639	0.0993
	0.0130	0.0563	3.2560	**14.078**	0.0390	0.1117	0.0367	0.0964
	0.0003	0.0029	2.8400	**28.680**	0.0011	0.0050	0.0001	0.0041
σ^2_{Slope}	0.0029	0.0050	2.8960	**5.048**	0.0095	0.0102	0.0090	0.0088
	0.0108	0.0145	1.2700	3.628	0.0371	0.0317	0.0355	0.0282

Note: Same as Table 3a.

The results for the fifth simulated data design condition for the variance-covariance matrices Ψ_1, Ψ_2, and Ψ_3 are presented in Table 7a. This simulated data design condition comprised of observations taken across 3 assessment occasions, measured every six months, starting from age 4. This simulated condition reflected the feature of early age range of development in a small breadth of measurement years. The results presented in Table 7a again correspond to those obtained when a sample size of 1000 observations is incorporated as the first dataset and then as the last dataset while randomly varying the order of the other 5 datasets each with 50 observations.

Table 5b: Data Condition 3 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ^2_e	0.0004[a]	0.0009	0.4360[a]	0.9120	0.0050[a]	0.0032	0.0050[a]	0.0031
	0.0004[b]	0.0003	0.4080[b]	0.2720	0.0051[b]	0.0021	0.0050[b]	0.0021
	0.0005[c]	0.0002	0.4680[c]	0.2120	0.0049[c]	0.0018	0.0049[c]	0.0018
σ_{IS}	0.0000	-0.0006	−−	−−	0.0032	0.0119	0.0032	0.0119
	0.0016	0.0043	3.2080	**8.6160**	0.0186	0.0244	0.0185	0.0240
	0.0063	0.0066	3.1480	3.3060	0.0310	0.0325	0.0304	0.0318
$\beta_{Intercept}$	0.0002	0.0002	-0.0088	-0.0106	0.0140	0.0146	0.0140	0.0146
	-0.0002	-0.0006	0.0078	0.0294	0.0253	0.0255	0.0253	0.0255
	0.0001	-0.0006	-0.0040	0.0276	0.0194	0.0205	0.0194	0.0205
β_{Slope}	-0.0002	0.0000	-0.0450	-0.0090	0.0037	0.0039	0.0037	0.0039
	-0.0003	-0.0004	-0.0830	-0.1000	0.0103	0.0097	0.0103	0.0097
	-0.0006	-0.0013	-0.1570	-0.3240	0.0197	0.0192	0.0197	0.0192
$\sigma^2_{Intercept}$	0.0088	0.0126	4.3800	**6.3160**	0.0208	0.0703	0.0189	0.0692
	0.0298	0.0169	4.2560	2.4177	0.0726	0.1027	0.0662	0.1013
	0.0177	0.0166	4.4180	4.1420	0.0419	0.0815	0.0380	0.0798
σ^2_{Slope}	0.0002	0.0006	1.7600	**6.0800**	0.0011	0.0038	0.0011	0.0037
	0.0030	0.0014	3.0040	1.3720	0.0096	0.0098	0.0091	0.0097
	0.0129	0.0091	3.2320	2.2650	0.0382	0.0306	0.0360	0.0293

Note: Same as Table 3a.

