Applying Modern Methods for Missing Data Analysis to the Social Relations Model

Terrence D. Jorgensen
University of Amsterdam

Social Relations Model

- Dyadic data from round-robin design (Kenny, 1994)
  - e.g., all students in a classroom rate each other
  - Block design: heterosexual speed-daters rate opposite sex
- Originally used ANOVA to partition ratings into person- and dyad-level components: $x_{ij} = \mu + P_i + T_j + R_{ij}$
  - μ = average rating
  - P_i = perceiver *i*'s tendency to rate above/below μ
  - T_j = target *j*'s tendency to elicit ratings above/below μ
  - R_{ij} = residual, contains dyadic relationship effect and error

SRM Data

- All participants rate their perceptions of one another

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- μ is estimated from the average of all ratings ($\bar{X}$)

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Row means indicate each perceiver's average rating

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Perceiver effects (P_i) are their averages relative to the grand mean

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Column means indicate the average rating received by each target

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Target effects (T_j) are their averages relative to the grand mean

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Residuals (R_{ij}) are the differences between observed and expected ratings, given μ, P_i, and T_j
  - $R_{\text{Alice}\rightarrow\text{Betty}}$ = 11 – (5 + 3.25 + 3.25) = 11 – 11.5 = –0.5

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

SRM Data

- Self-ratings (or expert observations) can also be recorded, to calculate self–other agreement (or accuracy)

Targets → ↓ Perceivers	Alice	Betty	Cathy	Daria	Ellen	$\bar{X}_{Row}$	Perceiver Effect
Alice	Self	11	6	9	7	8.25	3.25
Betty	2	Self	2	5	4	3.25	−1.75
Cathy	3	11	Self	8	4	6.50	1.50
Daria	4	4	1	Self	2	2.75	−2.25
Ellen	4	7	2	4	Self	4.25	−0.75
$\bar{X}_{Column}$	3.25	8.25	2.75	6.50	4.25	$\bar{X} = 5$	
Target Effect	−1.75	3.25	−2.25	1.50	−0.75		

Estimating SRM with ML

- More recently, maximum likelihood estimation is available by specifying SRM as a random-effects / multilevel model (Snijders & Kenny, 1999)
  - Requires many dummy codes and tedious equality constraints in MLwiN (unavailable in most other multilevel software)
  - Restrictive assumptions in SAS PROC MIXED
- Can also be specified as an "*n*-level SEM" (Brunson et al., 2016) in R package `xxM` (Mehta, 2013)
- Restricted ML recently proposed (Nestler, 2016)

Estimating SRM with MCMC

- Bayesian estimation first proposed by Hoff (2005), more recently by Lüdtke et al. (2013)
- Using MCMC estimation, random effects of perceiver (P_i) and target (T_j) are estimated as parameters, along with variance-component hyperparameters
  - Data augmentation
  - Differs from marginalizing over random effects in ML estimation of random-effects / multilevel models
- Major advantage: posterior distribution of (functions of) all parameters

Missing Data

- ANOVA calculations using all available data might work if data are missing completely at random (MCAR)
  - Very restrictive assumption
- State-of-the-art missing-data methods are maximum likelihood and multiple imputation
  - Only assumes data are missing at random (MAR), given the observed data in the model
  - More defensible if the model includes variables that explain missingness or correlate with the missing values
  - **Auxiliary variables**: not of theoretical interest, but useful to justify MAR assumption

Missing Data

- Multiple imputation has a Bayesian foundation
- Can be done by augmenting observed data with missing data, just like estimating latent variables
  - e.g., random effects, factor scores
  - Usually an "unrestricted" imputation model (e.g., NORM)
    - Freely estimated mean vector and covariance matrix
- Can easily incorporate into the SRM model
  - More efficient than unrestricted imputation model (Merkle, 2011)

Missing Data

- ML simply evaluates the likelihood function using all available data
  - Advantage: no need to "do anything" with the missing data, but may need to incorporate auxiliaries
  - Disadvantage: exogenous predictors must be complete
- Handling of missing data has been touted as an advantage of ML and MCMC estimation
  - Brunson et al. (2016), Hoff (2005), Lüdtke et al. (2013), Nestler (2016), Snijders & Kenny (1999)
  - No one yet described how to do so in MCMC, so I submitted an application to *Social Networks* (under review)

