ANALYZING INTENSIVE LONGITUDINAL DATA (AILD)

Tilburg University
December 13 - 17, 2021

Niall Bolger, Ph.D.
Columbia University

Jean-Philippe Laurenceau, Ph.D.
University of Delaware

Daily Schedule

Morning:

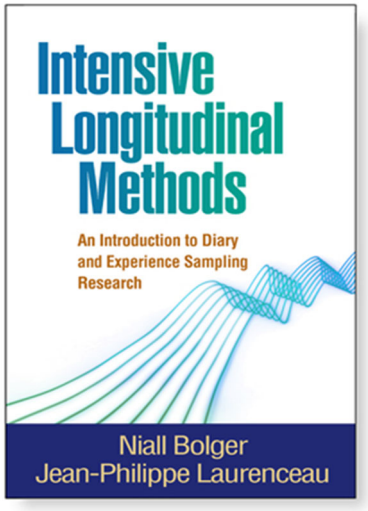
9:00a-10:10a: Lecture (70 min)
10:10a-10:25a: Break (15 min)
10:25a-11:25p: Lecture (60 min)
11:25a-11:40p: Break (15 min)
11:40a-12:30p: Lecture (50 min)

12:30a-1:30p: LUNCH (60 min)

Afternoon:

1:30p-2:40p: Lecture (70 min)
2:40p-2:55: Break (15 min)
2:55p-3:55p: Lecture (60 min)
3:55p-4:10p: Break (15)
4:10p-5:00p: Lecture (50 min)

Bolger & Laurenceau (2013)



Guilford Press *Methodology in the Social Sciences* Series

www.intensivelongitudinal.com

3

Chapter 6

Modeling IL Categorical Outcomes

Previously

- Modeling the causal process of a continuous outcome
- What if you want to understand causal processes for categorical outcomes? For count outcomes?
- We will need to use what are called Generalized Linear Mixed Models (GLMMs)

5

Objectives

- Introduce the example 28-day intensive longitudinal dataset, the Grad Couples study
- Introduce the research question that is the focus of this chapter
- Show the structure of the dataset and plot the focal variables separately for each couple

6

Dataset and Research Question

- The Grad Couples dataset is the same one we used in the psychometrics lecture to examine the reliability of change in daily mood; we found that R_c for daily anger was 0.82
- We want to examine a possible causal process that begins with **anger/irritability of the female partner** in her morning diary report and her **male partner's report of a conflict** later that day

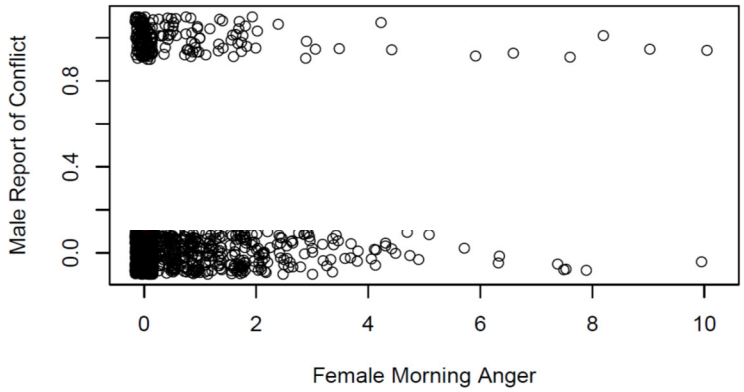
7

Grad Couples Dataset with Key Variables

id	time	pconf	amang
		(Y_{ij})	(X_{ij})
1	1	0	0.0
1	2	0	0.4
1	3	0	0.0
1	4	0	0.0
1	5	0	0.0
1	6	1	0.0
1	7	0	0.0
1	8	0	0.0
.	.	.	.
.	.	.	.
1	25	0	0.0
1	26	0	0.0
1	27	0	0.0
1	28	0	0.0
3	1	0	0.0
3	2	0	0.0
3	3	0	0.0
.	.	.	.
.	.	.	.
3	26	0	0.0
3	27	0	0.4
3	28	0	0.0
5	1	0	2.5
5	2	0	2.1
5	3	0	2.5
.	.	.	.
.	.	.	.
102	26	0	0.0
102	27	0	0.0
102	28	0	0.0

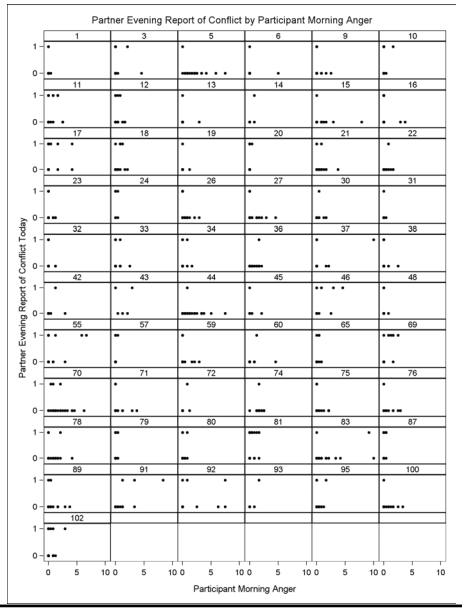
8

Male Report of Conflict by Female Morning Anger



9

Panel Plot of Raw Data



10

Research Question

- Does morning anger of the female partner predict male partner reports of conflict later that day?