Table 6a: Data Condition 4 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	$0.0004^{\mathbf{a}}$	-0.0002	$0.380^{\mathbf{a}}$	-0.2040	$0.0052^{\mathbf{a}}$	0.0017	$0.0052^{\mathbf{a}}$	0.0017
	$0.0003^{\mathbf{b}}$	0.0000	$0.304^{\mathbf{b}}$	0.0080	$0.0053^{\mathbf{b}}$	0.0016	$0.0053^{\mathbf{b}}$	0.0016
	$0.0003^{\mathbf{c}}$	-0.0001	$0.304^{\mathbf{c}}$	-0.1320	$0.0053^{\mathbf{c}}$	0.0017	$0.0053^{\mathbf{c}}$	0.0017
σ_{IS}	-0.0001	-0.0079	–	–	0.0035	0.0158	0.0035	0.0137
	0.0004	-0.0053	0.832	**-10.544**	0.0179	0.0224	0.0179	0.0218
	0.0048	-0.0022	2.382	-1.1160	0.0291	0.0336	0.0287	0.0335
$\beta_{Intercept}$	-0.0012	-0.0004	0.0596	0.0194	0.0195	0.0191	0.0194	0.0191
	-0.0026	-0.0020	0.1318	0.1020	0.0282	0.0285	0.0281	0.0284
	-0.0016	-0.0015	0.0802	0.0728	0.0231	0.0241	0.0230	0.0240
β_{Slope}	0.0001	0.0001	0.0130	0.0190	0.0035	0.0034	0.0035	0.0034
	0.0000	-0.0003	-0.0050	0.0640	0.0090	0.0090	0.0090	0.0090
	-0.0008	0.0000	-0.1980	0.0040	0.0176	0.0175	0.0175	0.0175
$\sigma^2_{Intercept}$	0.0077	0.0535	3.8260	**26.726**	0.0225	0.1107	0.0212	0.0969
	0.0235	0.0605	3.3566	**8.6389**	0.0677	0.1160	0.0634	0.0989
	0.0144	0.0577	3.6030	**14.424**	0.0406	0.1138	0.0379	0.0980
σ^2_{Slope}	0.0003	0.0024	2.5200	**23.760**	0.0011	0.0042	0.0011	0.0034
	0.0030	0.0044	2.9600	4.384	0.0094	0.0090	0.0089	0.0078
	0.0110	0.0125	2.7420	3.115	0.0355	0.0279	0.0338	0.0250

Note: Same as Table 3a.

Table 6b: Data Condition 4 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	0.0003[a]	0.0002	0.3000[a]	0.2440	0.0051[a]	0.0019	0.0051[a]	0.0019
	0.0004[b]	0.0002	0.3600[b]	0.2040	0.0051[b]	0.0021	0.0051[b]	0.0021
	0.0004[c]	0.0001	0.3520[c]	0.0720	0.0053[c]	0.0018	0.0052[c]	0.0018
σ_{IS}	0.0001	-0.0014	−−	−−	0.0036	0.0113	0.0036	0.0112
	0.0019	0.0025	3.7200	**5.0320**	0.0181	0.0231	0.0180	0.0230
	0.0069	0.0038	3.4380	1.9080	0.0302	0.0317	0.0295	0.0315
$\beta_{Intercept}$	-0.0006	0.0001	0.0292	-0.0038	0.0202	0.0213	0.0202	0.0213
	-0.0010	-0.0006	0.0518	0.0320	0.0291	0.0293	0.0291	0.0293
	-0.0006	-0.0006	0.0302	0.0280	0.0242	0.0264	0.0242	0.0264
β_{Slope}	-0.0001	0.0000	-0.0250	-0.0110	0.0038	0.0041	0.0038	0.0041
	-0.0003	-0.0002	-0.0740	-0.0550	0.0102	0.0101	0.0102	0.0101
	-0.0009	-0.0005	-0.2340	-0.1270	0.0195	0.0197	0.0195	0.0197
$\sigma^2_{Intercept}$	0.0084	0.0162	4.2240	**8.1120**	0.0228	0.0695	0.0212	0.0675
	0.0304	0.0178	4.3463	2.5463	0.0706	0.0990	0.0637	0.0974
	0.0178	0.0189	4.4480	4.7240	0.0418	0.0784	0.0378	0.0761
σ^2_{Slope}	0.0003	0.0007	3.4000	**6.6400**	0.0012	0.0032	0.0011	0.0031
	0.0042	0.0021	4.1560	2.0960	0.0104	0.0092	0.0095	0.0089
	0.0165	0.0099	4.1200	2.4740	0.0399	0.0292	0.0363	0.0275

Note: Same as Table 3a.

In general, the obtained values for the raw bias, accuracy, and efficiency criteria reflect similar results to those observed in the other simulated data design conditions. Biased values were observed when comparing the obtained estimate to the true population values when applying the Bayesian Synthesis approach and regardless of the specified variance-covariance matrix examined. Focusing again on the estimated intercept variance ($\sigma^2_{Intercept}$) for the first dataset condition in Table 7a, the obtained values for the raw bias (.0069, .0212, and .0125), accuracy (.0190, .0643, and .0368), and efficiency (.0177, .0607, .0346) when the large dataset was fused first are negligible. However, when the large dataset is fused last, the estimated intercept variance values increase by a little for the raw bias (.0475, .0546, and .0532), accuracy (.0929, .1059, and .1027), and efficiency (.0798, .0907, .0878).

