Multivariate Meta-Analysis of Vector Autoregressive Model Coefficients

Ivan Jacob Agaloos Pesigan¹, Michael A. Russell^{1, 2}, and Sy-Miin Chow³

¹Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University

²Department of Biobehavioral Health, The Pennsylvania State University

³Department of Human Development and Family Studies, The Pennsylvania State

University

Abstract

TODO

Keywords: vector autoregressive model, multilevel model, random-effects meta-analysis, mixed-effects meta-analysis, distal outcomes

Recent years have witnessed a surge of interest in intensive longitudinal data (ILD) and dynamic modeling techniques to examine within-person processes across time. Among these, discrete-time vector autoregressive (VAR) models have gained traction for modeling temporal dynamics in psychological systems. While powerful, such models often grapple with the challenge of balancing individual-level precision with generalizable population-level inference—particularly when dealing with the hierarchical structure of repeated measures nested within individuals.

Traditional approaches often apply either fully pooled multilevel models or separate idiographic models per person. However, each comes with trade-offs. Fully pooled models may obscure person-specific dynamics due to shrinkage toward the group mean, whereas single-subject analyses struggle to generalize due to the non-ergodic nature of psychological processes (Hamaker et al., 2005; Molenaar, 2004). A growing literature has therefore advocated for hybrid methods that incorporate both within-person dynamics and between-person variability (Bringmann et al., 2013; Castro-Schilo & Ferrer, 2013).

This research was made possible by the Prevention and Methodology Training Program (PAMT) funded by a T32 training grant (T32 DA017629 MPIs: J. Maggs & S. Lanza) from the National Institute on Drug Abuse (NIDA).

Computations for this research were performed on the Pennsylvania State University's Institute for Computational and Data Sciences' Roar supercomputer.

Correspondence concerning this article should be addressed to Ivan Jacob Agaloos Pesigan, Edna Bennett Pierce Prevention Research Center, College of Health and Human Development, The Pennsylvania State University, 320 Biobehavioral Health Building, University Park, PA 16802 or by email (ijapesigan@psu.edu).

Building on this foundation, we propose a two-step meta-analytic framework to model multilevel dynamic systems using discrete-time VAR models. In the first step, idiographic VAR parameters are estimated at the individual level. In the second step, these estimates are aggregated and analyzed using a meta-analytic structural equation modeling (MASEM) approach or a two-stage random effects meta-analysis (Cheung, 2008; Lee & Gates, 2023). This allows for population-level inference while properly accounting for the uncertainty associated with person-specific estimates.

A particularly promising extension of this framework is the ability to incorporate distal outcomes—long-term or external variables predicted by dynamic features of individuals' intraindividual processes. By treating estimated VAR parameters as predictors in a second-stage model, researchers can examine how individual differences in temporal dynamics (e.g., affective inertia, cross-lagged coupling) relate to later outcomes such as well-being, relapse, or performance. This approach supports hypothesis-driven investigation of dynamic processes as mechanisms or risk markers, while preserving the statistical rigor of properly propagating uncertainty from step one.

This two-step framework offers several advantages. First, it provides a principled way to quantify uncertainty at both the within- and between-person levels. Individual parameter estimates are not treated as fixed values but are accompanied by standard errors, which are incorporated into the meta-analytic step. Second, we show through simulation that failure to account for estimation uncertainty in the first step leads to biased or overconfident group-level inferences—distorting estimated population-level means and variances and obscuring heterogeneity. Third, the framework facilitates exploration of moderators that explain variability in dynamic parameters across individuals. Fourth, the approach naturally supports the inclusion of distal outcomes in the second-stage model, allowing researchers to examine how individual differences in dynamic processes (e.g., affective inertia, cross-lag coupling) predict later outcomes such as well-being, relapse, or achievement. Together, these features make the two-step framework a flexible and powerful tool for drawing generalizable inferences about intraindividual dynamics.

In sum, this paper contributes a general, modular framework for modeling multilevel time series data that bridges idiographic and nomothetic traditions in psychology. By explicitly modeling and integrating uncertainty across levels of analysis, and allowing for the inclusion of distal outcomes, our approach improves statistical inference and fosters more robust generalizations about dynamic psychological processes and their long-term consequences.