- Specifically: On days when female partners begin the day more or less angry than they typically are, are their male partners that evening more likely than usual to say that they had a relationship conflict?

11

Comments

- Our research question concerns a within-couple causal process linking the morning anger of the female partner to the male partner's report of a conflict later that day
- We saw that the data structure was sparse; daily conflicts are relatively rare and morning anger is even rarer
- We saw that couples differ in levels of morning anger and conflict
- These facts will become relevant when we attempt to estimate a model that allows for couple-specific causal effects

12

Objectives

- Introduce Generalized Linear Models, a form of linear model that allows the analysis of categorical (and count) outcomes
- Introduce the central features of GLMs, namely, the link function and the error distribution

13

Generalized Linear Models

Greatest Hits of the 70s (in Statistics)

- 1970: Tukey: Exploratory Data Analysis
- **1972: Nelder & Wedderburn: Generalized Linear Models**
- 1977: Dempster, Laird & Rubin: The EM algorithm
- 1979: Efron: Resampling methods: The Bootstrap and others

14

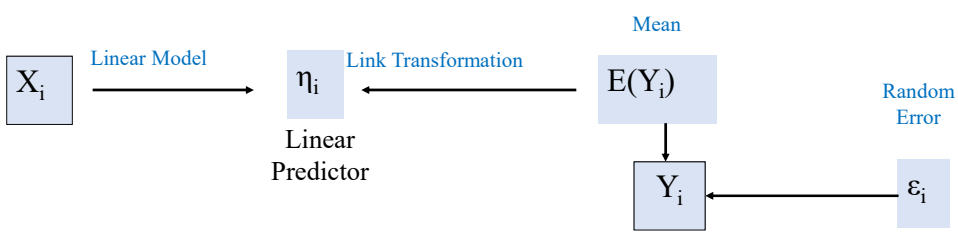
Generalized Linear Models

- Nelder and Wedderburn (1972) developed a class of models that assume that
  - Y belongs to a *class of distributions*, called the *exponential family* (e.g. normal, binomial)
  - a **transformation of the mean of Y** can be described using a linear (additive) model
  - the **variance of Y** can be described by a separate variance function

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General), 135*, 370-384.

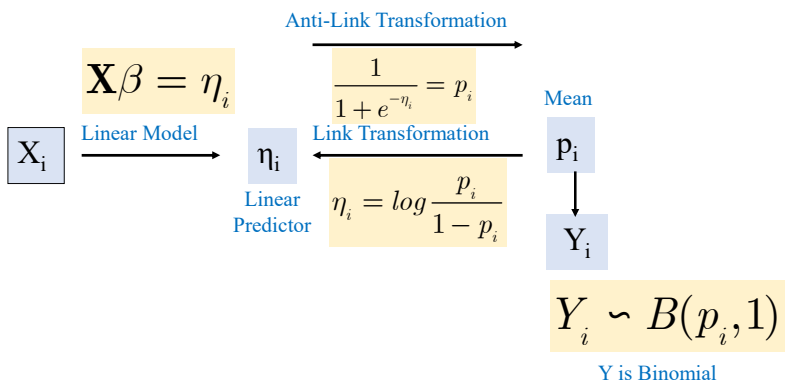
15

Generalized Linear Models



16

Generalized Linear Models



17

Generalized Linear Models

$$Y_i \sim B(p_i,1)$$

$$E(Y_i) = p_i$$

$$log\frac{p_i}{1-p_i} = \eta_i$$

$$\eta_i = \beta_0 + \beta_1 X_{1i}$$

When Y is 0 or 1, the result of a binomial process, with one trial

- $E(Y_i) = p_i$, a probability of getting a score of 1
- $Var(Y_i) = p_i(1-p_i)$
- A logit *link function*
  - $\eta_i = \log(p_i/(1-p_i))$
- The linear predictor η_i is a linear function of Xs:
  - $\eta_i = \beta_0 + \beta_1 X_{i1}$ + etc.