When examining the relative bias criterion under the three variance-covariance matrixes, some visible biased patterns of results again emerge. In this data design, it can again be seen that the relative bias for the estimated intercept variance ($\sigma^2_{Intercept}$) increases from ignorable sizes (3.472%, 3.033%, and 3.116%, respectively) when the large dataset was fused first into moderately biased and substantially biased values (23.726%, 7.793%, and 13.330%, respectively) when the large dataset was fused last. Interestingly, the estimates of the slope variance (σ^2_{Slope}) showed rather different patterns of results, with sizable magnitudes of relative bias both when fusing the larger sample dataset first and last. In particular, the slope variances displayed substantial bias (18.480%) for covariance matrix (Ψ_1) and ignorable bias (4.320% and 3.574%) for covariance matrixes (Ψ_2 and Ψ_3) when the large dataset was fused first, compared to substantially biased values (18.28%) for the small magnitude covariance matrix (Ψ_1) to moderately biased (6.35%) for medium magnitude covariance matrix (Ψ_2) to ignorable (4.42%) for the large magnitude covariance matrix (Ψ_3) when the larger dataset was fused last. A sizable amount of relative bias was also observed when examining the magnitude of the intercept slope covariance value (σ_{IS}) for the Ψ_2 variance-covariance matrix, shifting from ignorable bias (-2.104%) when the large dataset was fused first to substantially biased (-15.112%) when the larger dataset was fused last. This similar pattern was also observed in Table 7b when initial informative priors were used. Specifically, moderate bias for the intercept variance ($\sigma^2_{Intercept}$) for the first covariance matrix (Ψ_1) condition was observed both when the large sample dataset was fused first and last. In contrast, moderate bias for the estimate of the slope variance (σ^2_{Slope}) was only observed when the large dataset was fused last (6.3855%).

Table 8a presents the results for the final sixth simulated data design condition for the variance-covariance matrices Ψ_1, Ψ_2, and Ψ_3. This simulated data design condition comprised of observations collected across 3 assessment occasions, measured every year, starting from age 11. This simulated condition reflected the feature of a late age range of development in a small breadth of measurement years. The obtained values for the raw bias, accuracy, and efficiency criteria reflect similar results to those observed in other simulated data design conditions. Focusing again on the estimated intercept variance ($\sigma^2_{Intercept}$)

Table 7a: Data Condition 5 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
	0.0001[a]	-0.0005	-0.112[a]	-0.5160	0.0053[a]	0.0032	0.0053[a]	0.0032
σ_e^2	0.0001[b]	-0.0005	-0.052[b]	-0.4640	0.0052[b]	0.0035	0.0052[b]	0.0035
	0.0000[c]	-0.0001	-0.036[c]	-1.0200	0.0052[c]	0.0040	0.0052[c]	0.0039
	-0.0023	-0.0073	−−	−−	0.0049	0.0142	0.0035	0.0121
σ_{IS}	-0.0011	-0.0076	-2.104	**-15.112**	0.0156	0.0197	0.0156	0.0182
	0.0008	-0.0079	0.382	-3.9400	0.0251	0.0262	0.0251	0.0250
	-0.0008	-0.0006	0.0378	0.0312	0.0164	0.0152	0.0163	0.0152
$\beta_{Intercept}$	-0.0011	-0.0010	0.0538	0.4840	0.0263	0.0264	0.0263	0.0264
	-0.0015	-0.0010	0.0764	0.0482	0.0208	0.0214	0.0208	0.0213
	0.0009	0.0009	0.0232	0.2320	0.0062	0.0065	0.0062	0.0064
β_{Slope}	0.0015	-0.0020	0.3810	0.4960	0.0131	0.0132	0.0130	0.0131
	0.0017	0.0028	0.4170	0.7050	0.0207	0.0211	0.0207	0.0209
	0.0069	0.0475	3.4720	**23.726**	0.0190	0.0929	0.0177	0.0798
$\sigma^2_{Intercept}$	0.0212	0.0546	3.0331	**7.7931**	0.0643	0.1059	0.0607	0.0907
	0.0125	0.0532	3.1160	**13.300**	0.0368	0.1027	0.0346	0.0878
	0.0018	0.0018	**18.480**	**18.280**	0.0029	0.0029	0.0023	0.0023
σ^2_{Slope}	0.0043	0.0064	4.3200	**6.353**	0.0100	0.0114	0.0090	0.0095
	0.0143	0.0177	3.5740	4.416	0.0367	0.0344	0.0338	0.0295

Note: Same as Table 3a.