Missing Data in SRM

- From the SRM equation

$$Y_{ij} = \mu + P_i + T_j + E_{ij}$$

- The expected value of the vector of both observations within a dyad is

$$\hat{\mathbf{Y}}_{\{ij\}} = \begin{bmatrix} \hat{Y}_{ij} \\ \hat{Y}_{ji} \end{bmatrix} = \begin{bmatrix} \mu + P_i + T_j \\ \mu + P_j + T_i \end{bmatrix}$$

- Data might be missing for one or both observations within a dyad

Missing Data in SRM

- The likelihood of a dyad's vector is bivariate normal with mean equal to the expected value of the vector, and covariance matrix of residuals:

$$\begin{bmatrix} Y_{ij} \\ Y_{ji} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \hat{Y}_{ij} \\ \hat{Y}_{ji} \end{bmatrix}, \begin{bmatrix} \sigma_E^2 & \\ \rho_E \sigma_E & \sigma_E^2 \end{bmatrix} \right)$$

- Observed data can be augmented with estimates of missing values by using this likelihood of observed data as the prior for missing-data estimates
  - Assumes MAR, conditional on expected values

Missing Data in SRM

- Even complete data are augmented with estimates of random effects, distributed bivariate normally:

$$\begin{bmatrix} P_i \\ T_i \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_P^2 & \\ \rho_{PT}\sigma_P\sigma_T & \sigma_T^2 \end{bmatrix} \right)$$

- This can be extended to include 1 or more auxiliary covariates (X_i):

$$\begin{bmatrix} P_i \\ T_i \\ X_i \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ \mu_X \end{bmatrix}, \begin{bmatrix} \sigma_P^2 & & \\ \rho_{TP}\sigma_T\sigma_P & \sigma_T^2 & \\ \rho_{XP}\sigma_X\sigma_P & \rho_{XT}\sigma_X\sigma_T & \sigma_X^2 \end{bmatrix} \right)$$

Missing Data in SRM

- Covariates might also be substantively interesting as predictors of the random effects:

$$P_i = \beta_1^P X_i + \varepsilon_i \qquad , \qquad T_j = \beta_1^T X_j + \delta_j$$

- In which case the covariate(s) would be independent of the residual perceiver and target effects:

$$\begin{bmatrix} \varepsilon_i \\ \delta_i \\ X_i \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ \mu_X \end{bmatrix}, \begin{bmatrix} \sigma_\varepsilon^2 & & \\ \sigma_{\varepsilon\delta} & \sigma_\delta^2 & \\ 0 & 0 & \sigma_X^2 \end{bmatrix} \right)$$

Missing Data in SRM

- The MAR assumption is more defensible if explanatory or auxiliary covariates can be included that either:
  - Explain missingness
  - Correlate with missing values
- Dyad-level covariates can also be incorporated in the model, can could be either:
  - Constant within dyad ($V_{ij} = V_{ji}$; e.g., "How long have you known each other?")
  - Vary within dyad ($W_{ij} \neq W_{ji}$; e.g., "How well does [this friend] know you?")

Missing Data in SRM

- Dyad-level covariates can be merely auxiliary:

$$\begin{bmatrix} Y_{ij} \\ Y_{ji} \\ V_{\{ij\}} \\ W_{ij} \\ W_{ji} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \hat{Y}_{ij} \\ \hat{Y}_{ji} \\ \mu_V \\ \mu_W \\ \mu_W \end{bmatrix}, \begin{bmatrix} \sigma_R^2 & & & & \\ \rho_R\sigma_R & \sigma_R^2 & & & \\ \rho_{VR}\sigma_V\sigma_R & \rho_{VR}\sigma_V\sigma_R & \sigma_V^2 & & \\ \rho_{WR}\sigma_W\sigma_R & \rho_{WR}\sigma_W\sigma_R & \rho_{WV}\sigma_W\sigma_V & \sigma_W^2 & \\ \rho_{WR}\sigma_W\sigma_R & \rho_{WR}\sigma_W\sigma_R & \rho_{WV}\sigma_W\sigma_V & \rho_W\sigma_W & \sigma_W^2 \end{bmatrix} \right)$$

- So including them would make the MAR assumption more tenable
- Note that the equality constraints are unnecessary if persons *i* and *j* have roles (e.g., men rating women)