18

Objectives

- Introduce Generalized Linear Mixed Models that are suitable for intensive longitudinal data with binary and count outcomes
- Show that GLMMs allow for upper-level random effects analogous to regular multilevel models
- Use a GLMM to analyze daily conflict in couples
- Interpret the key fixed and random effects
- See plots of model predictions

19

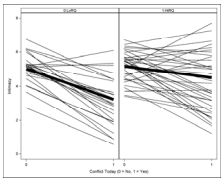
Generalized Linear Mixed Models (GLMMs)

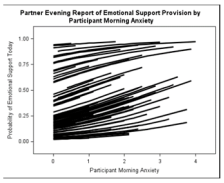
$$\eta_{ij} = \mathbf{X}\beta + \mathbf{Z}u$$

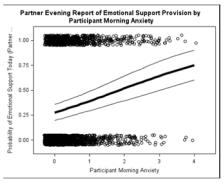
$$p_{ij} = \frac{1}{1 + e^{-\eta_{ij}}}$$

$$Y_{ij} \sim B(p_{ij}, 1)$$

$$log(\frac{p_{ij}}{1 - p_{ij}}) = \eta_{ij}$$







20

GLMM Random Intercepts Specification

$$\eta_{ij} = \gamma_{00} + \gamma_{01}\breve{X}_{\mathbf{B}} + \boxed{\gamma_{10}}\breve{X}_{\mathbf{W}} + \gamma_{20}\breve{Y}_{ij-1} + \gamma_{30}\breve{T}_i + u_{0j} \qquad 6.1$$

$$u_{0j} \sim N(0, \tau_{00}) \qquad 6.2$$

$$p_{ij} = 1/(1 + e^{-\eta_{ij}}) \qquad 6.3$$

$$Y_{ij} \backsim B(p_{ij}, 1) \qquad 6.4$$

21

Daily Conflict Example: All Variables

- time: Diary day (1, 2, …, 28)
- **time7c**: Time in units of one week, centered
- amang: Morning Anger (0-10)
- amangc: Centered morning anger
- **amangcb**: Centered morning anger, between
- **amangcw**: Centered morning anger, within
- **pconf**: Today's partner report of conflict (0-1)
- lpconf: Yesterday's partner report of conflict (0-1)
- **lpconfc**: Centered, yesterday's partner report of conflict

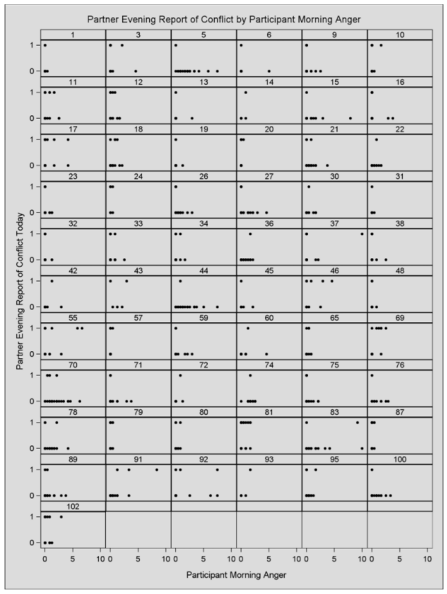
22

Full Categorical Dataset

Id	time	time7c	pconf	lpconfc	amang	amangc	amangcb	amangcw
		$(\tilde{T}_i)$	(Y_{ij})	$(\tilde{Y}_{ij-1})$	(X_{ij})	$(\tilde{X}_{ij})$	$(\tilde{X}_{.j})$	$(\tilde{X}_{ij} - \tilde{X}_{.j})$
1	1	-	-	-	-	-	-	-
1	2	-1.88	0	-.157	0.4	-.07	-.47	.40
1	3	-1.73	0	-.157	0.0	-.49	-.47	-.02
1	4	-1.58	0	-.157	0.0	-.49	-.47	-.02
1	5	-1.43	0	-.157	0.0	-.49	-.47	-.02
1	6	-1.28	1	-.157	0.0	-.49	-.47	-.02
1	7	-1.13	0	.843	0.0	-.49	-.47	-.02
1	8	-0.99	0	-.157	0.0	-.49	-.47	-.02
.	.	.	.		.			.
.	.	.	.		.			.
1	25	1.53	0	-.157	0.0	-.49	-.47	-.02
1	26	1.68	0	-.157	0.0	-.49	-.47	-.02
1	27	1.83	0	-.157	0.0	-.49	-.47	-.02
1	28	1.98	0	-.157	0.0	-.49	-.47	-.02
3	1	-	-	-	-		-	-
3	2	-1.88	0	-.157	0.0	-.49	-.22	-.26
3	3	-1.73	0	-.157	0.0	-.49	-.22	-.26
.	.	.	.		.			.
.	.	.	.		.			.
3	26	1.68	0	-.157	0.0	-.49	-.22	-.26
3	27	1.83	0	-.157	0.4	-.07	-.22	.15
3	28	1.98	0	-.157	0.0	-.49	-.22	-.26
5	1	-	-	-	-		-	-
5	2	-1.88	0	-.157	2.1	1.60	1.61	-.02
5	3	-1.73	0	-.157	2.5	2.01	1.61	.40
.	.	.	.		.			.
.	.	.	.		.			.
102	26	26	0	-.157	0.0	-.49	-.22	-.26
102	27	27	0	-.157	0.0	-.49	-.22	-.26
102	28	28	0	-.157	0.0	-.49	-.22	-.26