In meta-analysis, the unit of analysis is typically the study, where correlations or standardized mean differences from each study are pooled. However, in this case, the unit of analysis is the individual, with parameters estimated for each individual being pooled. In the first stage, parameters for a vector autoregressive model (discrete or continuous) are estimated for each individual, denoted as \mathbf{y}_i , with the corresponding sampling variance-covariance matrix \mathbf{V}_i^2 for individual i. In the second stage, a multivariate meta-analysis is performed to pool the estimates obtained in the first stage. Additionally, we can estimate the effects of time-invariant covariates on the within-individual level estimates from stage one, as well as the effects of the within-level estimates from stage one on time-invariant distal outcomes.

Multivariate Mixed-Effects Meta-Analysis as a Structural Equation Model

The multivariate mixed-effects meta-analysis as a structural equation model is given by

$$\mathbf{y}_i = \alpha + \beta \mathbf{x}_i + \mathbf{V}_i^2 \tag{1}$$

where \mathbf{y}_i is a vector of effect sizes for individual i, $\boldsymbol{\alpha}$ is the vector of average effect sizes without the effect of the covariate vector \mathbf{x}_i , $\boldsymbol{\beta}$ is the matrix of the effect of the covariate vector \mathbf{x}_i on \mathbf{y}_i , and \mathbf{V}_i^2 is a symmetric matrix representing the sampling covariance matrix of \mathbf{y}_i , which is also the within-individual covariance matrix for individual i. In meta-analysis parlance, \mathbf{V}_i^2 is also the weight matrix for individual i. The expected mean vector and covariance matrix of \mathbf{y}_i are given by

$$\mu_{\mathbf{y}_i}(\boldsymbol{\theta}) = \boldsymbol{\alpha} + \boldsymbol{\beta} \mathbf{x}_i, \tag{2}$$

and

$$\Sigma_{\mathbf{V}_i}(\boldsymbol{\theta}) = \boldsymbol{\tau}^2 + \mathbf{V}_i^2,\tag{3}$$

respectively, where $\boldsymbol{\tau}^2$ is the between-individual covariance matrix, and $\boldsymbol{\theta}$ is the parameter vector given by

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\alpha} \\ \operatorname{Vec}(\boldsymbol{\beta}) \\ \operatorname{Vech}(\boldsymbol{\tau}^2) \end{pmatrix}, \tag{4}$$

where $\text{Vec}(\cdot)$ converts a matrix into a vector column-wise, and $\text{Vech}(\cdot)$ converts the unique elements of a symmetric matrix into a vector column-wise.

Notice in Equation 3 that the total variability is a combination of the between-individual variance \mathbf{V}_{i}^{2} . In meta-analysis, it is customary to quantify the heterogeneity of effect sizes using these two sources of variability. A popular index for this purpose is \mathbf{I}^{2} (Higgins & Thompson, 2002), which is given by

$$\mathbf{I}^2 = \frac{\hat{\boldsymbol{\tau}}^2}{\hat{\boldsymbol{\tau}}^2 + \tilde{\mathbf{V}}^2} \tag{5}$$

where $\hat{\tau}^2$ is the sample estimate of τ^2 , and $\tilde{\mathbf{V}}^2$ is the typical within-individual sampling variance for n individuals, given by

$$\tilde{\mathbf{V}}^2 = \frac{(n-1)\sum_{i=1}^n \frac{1}{\mathbf{V}_i^2}}{\left(\sum_{i=1}^n \frac{1}{\mathbf{V}_i^2}\right)^2 - \left(\sum_{i=1}^n \frac{1}{\left(\mathbf{V}_i^2\right)^2}\right)}.$$
 (6)

 \mathbf{I}^2 is interpreted as the proportion of the total variation in effect sizes that is due to between-individual heterogeneity.

If β is a null matrix, the model simplifies to a random-effects meta-analysis model. Furthermore, if both β and τ^2 are null matrices, the model simplifies to a fixed-effect meta-analysis model. For more on this parameterization, see Cheung (2008, 2013). For a general perspective on meta-analysis as a structural equation model, see Cheung (2015).