Table 7b: Data Condition 5 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
	0.0008[a]	0.0003	0.8273[a]	0.3052	0.0059[a]	0.0037	0.0059[a]	0.0037
σ_e^2	0.0001[b]	-0.0001	0.1360[b]	-0.0920	0.0046[b]	0.0035	0.0046[b]	0.0035
	0.0001[c]	0.0002	0.0760[c]	0.2000	0.0046[c]	0.0043	0.0046[c]	0.0043
	0.0003	-0.0016	−−	−−	0.0058	0.0110	0.0059	0.0109
σ_{IS}	-0.0007	-0.0012	-1.4160	-2.3200	0.0161	0.0192	0.0161	0.0191
	0.0020	-0.0009	1.0160	-0.4480	0.0257	0.0293	0.0256	0.0293
	0.0016	0.0018	-0.0795	-0.0878	0.0165	0.0155	0.0165	0.0154
$\beta_{Intercept}$	0.0025	0.0026	-0.1234	-0.1288	0.0267	0.0261	0.0266	0.0259
	0.0020	0.0016	-0.1000	-0.0802	0.0211	0.0217	0.0210	0.0216
	-0.0002	-0.0001	-0.0562	-0.0351	0.0068	0.0074	0.0068	0.0074
β_{Slope}	0.0001	0.0001	0.0320	0.0310	0.0140	0.0137	0.0140	0.0137
	0.0007	0.0002	0.1870	0.0600	0.0213	0.0220	0.0213	0.0220
	0.0118	0.0127	**5.8976**	**6.3394**	0.0235	0.0613	0.0204	0.0601
$\sigma^2_{Intercept}$	0.0267	0.0140	3.8080	2.0000	0.0632	0.0824	0.0573	0.0812
	0.0161	0.0144	4.0240	3.5950	0.0368	0.0705	0.0331	0.0690
	0.0000	0.0006	0.0402	**6.3855**	0.0028	0.0022	0.0029	0.0022
σ^2_{Slope}	0.0028	0.0029	2.8080	2.9080	0.0095	0.0099	0.0091	0.0095
	0.0105	0.0091	2.6280	2.2650	0.0361	0.0318	0.0345	0.0305

Note: Same as Table 3a.

in Table 8a, the same patterns of results are evident. For instance, the obtained values for the raw bias (.0038, .0214, and .0251), accuracy (.0552, .0972, and .0613), and efficiency (.0399, .0762, .0559) when the large dataset was fused first are negligible. Similarly, when the large dataset is fused last, the estimated intercept variance values again slightly increase (though are still rather close to zero) for the raw bias (.0145, .0335, and .0206), accuracy (.0357, .0808, and .0529), and efficiency (.0326, .0735, .0487).

When carefully examining the relative bias criterion under the three variance-covariance matrixes, some novel biased results emerge. Specifically, it can be seen in Table 8a that the relative bias for the estimated intercept variance ($\sigma^2_{Intercept}$) ranges from substantial to ignorable to moderately biased (19.020%, 3.054%, and 6.286%, respectively) when the large dataset was fused first, and similarly (7.226%, 4.785%, and 5.162%, respectively) when the large dataset was fused last. The estimates of the slope variance (σ^2_{Slope}) also showed somewhat similar patterns of results, with sizable magnitudes of relative bias both when fusing the larger sample dataset first and last. The slope variances also displayed ranges from substantial bias to ignorable bias (13.20%, 4.63%, and 4.04%, respectively) when the large dataset was fused first, compared to (15.52%, 2.93%, and 1.66%, respectively) when the larger dataset was fused last. A sizable amount of relative bias was also observed when examining the magnitude of the intercept slope covariance value (σ_{IS}) for the Ψ_2 variance-covariance matrix, shifting from ignorable bias (-2.016%) when the large dataset was fused first to substantially biased (-13.76%) when the larger dataset was fused last. Interestingly, this was the only data design condition where moderate bias was observed when the large dataset was analyzed first and informative priors were used to analyze the first dataset. In table 8b, we see that there was moderate bias (7.1478% and 5.5466% respectively) for the estimated intercept ($\sigma^2_{Intercept}$) and slope (σ^2_{Slope}) variance for the first variance-covariance matrix condition (Ψ_1).