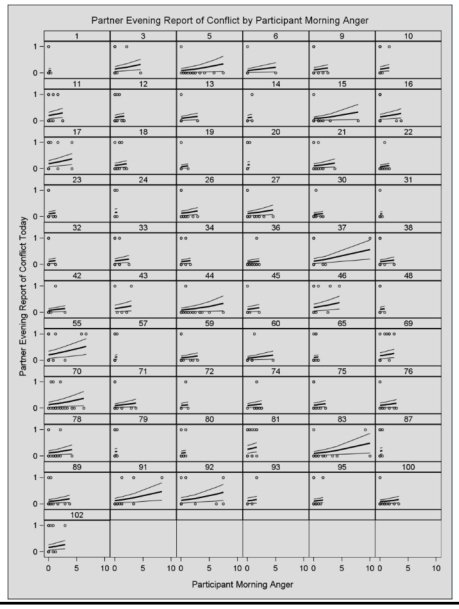
23

Panel Plot of Raw Data



24

Raw Data & Model Predictions on Probability Scale



25

Generalized Linear Mixed Model (Frequentist)

(R package *lme4*, *glmer*)

Code:

```
catmodel <- glmer(pconf ~ amangcw + amangcb + lpconfc +
time7c + (1 | id), family = binomial, data = categorical)
summary(catmodel)
```

Output:

```
> summary(catmodel)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
Family: binomial  ( logit )
Formula: pconf ~ amangcw + amangcb + lpconfc + time7c + (1 | id)
Data: categorical
     AIC      BIC   logLik deviance df.resid
  1084.8   1116.0   -536.4   1072.8     1339
```

26

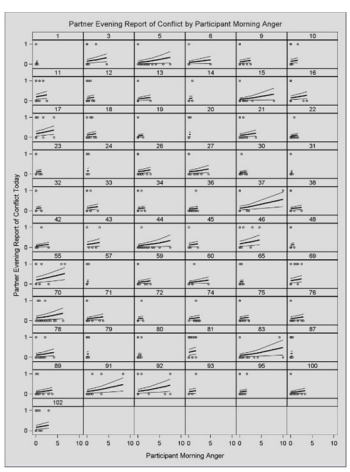
Output (con't):

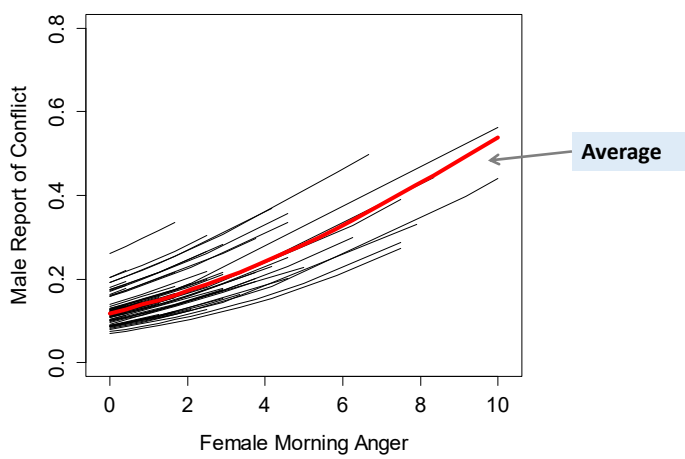
```
Random effects:

 Groups Name        Variance Std.Dev.
 id     (Intercept) 0.2475   0.4975

Number of obs: 1345, groups:  id, 61

Fixed effects:

            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.90161    0.10894 -17.456  < 2e-16 ***
amangcw      0.21565    0.06738   3.200  0.00137 **
amangcb     -0.22437    0.22398  -1.002  0.31649
lpconfc      0.31826    0.20892   1.523  0.12767
time7c      -0.19222    0.07073  -2.718  0.00657 **
```

	2.5 %	97.5 %
amangcw	0.08	0.35

27

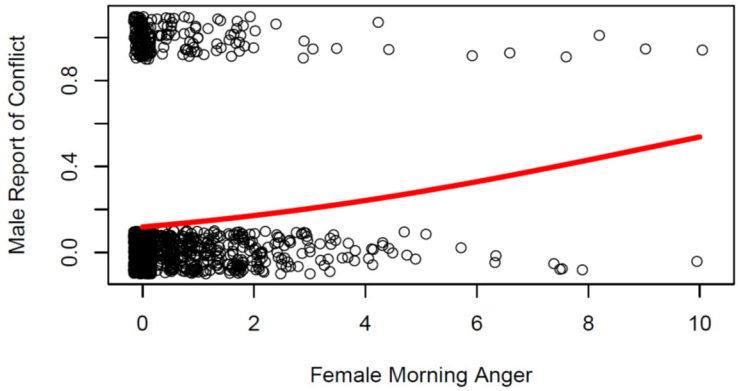
Predicted Probabilities (from random intercept, fixed slope model)