Multivariate Mixed-Effects Meta-Analysis as a Structural Equation Model with Distal Outcomes

We extend the multivariate mixed-effects meta-analysis as a structural equation model by including a vector of distal outcomes. Since the meta-analysis model is essentially a structural equation model, adding a vector of distal outcomes is straightforward. In this parameterization, the vector of effect sizes \mathbf{y}_i is used to predict a vector of distal outcomes \mathbf{z}_i , given by

$$\mathbf{z}_{i} = \boldsymbol{\kappa} + \mathbf{\Gamma} \mathbf{y}_{i} + \boldsymbol{\varepsilon}_{i}, \quad \boldsymbol{\varepsilon}_{i} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Psi}^{2}\right)$$
 (7)

where κ is a vector of the values of \mathbf{z}_i without the effect of \mathbf{y}_i , Γ is a matrix of the effects of \mathbf{y}_i on \mathbf{z}_i , and ε_i is the vector of residuals. We augment the parameter vector to include the new parameters as follows:

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\kappa} \\ \operatorname{Vec} (\boldsymbol{\Gamma}) \\ \operatorname{Vech} (\boldsymbol{\Psi}^2) \\ \boldsymbol{\alpha} \\ \operatorname{Vec} (\boldsymbol{\beta}) \\ \operatorname{Vech} (\boldsymbol{\tau}^2) \end{pmatrix}. \tag{8}$$

The expected mean vector and covariance matrix of \mathbf{z}_i are given by

$$\mu_{\mathbf{z}_{i}}(\boldsymbol{\theta}) = \kappa + \Gamma (\alpha + \beta \mathbf{x}_{i}) = \kappa + \Gamma \mu_{\mathbf{v}_{i}}(\boldsymbol{\theta}), \tag{9}$$

and

$$\Sigma_{\mathbf{z}_{i}}(\boldsymbol{\theta}) = \Psi^{2} + \Gamma\left(\boldsymbol{\tau}^{2} + \mathbf{V}_{i}^{2}\right)\Gamma' = \Psi^{2} + \Gamma\left(\Sigma_{\mathbf{y}_{i}}(\boldsymbol{\theta})\right)\Gamma'. \tag{10}$$

respectively.

To get the overall expected mean vector and covariance matrix, we stack \mathbf{z}_i and \mathbf{y}_i into the vector \mathbf{a}_i with as follows:

$$\mathbf{a}_i = \begin{pmatrix} \mathbf{z}_i \\ \mathbf{y}_i \end{pmatrix}. \tag{11}$$

The expected mean vector of \mathbf{a}_i can be obtained by stacking Equations 2 and 9 as follows:

$$\mu_{i}(\boldsymbol{\theta}) = \begin{pmatrix} \kappa + \Gamma(\alpha + \beta \mathbf{x}_{i}) \\ \alpha + \beta \mathbf{x}_{i} \end{pmatrix} = \begin{pmatrix} \mu_{\mathbf{z}_{i}}(\boldsymbol{\theta}) \\ \mu_{\mathbf{y}_{i}}(\boldsymbol{\theta}) \end{pmatrix}.$$
(12)

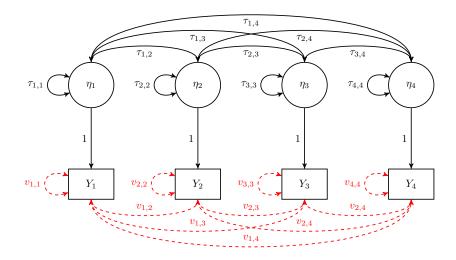
For the expected covariance matrix, the diagonal elements of the stacked covariance matrix are given by Equations 3 and 10. The lower off-diagonal elements are given by $(\tau^2 + \mathbf{V}_i^2) \Gamma'$. Combining these elements, we arrive at