Missing Data in SRM

- Dyad-level covariates can also be explanatory:

$$\hat{\mathbf{Y}}_{\{ij\}} = \begin{bmatrix} \hat{Y}_{ij} \\ \hat{Y}_{ji} \end{bmatrix} = \begin{bmatrix} \mu + \gamma_1 V_{\{ij\}} + \gamma_2 W_{ij} + P_i + T_j \\ \mu + \gamma_1 V_{\{ij\}} + \gamma_2 W_{ji} + P_j + T_i \end{bmatrix}$$

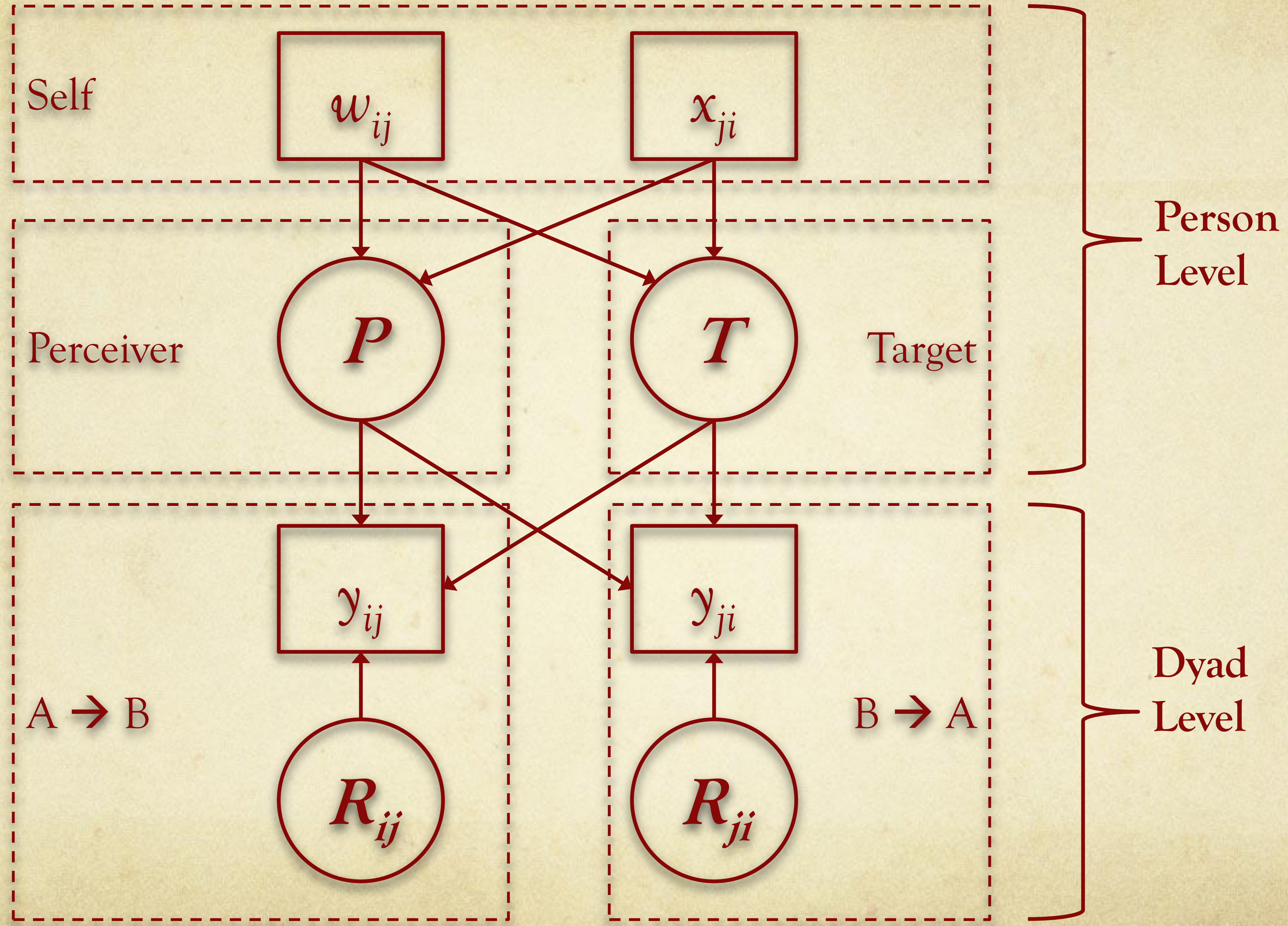
- In which case they should not correlate with residuals:

$$\begin{bmatrix} Y_{ij} \\ Y_{ji} \\ V_{\{ij\}} \\ W_{ij} \\ W_{ji} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \hat{Y}_{ij} \\ \hat{Y}_{ji} \\ \mu_V \\ \mu_W \\ \mu_W \end{bmatrix}, \begin{bmatrix} \sigma_R^2 & & & & \\ \rho_R \sigma_R & \sigma_R^2 & & & \\ 0 & 0 & \sigma_V^2 & & \\ 0 & 0 & \rho_{WV} \sigma_W \sigma_V & \sigma_W^2 & \\ 0 & 0 & \rho_{WV} \sigma_W \sigma_V & \rho_W \sigma_W & \sigma_W^2 \end{bmatrix} \right)$$