28

Male Report of Conflict by Female Morning Anger



29

GLMM Random Slopes Specification

$$\eta_{ij} = \gamma_{00} + \gamma_{01}\breve{X}_{\mathbf{B}} + \boxed{\gamma_{10}}\breve{X}_{\mathbf{W}} + \gamma_{20}\breve{Y}_{ij-1} + \gamma_{30}\breve{T}_i + u_{0j} + \boxed{u_{1j}\breve{X}_{\mathbf{W}}} \qquad 6.1$$

$$u_{0j} \sim N(0, \tau_{00}) \qquad 6.2$$

$$u_{1j} \sim N(0, \boxed{\tau_{11}}) \qquad 6.3$$

$$p_{ij} = 1/(1 + e^{-\eta_{ij}}) \qquad 6.4$$

$$Y_{ij} \backsim B(p_{ij}, 1) \qquad 6.5$$

30

Comments

- GLMMs provide the same richness as multilevel models when applied to IL data with binary outcomes
- The functional form of the model is sigmoidal in the probability scale but linear in the log-odds scale
- Given the sparseness of our data, our model using ML estimation was unable to estimate random slopes

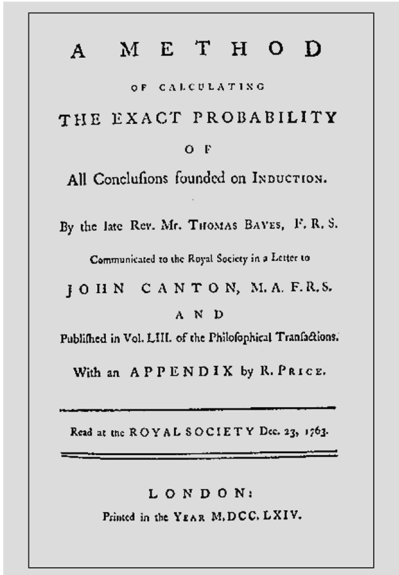
31

Objectives

- Introduce the key ideas underlying Bayesian estimation
- Compare Bayesian and Frequentist (ML) models of the daily conflict data
- Show that the Bayesian approach allows one to estimate random slopes
- Show how different the conclusions are depending on the method used

32

Philosophical Transactions of the Royal Society (1764)



A METHOD

OF CALCULATING

THE EXACT PROBABILITY

OF

All Conclusions founded on INDUCTION.

By the late Rev. Mr. THOMAS BAYES, F. R. S.

Communicated to the Royal Society in a Letter to

JOHN CANTON, M. A. F. R. S.

AND

Published in Vol. LIII. of the Philosophical Transactions.

With an APPENDIX by R. PRICE.

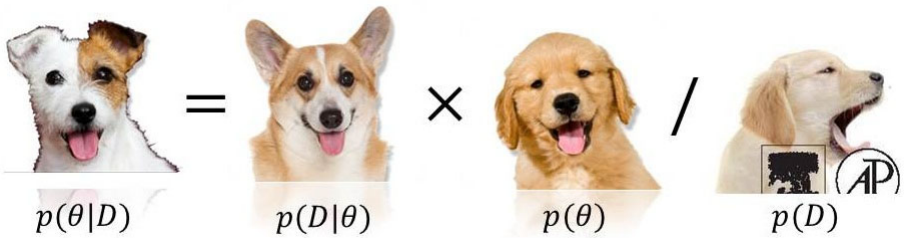
Read at the ROYAL SOCIETY Dec. 23, 1763.

LONDON:

Printed in the YEAR M.DCC.LXIV.

33

Bayes Theorem: The Canine Version



Kruschke, J. K. (2014). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan. Academic Press.*

34

Comparison of Frequentist and Bayes

- We will be using a computational approach to Bayesian estimation called Markov-Chain Monte Carlo (Winbugs, Jags, STAN, Mplus 8.0)
- When non-informative priors are used, both approaches tend to give similar estimates and uncertainty intervals
- When Frequentist estimation (ML) methods fail to converge, Bayesian equivalents tend to work

35

Comparison of *lme4* & *brms* Specification

Frequentist:

```
catmodel <- glmer(pconf ~ amangcw + amangcb + lpconfc + time7c
+ (1 | id), family = binomial, data = categorical)
summary(catmodel)
```

Bayesian (w. no random slopes):

```
catmodB<- brm(pconf ~ amangcw + amangcb + lpconfc + time7c +
(1 | id), family = binomial, data = categorical)
summary(catmodB)
```

Bayesian (w. random amangcw slope):

```
catmod1B<-brm(pconf ~ amangcw + amangcb + lpconfc + time7c +
(amangcw| id), family = binomial, data = categorical, cores=4)
summary(catmod1B)
```