$$\Sigma_{i}\left(\boldsymbol{\theta}\right) = \left(\begin{array}{c|c} \boldsymbol{\Psi}^{2} + \boldsymbol{\Gamma}\left(\boldsymbol{\tau}^{2} + \mathbf{V}_{i}^{2}\right)\boldsymbol{\Gamma}' & \boldsymbol{\Gamma}\left(\boldsymbol{\tau}^{2} + \mathbf{V}_{i}^{2}\right) \\ \hline \left(\boldsymbol{\tau}^{2} + \mathbf{V}_{i}^{2}\right)\boldsymbol{\Gamma}' & \boldsymbol{\tau}^{2} + \mathbf{V}_{i}^{2} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{\mathbf{z}_{i}}\left(\boldsymbol{\theta}\right) & \boldsymbol{\Gamma}\boldsymbol{\Sigma}_{\mathbf{y}_{i}}\left(\boldsymbol{\theta}\right) \\ \hline \boldsymbol{\Sigma}_{\mathbf{y}_{i}}\left(\boldsymbol{\theta}\right)\boldsymbol{\Gamma}' & \boldsymbol{\Sigma}_{\mathbf{y}_{i}}\left(\boldsymbol{\theta}\right) \end{array}\right). \tag{13}$$

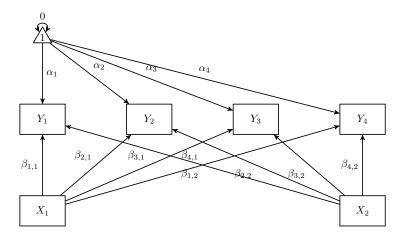
Figure 1

Multivariate Mixed-Effects Meta-Analysis as a Structural Equation Model with Distal Outcomes

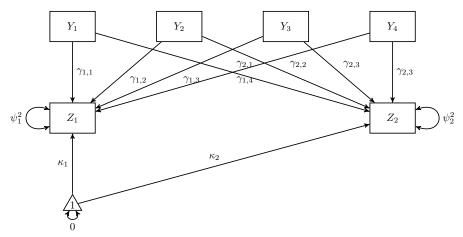
(a) Between- and Within-Level Variability



(b) Covariate Model



(c) Distal Outcome Model



Note: The Panels 1a, 1b, and 1c are estimated simultaneously in one model. They are separated here for ease of presentation. $v_{1,1}$ to $v_{4,4}$ are elements of the \mathbf{V}_i^2 matrix that are fixed per individual i and are specified as definition variables. The rest of the parameters are estimated.

Model Specification and Estimation

The parameter vector $\boldsymbol{\theta}$ can be estimated using normal theory maximum likelihood, based on the expected mean vector and covariance matrix given in Equations 12 and 13, respectively. The log-likelihood function is given by

$$\ell(\boldsymbol{\theta}; \mathbf{a}_i) = -\frac{1}{2} \left(p \log (2\pi) + \log |\boldsymbol{\Sigma}_i(\boldsymbol{\theta})| + (\mathbf{a}_i - \boldsymbol{\mu}_i(\boldsymbol{\theta}))' (\boldsymbol{\Sigma}_i(\boldsymbol{\theta}))^{-1} (\mathbf{a}_i - \boldsymbol{\mu}_i(\boldsymbol{\theta})) \right)$$
(14)

where p represents the number of elements in \mathbf{a}_i , $|\cdot|$ the determinant of a matrix, $(\cdot)'$ the matrix transpose, and $(\cdot)^{-1}$ the matrix inverse. An important consideration in specifying the model is that the \mathbf{V}_i^2 matrix is unique to each individual but is treated as fixed values in the estimation. This can be achieved by treating the elements of \mathbf{V}_i^2 as definition variables. Definition variables are observed variables that modify the statistical model for each individual case. Essentially, these variables from the raw data vectors are used to tailor the statistical model to the specific individual (Cheung, 2013; Mehta & Neale, 2005). This feature is readily available in the OpenMx package (Neale et al., 2015) within the R statistical environment (R Core Team, 2025). The metaVAR package simplifies model specification and estimation by utilizing the OpenMx package under the hood.