These results appear to collectively highlight not only the importance that the order that fusing the datasets can play in the estimation of parameters but also the potential impact that data design characteristics can exert when implementing Bayesian synthesis strategies. It appears that in data design settings with fewer occasions of measurement covering wide range of ages, there can be sizable bias irrespective of whether a larger data set is fused first or last. Additionally, even in instances where there are sufficient assessment settings over a wider age range, the order of the fusing of the data sets can again play a key role. These results would collectively suggest that the order in which datasets are incorporated in the Bayesian Synthesis process do in fact impact the results when one dataset is substantially larger than the rest.

Table 8a: Data Condition 6 Using Initial Non-Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
σ_e^2	0.0009[a]	-0.0003	-0.888[a]	-0.3040	0.0051[a]	0.0032	0.0050[a]	0.0032
	0.0002[b]	-0.0004	-0.024[b]	-0.3920	0.0051[b]	0.0034	0.0051[b]	0.0034
	0.0004[c]	-0.0002	-0.408[c]	-0.2440	0.0052[c]	0.0033	0.0052[c]	0.0033
σ_{IS}	-0.0058	-0.0047	−−	−−	0.0083	0.0084	0.0060	0.0069
	-0.0010	-0.0069	-2.016	**-13.760**	0.0170	0.0168	0.0169	0.0153
	-0.0018	-0.0064	-0.876	-3.2200	0.0266	0.0241	0.0265	0.0232
$\beta_{Intercept}$	-0.0004	-0.0007	-0.0212	0.0352	0.0308	0.0287	0.0308	0.0286
	-0.0006	-0.0010	0.0288	0.0490	0.0423	0.0425	0.0423	0.0425
	-0.0014	-0.0050	0.0696	0.0246	0.0342	0.0369	0.0341	0.0369
β_{Slope}	0.0006	0.0003	0.1440	0.0670	0.0061	0.0045	0.0060	0.0045
	0.0010	-0.0090	0.3520	0.2300	0.0110	0.0105	0.0109	0.0105
	0.0014	0.0019	0.4170	0.4700	0.0189	0.0190	0.0189	0.0189
$\sigma^2_{Intercept}$	0.0038	0.0145	**19.020**	**7.2260**	0.0552	0.0357	0.0399	0.0326
	0.0214	0.0335	3.0543	4.7851	0.0792	0.0808	0.0762	0.0735
	0.0251	0.0206	**6.2860**	**5.1620**	0.0613	0.0529	0.0559	0.0487
σ^2_{slope}	0.0013	0.0016	**13.200**	**15.520**	0.0019	0.0027	0.0013	0.0022
	0.0046	0.0029	4.6320	2.936	0.0094	0.0060	0.0082	0.0052
	0.0162	0.0066	4.0440	1.662	0.0346	0.0187	0.0305	0.0175

Note: Same as Table 3a.

Table 8b: Data Condition 6 Using Initial Informative Priors – Parameter Evaluation Criteria Results

Parameter	Raw Bias		Relative Bias		Accuracy		Efficiency	
	FIRST	LAST	FIRST	LAST	FIRST	LAST	FIRST	LAST
	0.0006[a]	-0.0002	0.6032[a]	-0.1960	0.0059[a]	0.0031	0.0059[a]	0.0031
σ_e^2	0.0003[b]	-0.0005	0.2600[b]	-0.4920	0.0053[b]	0.0031	0.0053[b]	0.0031
	0.0007[c]	-0.0004	0.6960[c]	0.3800	0.0056[c]	0.0030	0.0055[c]	0.0030
	-0.0012	-0.0005	−−	−−	0.0086	0.0078	0.0086	0.0078
σ_{IS}	0.0007	-0.0017	1.3200	-3.3200	0.0181	0.0156	0.0181	0.0155
	0.0041	0.0010	2.0560	0.5020	0.0281	0.0238	0.0278	0.0238
	0.0031	0.0022	-0.1547	-0.1094	0.0361	0.0306	0.0362	0.0305
$\beta_{Intercept}$	0.0034	0.0025	-0.1680	-0.1234	0.0471	0.0430	0.0470	0.0430
	0.0027	0.0001	-0.1326	-0.0050	0.0374	0.0384	0.0373	0.0384
	-0.0002	0.0000	-0.0435	0.0070	0.0065	0.0047	0.0066	0.0047
β_{Slope}	0.0003	0.0002	0.0700	0.0420	0.0111	0.0105	0.0111	0.0105
	0.0008	0.0002	0.2080	0.0610	0.0190	0.0194	0.0190	0.0194
	0.0143	0.0042	**7.1478**	2.0960	0.0523	0.0375	0.0506	0.0373
$\sigma^2_{Intercept}$	0.0151	0.0188	2.1503	2.6880	0.0846	0.0719	0.0832	0.0694
	0.0069	0.0114	1.7160	2.8560	0.0691	0.0498	0.0688	0.0484
	0.0006	0.0002	**5.5466**	2.2000	0.0016	0.0020	0.0015	0.0020
σ^2_{Slope}	0.0030	0.0012	3.0160	1.1520	0.0091	0.0061	0.0086	0.0060
	0.0112	0.0035	2.8040	0.8810	0.0355	0.0192	0.0336	0.0189