36

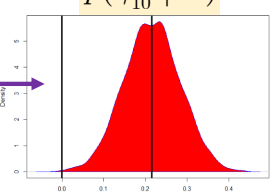
Bayesian Estimation
(Rerun random intercepts model in brms)

Family: binomial (logit) Prior: $p(\gamma_{10}) \sim N(0,10))$

```
Formula: pconf ~ amangcw + amangcb + lpconfc + time7c + (1 | id)

Data: categorical (Number of observations: 1345)

Samples: 2 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 2000

Group-Level Effects:

~id (Number of levels: 61)
                 Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat

sd(Intercept)      0.53      0.14     0.27      0.8        482 1.01

Population-Level Effects:

           Estimate Est.Error l-95% CI u-95% CI Ef

Intercept     -1.92      0.11    -2.14    -1.70
amangcw        0.21      0.07     0.08     0.35
amangcb       -0.25      0.24    -0.72     0.20
lpconfc        0.30      0.21    -0.12     0.71
time7c        -0.19      0.07    -0.33    -0.05
```

$\hat{\gamma}_{10}$

$p(\gamma_{10} \mid D)$



37

Run random slope model in *brms*

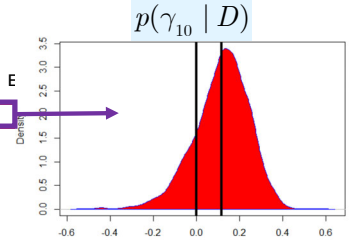
Family: binomial (logit) Prior: $p(\gamma_{10}) \sim N(0,10))$

```
Formula: pconf ~ amangcw + amangcb + lpconfc + time7c + (amangcw | id)

Data: categorical (Number of observations: 1345)

Samples: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup samples = 4000

Group-Level Effects:
~id (Number of levels: 61)

                        Estimate Est.Error l-95% CI u-95% CI Eff.Sample Rhat
sd(Intercept)               0.56      0.13     0.31     0.83       1599    1
sd(amangcw)                 0.37      0.19     0.05     0.77       1071    1
cor(Intercept,amangcw)      0.68      0.30    -0.12     0.99       1650    1

Population-Level Effects:

           Estimate Est.Error l-95% CI u-95% CI E

Intercept     -1.94      0.12    -2.17    -1.71
amangcw        0.11      0.13    -0.17     0.33
amangcb       -0.23      0.23    -0.69     0.22
lpconfc        0.27      0.22    -0.16     0.68
time7c        -0.20      0.07    -0.35    -0.05
```