(Castro-Schilo & Ferrer, 2013)

(Lee & Gates, 2023)

$$\mathbf{y}_{i,t} = \boldsymbol{\nu} + \boldsymbol{\Lambda} \boldsymbol{\eta}_{i,t} + \boldsymbol{\varepsilon}_{i,t}, \qquad \boldsymbol{\varepsilon}_{i,t} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Theta})$$
 (15)

$$\boldsymbol{\eta}_{i,t} = \boldsymbol{\alpha}_i + \boldsymbol{\beta}_i \boldsymbol{\eta}_{i,t-1} + \boldsymbol{\zeta}_{i,t}, \qquad \boldsymbol{\zeta}_{i,t} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Psi}\right)$$
(16)

Simulation factors from (Bringmann et al., 2013)

Sample size $n = \{20, 129, 500\}$

Measurement occasion $m = \{20, 60, 500\}$

$$\boldsymbol{\nu} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\Theta} = \begin{pmatrix} 0.20 & 0.00 \\ 0.00 & 0.20 \end{pmatrix}.$$

Intercept

$$\alpha_i = \mu_{\alpha} + v_{\alpha_i}, \qquad v_{\alpha_i} \sim \mathcal{N}(0, \Sigma_{\alpha})$$
 (17)

$$\mu_{\alpha} = \begin{pmatrix} 2.87 \\ 2.04 \end{pmatrix}, \Sigma_{\alpha} = \begin{pmatrix} 1.20000 & 0.45957 \\ 0.45957 & 1.10000 \end{pmatrix}$$
(18)

Matrix of lagged coefficients

$$\Gamma_i = \text{Vec}\left(\beta_i\right) \tag{19}$$

$$\Gamma_i = \mu_{\Gamma} + \upsilon_{\Gamma_i}, \qquad \upsilon_{\Gamma_i} \sim \mathcal{N}(0, \Sigma_{\Gamma})$$
 (20)

$$\boldsymbol{\mu}_{\Gamma} = \begin{pmatrix} 0.28 \\ -0.04 \\ -0.05 \\ 0.26 \end{pmatrix}, \boldsymbol{\Sigma}_{\Gamma} = \begin{pmatrix} 0.01690 & 0.00468 & 0.00146 & 0.00832 \\ 0.00468 & 0.00810 & 0.00101 & 0.00576 \\ 0.00146 & 0.00101 & 0.00078 & 0.00179 \\ 0.00832 & 0.00576 & 0.00179 & 0.02560 \end{pmatrix}$$
(21)