Note: Same as Table 3a.

4 Discussion

Bayesian estimation operates by using prior information about the characteristics of parameters and the conditional likelihood of the data given the model parameters to arrive at a posterior distribution. The Bayesian Synthesis approach is based on this Bayesian estimation framework in which information obtained from one dataset serves to provide prior information for the analysis of the next dataset and this process continues sequentially until a single posterior distribution is created using all available datasets. While the benefits of using fused datasets have been repeatedly demonstrated in the literature (e.g., Curran & Hussong, 2009; Du et al., 2020; Hofer & Piccinin, 2009; Marcoulides, 2017b; Marcoulides & Grimm, 2017), and the estimates computed via a sequentially obtained final posterior distribution like those in the Bayesian Synthesis approach have also been shown to effectively aid in the accuracy of the estimation process (Du et al., 2020; Marcoulides, 2017b), what had not be determined was whether the order in which the data are sequentially analyzed has an impact on the obtained results.

The commonly accepted view in Bayesian estimation is that the order in which the data are analyzed should not be a concern due to the exchangeability assumption (de Finetti, 1972, 1974). Nevertheless, because Bayesian Synthesis utilizes point summary estimates of the posterior distributions instead of the full posterior distribution as required in standard Bayesian estimation, it possible that using point summary estimates of the posterior distributions may conceivably introduce some bias in the parameter estimates. Although past research has confirmed that the order of analysis does not meaningfully impact the final data fusion results obtained via Bayesian Synthesis (Marcoulides, 2017b), these conclusions were determined on the basis of analyzing datasets that were from similarly-sized and large samples. What was unresolved in the literature is whether exchangeability matters when the datasets being fused have substantially different sample sizes, as regularly occurs in empirical settings. Does beginning or ending the Bayesian Synthesis approach with the analysis of a large dataset produce a biased final posterior distribution when the other sequentially analyzed datasets are much smaller? This study examined via simulation the impact that the ordering of datasets might have on parameter estimates obtained when making use of the Bayesian Synthesis process in such data fusion design settings.

The results of the simulation study collectively highlighted the importance that the ordering of datasets can have on the estimation of growth model parameters when using the Bayesian Synthesis process. When the datasets being fused are of markedly different and much smaller sizes, ending the fusion and Bayesian estimation based on a large dataset produces a substantially biased (according to the measure of relative bias) final posterior distribution, particularly for the intercept and slope variance. Dissimilar longitudinal data design characteristics were also sometimes found to produce substantially biased final posterior distribution when implementing Bayesian Synthesis strategies. In longitudinal data design settings with fewer occasions of measurement and covering varying ranges of ages, sizeable biased estimates were observed irrespective of whether a larger data set is fused first or last. Even in instances where there were numerous assessment occasions over a wider age range, the order of the fusing of the data sets still played a key role in the estimation, primarily with more sizable bias present when the large dataset was fused last. The results revealed that the order datasets of differing size are incorporated into the Bayesian Synthesis process along with the data design characteristics can impact the resulting parameter estimates and clearly calls into question the previously accepted notion of exchangeability of Bayesian estimation within the Bayesian Synthesis process.