$\hat{\tau}_{11}$

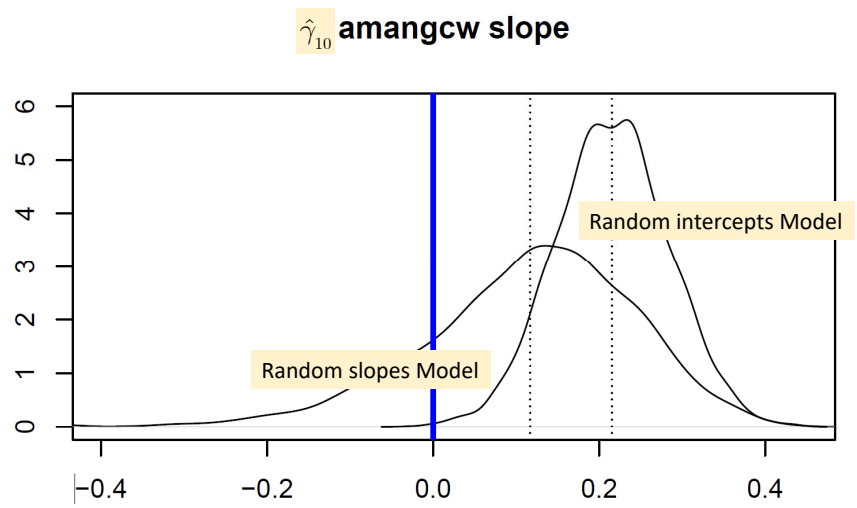
$\hat{\gamma}_{10}$

$p(\gamma_{10} \mid D)$



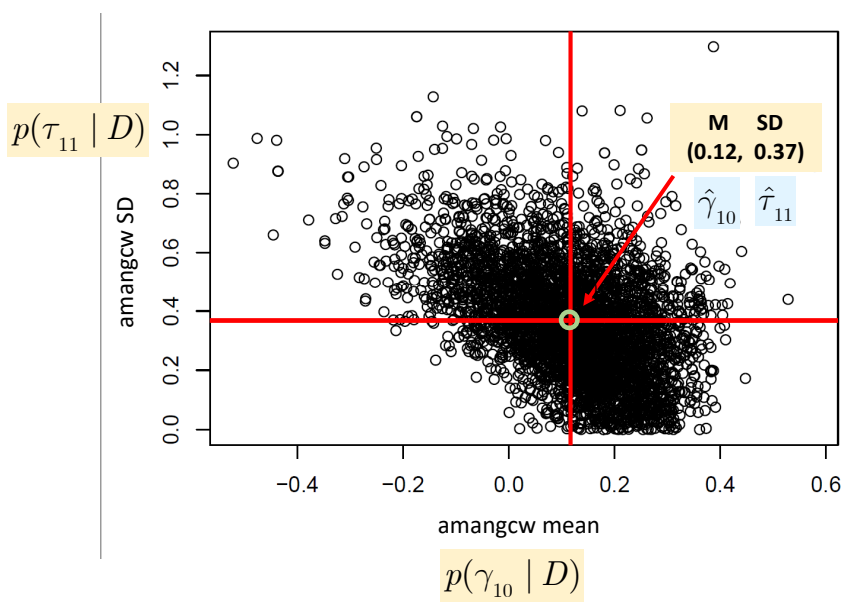
38

amangcw Fixed Effect: Posterior Distributions



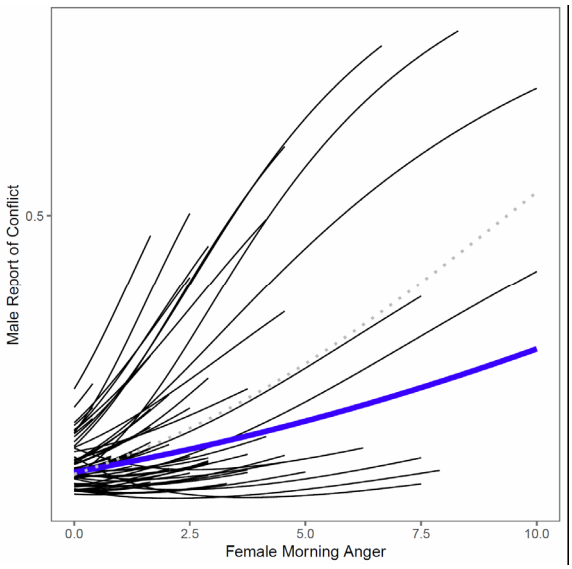
39

Posterior for amangcw Random Slopes Model



40

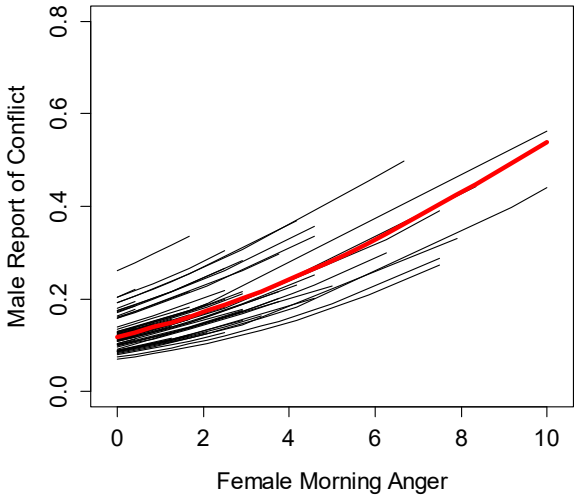
Random Slopes Model: Sample Predictions



M	SD
(0.12,	0.37)
$\hat{\gamma}_{10}$	$\hat{\tau}_{11}$

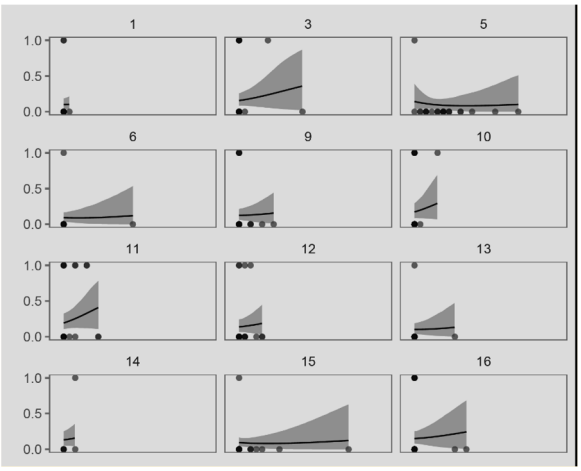
41

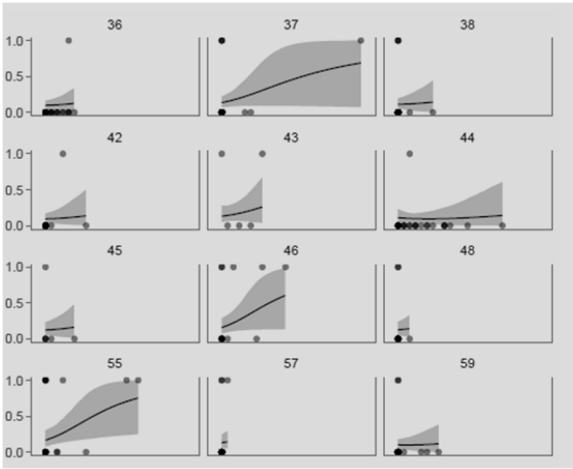
Compare with Fixed Slope Model



42

Panel Plot of Specific Sample Predictions





43

Comments

- Bayesian estimation of GLMMs requires the use of complex algorithms from computational statistics
- These algorithms are indispensable for researchers who work with binary IL outcomes, because they overcome a serious limitation of traditional ML estimation, its tendency toward non-convergence when estimating random effects
- Bayesian model output is voluminous and is best studied using a combination of numerical and graphical summaries