$$\Psi = \begin{pmatrix}
1.30000 & 0.56963 \\
0.56963 & 1.56000
\end{pmatrix}$$
(22)

```
model
#> $p
#> [1] 2
#>
#> $k
#> [1] 2
#> $mu0
#> [1] 3.814355 2.576348
#>
#> $sigma0
#> [,1] [,2]
#> [1,] 1.3978842 0.5782369
#> [2,] 0.5782369 1.6636513
#> $sigma0_1
#> [,1] [,2]
#> [1,] 1.1823215 0.000000
#> [2,] 0.4890691 1.193509
#>
#> $tau_sigma0
#> [,1] [,2]
#> [1,] 0.8354876 -0.2903913
#> [2,] -0.2903913 0.7020191
#>
#> $alpha_mu
#> [1] 2.87 2.04
#>
#> $alpha_sigma
#> [,1] [,2]
#> [1,] 1.200000 0.459565
#> [2,] 0.459565 1.100000
#>
#> $alpha_sigma_l
```

```
#> [,1] [,2]
#> [1,] 1.0954451 0.0000000
#> [2,] 0.4195235 0.9612492
#> $beta_mu
#> [,1] [,2]
#> [1,] 0.280 -0.048
#> [2,] -0.035 0.260
#>
#> $beta_sigma
#> [,1] [,2] [,3] [,4]
#> [1,] 0.016900 0.004680 0.001456 0.008320
#> [2,] 0.004680 0.008100 0.001008 0.005760
#> [3.] 0.001456 0.001008 0.000784 0.001792
#> [4,] 0.008320 0.005760 0.001792 0.025600
#> $beta_sigma_l
#> [,1] [,2] [,3] [,4]
#> [1,] 0.1300 0.000000000 0.00000000 0.0000000
#> [2,] 0.0360 0.082486363 0.00000000 0.0000000
#> [3,] 0.0112 0.007332121 0.02459268 0.0000000
#> [4,] 0.0640 0.041897835 0.03122880 0.1370158
#>
#> $psi
#> [,1] [,2]
#> [1,] 1.3000000 0.5696315
#> [2,] 0.5696315 1.5600000
#>
#> $psi l
#> [,1] [,2]
#> [1,] 1.1401754 0.000000
#> [2,] 0.4995998 1.144727
#>
#> $nu
#> [1] 0 0
#> $lambda
#> [,1] [,2]
#> [1,] 1 0
#> [2,] 0 1
#>
#> $theta
#> [,1] [,2]
#> [1,] 0.2 0.0
```

```
#> [2,] 0.0 0.2
#>
#> $theta_1
#>
             [,1]
                       [,2]
#> [1,] 0.4472136 0.0000000
#> [2,] 0.0000000 0.4472136
#>
#> $beta0
        0.280 -0.035 -0.048 0.260 2.870 2.040
#> [1]
#>
#> $tau_sqr
            [,1]
                               [,3]
                                        [,4]
                                                 [,5]
                                                           [,6]
#>
                     [,2]
#> [1,] 0.016900 0.004680 0.001456 0.008320 0.000000 0.000000
#> [2,] 0.004680 0.008100 0.001008 0.005760 0.000000 0.000000
#> [3,] 0.001456 0.001008 0.000784 0.001792 0.000000 0.000000
#> [4,] 0.008320 0.005760 0.001792 0.025600 0.000000 0.000000
#> [5,] 0.000000 0.000000 0.000000 0.000000 1.200000 0.459565
#> [6,] 0.000000 0.000000 0.000000 0.000000 0.459565 1.100000
#>
#> $parameter
#>
   [1]
         0.280000 -0.035000 -0.048000
                                        0.260000
                                                  2.870000
                                                            2.040000
                                                                      0.016900
   [8]
         0.004680
                   0.001456
                             0.008320
                                        0.000000
                                                  0.000000
                                                            0.008100
                                                                      0.001008
#>
                                        0.000784
#> [15]
         0.005760
                   0.000000
                             0.000000
                                                  0.001792
                                                            0.000000
                                                                      0.000000
#> [22]
        0.025600 0.000000 0.000000
                                       1.200000 0.459565
                                                            1.100000
```

References

- Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D., & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, 8(4). https://doi.org/10.1371/journal.pone.0060188
- Castro-Schilo, L., & Ferrer, E. (2013). Comparison of nomothetic versus idiographic-oriented methods for making predictions about distal outcomes from time series data. *Multivariate Behavioral Research*, 48(2), 175–207. https://doi.org/10.1080/00273171. 2012.736042
- Cheung, M. W.-L. (2008). A model for integrating fixed-, random-, and mixed-effects metaanalyses into structural equation modeling. *Psychological Methods*, 13(3), 182–202. https://doi.org/10.1037/a0013163
- Cheung, M. W.-L. (2013). Multivariate meta-analysis as structural equation models. Structural Equation Modeling: A Multidisciplinary Journal, 20(3), 429–454. https://doi.org/10.1080/10705511.2013.797827
- Cheung, M. W.-L. (2015). *Meta-analysis: A structural equation modeling approach*. Wiley. https://doi.org/10.1002/9781118957813

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, 40(2), 207–233. https://doi.org/10.1207/s15327906mbr4002 3

- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. https://doi.org/10.1002/sim.1186
- Lee, S. A. W., & Gates, K. M. (2023). From the individual to the group: Using idiographic analyses and two-stage random effects meta-analysis to obtain population level inferences for within-person processes. *Multivariate Behavioral Research*, 59(6), 1220–1239. https://doi.org/10.1080/00273171.2023.2229310
- Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, 10(3), 259-284. https://doi.org/10.1037/1082-989x.10.3.259
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Inter-disciplinary Research & Perspective*, 2(4), 201–218. https://doi.org/10.1207/s15366359mea0204 1
- Neale, M. C., Hunter, M. D., Pritikin, J. N., Zahery, M., Brick, T. R., Kirkpatrick, R. M., Estabrook, R., Bates, T. C., Maes, H. H., & Boker, S. M. (2015). OpenMx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, 81(2), 535–549. https://doi.org/10.1007/s11336-014-9435-8
- R Core Team. (2025). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/