- Again, equality constraints on slopes and (co)variances unnecessary if *i* and *j* have roles

Evaluation in MCMC and ML

- I implemented the methods described above, mimicking a real-data analysis of sorority data
  - Described in APS slides, which can be downloaded from my Open Science Framework account, along with sorority data and syntax for `xxM` and `RStan`
  - https://osf.io/fmhg6/
- To evaluate the frequency properties of the method, I used parameter estimates as population values to simulate 200 samples of the same size, and imposed the same missing-data pattern (MCAR)
  - Also compared Bayes to (FI)ML using xxM package, specifying the SRM as a multilevel SEM



Self
w_ij
x_ji
Perceiver
P
T
Target
y_ij
y_ji
A → B
B → A
R_ij
R_ji
Person Level
Dyad Level

Full-Information ML (in xxM)

Parameter	True θ	$\hat{\theta}$	Bias	RMSE	Coverage	Power / α
μ (intercept)	2.00	1.99	−0.01	0.07	95%	100%
Perceiver σ^2	0.25	0.24	−0.01	0.05	95%	100%
β_1	0.30	0.30	0.00	0.07	95%	98%
β_2	0	0.00	0.00	0.07	94%	6%
Target σ^2	0.12	0.16	0.04	0.05	75%	100%
β_1	0	0.00	0.00	0.06	96%	4%
β_2	0.30	0.30	0.00	0.06	98%	100%
Generalized ρ	−0.02	−0.01	0.01	0.04	95%	7%
Residual σ^2	0.49	0.49	0.00	0.03	95%	100%
Dyadic ρ	0.15	0.00	−0.15	0.15	3%	5%

Impute/Augment Data (in RStan)

Parameter	True θ	$\hat{\theta}$	Bias	RMSE	Coverage	Power / α
μ (intercept)	2.00	1.93	−0.07	0.07	0%	100%
Perceiver σ^2	0.25	0.27	0.02	0.06	94%	100%
β_1	0.30	0.29	−0.01	0.07	95%	95%
β_2	0	0.00	0.00	0.06	94%	7%
Target σ^2	0.12	0.13	0.01	0.03	94%	100%
β_1	0	0.00	0.00	0.07	96%	4%
β_2	0.30	0.29	−0.01	0.05	97%	100%
Generalized ρ	−0.02	−0.02	0.00	0.03	96%	6%
Residual σ^2	0.49	0.50	0.01	0.03	95%	100%
Dyadic ρ	0.15	0.14	−0.01	0.05	96%	75%

References

Brunson, J. A., Øverup, C. S., & Mehta, P. D. (2016). A social relations examination of neuroticism and emotional support. *Journal of Research in Personality, 63*, 67–71. doi:10.1016/j.jrp.2016.05.012

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association, 100*(469), 286–295. doi:10.1198/016214504000001015

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford.

Lüdtke, O., Robitzsch, A., Kenny, D. A., & Trautwein, U. (2013). A general and flexible approach to estimating the social relations model using Bayesian methods. *Psychological Methods, 18*(1), 101–119. doi:10.1037/a0029252

Mehta, P. (2013). xxm: Structural equation modeling for dependent data (version 0.6.0) [Computer software]. Retrieved from http://xxm.times.uh.edu/

Merkle, E. C. (2011). A comparison of imputation methods for Bayesian factor analysis models. *Journal of Educational and Behavioral Statistics, 36*(2), 257–276. doi:10.3102/1076998610375833

Nestler, S. (2016). Restricted maximum likelihood estimation for parameters of the social relations model. *Psychometrika, 81*(4), 1098–1117. doi:10.1007/s11336-015-9474-9

Snijders, T. A. B., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships,* 6, 471–486. doi:10.1111/j.1475-6811.1999.tb00204.x