Researchers planning on using the Bayesian Synthesis approach to data fusion should therefore be very careful how they elect to begin and end planned data fusion activities, especially in instances that involve the analyses of substantially large datasets among other much smaller datasets. Because Bayesian Synthesis uses point summary estimates from the analysis of one dataset as priors for the analysis of the next dataset, it is likely that this can introduce some bias in the Bayesian estimation. One explanation for this bias is that when the small datasets are being incorporated first, the informative priors that result from these small datasets are less reliable (contain sizable bias) and that this bias is then inevitably carried over to subsequent samples, resulting in a more biased final posterior distribution. Although it is commonly accepted that sample size can play an important role in the estimation of parameters, it is unclear in this context how much smaller the datasets can be relative to the other datasets. In this study, it was unmistakably determined that fusing a large dataset with smaller ones biased many of the parameter estimates provided by Bayesian Synthesis (particularly when measured by the relative bias criterion). But it is unclear what the ideal sample size needs to be in order to be used in the approach and ensure sufficiently stable parameter estimates. The current study fixed some of the data design characteristics in order to keep the scope of the work manageable. Given that a major benefit of Bayesian Synthesis is that data from multiple sources can be analyzed to obtain estimates of overall effects, examining other data size conditions under which this approach does not operate well is a natural extension to the current study. There is overall agreement among researchers that larger samples provide more stable estimates, but must all the fused datasets meet this requirement in order for Bayesian estimation exchangeability to hold? Although the current results indicated that exchangeability did not always hold in the examined data design scenarios involving growth curve models where there were differences in the size of the samples, the number of measurement occasions, time of first assessment and between assessments, as well as magnitude of the intercept and slope variances and covariances, it is of course possible that when modeling other statistical paradigms that different results may be observed. Empirical applications of the Bayesian Synthesis approach must make certain that the data fusion activities will provide researchers with unbiased parameter estimates.

Without doubt prior specification may be the largest advantage, yet potentially the greatest drawback, of implementing Bayesian methods in Bayesian

Synthesis. Priors enable researchers to include information from different data sources in a systematic manner. While it is recognized that imposing informative priors improves parameter estimates, especially with small sample sizes (Depaoli, 2014; Little, 2006), given that the true prior distribution is unknown in practice, researchers must be cautious about the impact that inaccurate priors have on parameter estimation in Bayesian Synthesis (Marcoulides, 2018). We also examined the impact of the order of incorporation in the Bayesian Synthesis process using initial informative data-dependent priors to analyze the first dataset in the data fusion process. Across the various design conditions, moderate and substantial bias was primarily found when the large dataset was analyzed last. These results are consistent with those found when using initial non-informative or diffuse priors to analyze the first dataset in the data fusion process. Future research studies should therefore expand further on our findings and examine additional data design conditions and settings.

The process of sequentially updating information to arrive at conclusions undeniably has a substantiated place in data analyses and Bayesian Synthesis can play a key role in helping researchers address questions not always achievable with a single study. Although additional research needs to be done regarding when Bayesian Synthesis is most useful and when it might prove to be problematic, the foundations for the continued use of this data fusion process are evident. We caution researchers to remain mindful of the limitations identified in this study when integrating data from different sources.

References

Asparouhov, T., & Muthén, B. (2010). Bayesian analysis using mplus : Technical implementation..

Bandalos, D. L., & Gagne, P. (2012). Simulation methods in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 92–110). New York, NY: Guildford Press.

Bandalos, D. L., & Leite, W. L. (2013). Use of Monte Carlo studies in structural equation modeling research. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course (2nd ed)* (pp. 564–666). Greenwich, CT.

Bhattacharya, B., & Saha, B. (2015). Community model: A new data fusion filter paradigm. *American Journal of Advanced Computing*, *2*, 25–31. doi: https://doi.org/10.15864/ajac.1303

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple datasets. *Psychological Methods*, *14*, 81–100. doi: https://doi.org/10.1037/a0015914

de Finetti, B. (1972). *Probability, induction, and statistics.* New York, NY: John Wiley & Sons.

de Finetti, B. (1974). *Theory of probability (Vol. I and Vol. II).* New York, NY: John Wiley & Sons.

Depaoli, S. (2014). The impact of inaccurate “informative” priors for growth parameters in Bayesian growth mixture modeling. *Structural Equation Modeling*, *21*, 239–252. doi: https://doi.org/10.1080/10705511.2014.882686

Du, H., Bradbury, T. N., Lavner, J. A., Meltzer, A. L., McNulty, J. K., Neff, L. A., & Karney, B. R. (2020). A comparison of Bayesian synthesis approaches for studies comparing two means: A tutorial. *Research Synthesis Methods*, *11*, 36–65. doi: https://doi.org/10.1002/jrsm.1365

Finch, W. H., & Miller, J. E. (2019). The use of incorrect informative priors in the estimation of MIMIC model parameters with small sample sizes. *Structural Equation Modeling*, *26*, 497–508. doi: https://doi.org/10.1080/10705511.2018.1553111

Fujimoto, K. A., Gordon, R. A., Peng, F., & Hofer, K. G. (2018). Examining the category functioning of the ECERS-R across eight data sets. *AERA Open*, *4*, 1–16. doi: https://doi.org/10.1177/2332858418758299

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis (3rd Edition).* CRC.

Hallquist, M., & Wiley, J. (2014). *Mplusautomation: Automating Mplus model estimation and interpretation* (Vol. 06-3.).

Hofer, S. M., & Piccinin, A. M. (2009). Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*, *14*, 150–164. doi: https://doi.org/10.1037/a0015566

Johnson, R. M., & Guttmannova, K. (2019, jan). Marijuana use among adolescents and emerging adults in the midst of policy change: Introduction to the special issue. *Prevention Science*, *20*(2), 179–184. doi: https://doi.org/10.1007/s11121-019-0989-7

Little, R. J. (2006). Calibrated Bayes: A Bayes/frequentist roadmap. *The American Statistician*, *60*, 213–223. doi: https://doi.org/10.1198/000313006x117837

Liu, H., Zhang, Z., & Grimm, K. J. (2016). Comparison of inverse Wishart and separation-strategy priors for Bayesian estimation of covariance parameter matrix in growth curve analysis. *Structural Equation Modeling*, *23*, 354–367. doi: https://doi.org/10.1080/10705511.2015.1057285

Marcoulides, K. M. (2017a). *A Bayesian synthesis approach to data fusion using augmented data-dependent priors (Doctoral dissertation)* (Unpublished doctoral dissertation). (Retrieved from ProQuest Dissertations & Theses Global Database. (Accession No. 10276332).)

Marcoulides, K. M. (2017b). A Bayesian synthesis approach to data fusion using data-dependent priors. *Multivariate Behavioral Research*, *52*(1), 111–112. doi: https://doi.org/10.1080/00273171.2016.1263927

Marcoulides, K. M. (2018). Careful with those priors: A note on Bayesian estimation in two-parameter logistic item response theory models. *Measurement: Interdisciplinary Research and Perspectives*, *16*(2), 92–99. doi: https://doi.org/10.1080/15366367.2018.1437305

Marcoulides, K. M., & Grimm, K. J. (2017). Data integration approaches to longitudinal growth modeling. *Educational and Psychological Measurement*, *77*, 971–989. doi: https://doi.org/10.1177/0013164416664117

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? what does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. doi: https://doi.org/10.1037/a0039400

McNeish, D. (2016, jun). On using bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(5), 750–773. doi: https://doi.org/10.1080/10705511.2016.1186549

Miocevic, M., Levy, R., & Savord, A. (2020). The role of exhangeability in sequential updating of findings from small studies and the challenges of identifying exchangeable data sets. In . M. M. R. van de Schoot (Ed.), *Small sample size solutions: A guide for applied researchers and practitioners* (pp. 13–29). London, UK: Routledge.

Muthen, L. K., & Muthen, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*, 599–620. doi: https://doi.org/10.1207/s15328007sem0904_8

Muthen, L. K., & Muthen, B. O. (2017). *Mplus user's guide (8th edition)*. Los Angeles, CA: Muthen & Muthen.

Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., & Chen, F. (2001, apr). Monte carlo experiments: Design and implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(2), 287–312. doi: https://doi.org/10.1207/s15328007sem0802_7

Preston, K. S. J., Gottfried, A. W., Park, J. J., Manapat, P. D., Gottfried, A. E., & Oliver, P. H. (2018). Simultaneous linking of cross-informant and longitudinal data involving positive family relationships. *Educational and Psychological Measurement*, *78*, 409–429. doi: https://doi.org/10.1177/0013164417690198

R Development Core Team. (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Saris, W., & Satorra, A. (2018). The pooled data approach for the estimation of split-ballot multitrait-multimethod experiments. *Structural Equation Modeling*, *25*, 659–672. doi: https://doi.org/10.1080/10705511.2018.1431543

Shi, D., & Tong, X. (2017). The impact of prior information on Bayesian latent basis growth model estimation. *SAGE Open*, *7*. doi: https://doi.org/10.1177/2158244017727039

Shi, D., & Tong, X. (2018). Bayesian robust two-stage estimation with nonnormal missing data. *Multivariate Behavioral Research*, *53*, 127–127. doi: https://doi.org/10.1080/00273171.2017.1404894