44

Chapter Summary

- We showed an example of intensive longitudinal data with a binary outcome
- We used an extension of linear mixed models called generalized linear mixed models
- These work with appropriate transformations of Y and relate them linearly to Xs
- Interpreting the results, though, requires that we back-transform the results to the original Y scale
- Bayesian methods are likely necessary to estimate multilevel logistic regression models
- Need to distinguish posterior distributions of parameters from y predictions and fit of the model

45

Recommended Readings

- Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

- Sommet, N. and Morselli, D. (2017). Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. *International Review of Social Psychology, 30*, 203–218.

- Rabe-Hesketh, S., & Skrondal, A. (2009). Generalized linear mixed effects models. In G. M. Fitzmaurice, M. Davidian, G. Verbeke, & G. Molenberghs (Eds.), *Longitudinal data analysis* (pp. 79–106). Boca Raton: CRC Press.

46

Chapter 5

- **Psychometrics of Intensive Longitudinal Data**

Section 5.1

- Why traditional conception of reliability is not helpful for longitudinal measures

Objective

- Traditional psychometrics is concerned with measuring differences between people
- Intensive longitudinal studies allow us to study change
- Random-Effects Structure of ILD: Items, Persons, Time

49

Cronbach's Reliability Coefficient Alpha

- Most often used in situation where a sample of people are assessed on a fixed set of items
- Coefficient Alpha is defined as the proportion of the total variation that is true-score variation
- Alpha is a ratio of true-score variation to true-score plus error variation

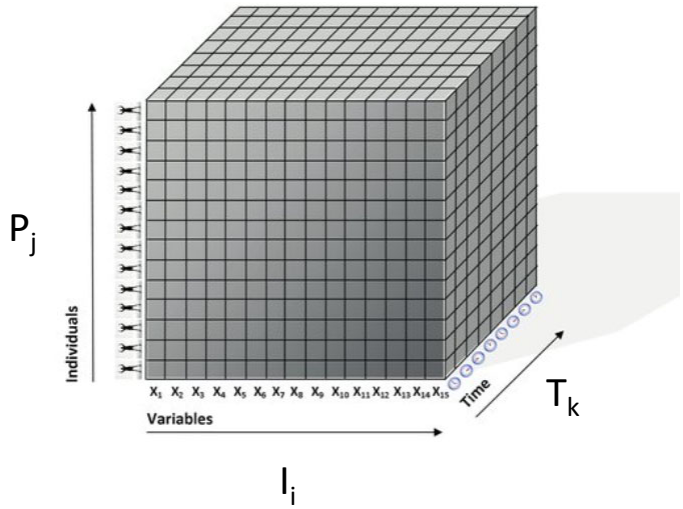
50

Structure of Intensive Longitudinal Data

- We have, as usual, people and items, but the time dimension is crucial
- How can one adapt traditional psychometrics so that we can be confident that our measures of change are reliable?

51

Three Dimensions of Variation in ILD



52

Measurement Error Model from Classical Test Theory

(one person, one time point, multiple items *i*)

$$M_i = \mu + e_i$$

Measurement Random Variable	Constant (fixed effect)	Error Random Variable

53

Measurement Error Model from Classical Test Theory (multiple people *j*, one time point, multiple items *i*)

$$M_{ij} = \mu + P_j + e_{ij}$$

Measurement Random Variable	Constant (fixed effect)	Person Random Variable	Error Random Variable

54

Measurement Error Model from Classical Test Theory (multiple people *j*, one time point, multiple items *i*, plus distinct error)

$$M_{ij} = \mu + I_i + P_j + e_{ij}$$

Measurement Random Variable	Constant (fixed effect)	Item Random Variable	Person Random Variable	Error Random Variable

For now, we assume no interaction between random factors.
A variation on this model is the basis for Cronbach's Alpha.

55

Measurement Error Model from Classical Test Theory (multiple people *j*, multiple time points *k*, multiple items *i*, distinct errors)

$$M_{ijk} = \mu + I_i + P_j + T_k + e_{ijk}$$

Measurement Random Variable	Constant (fixed effect)	Item Random Variable	Person Random Variable	Time Random Variable	Error Random Variable

For now, we assume no interaction between random factors.

56

Measurement Error Model from Classical Test Theory: Variance Decomposition

$$V(M_{ijk}) = V(I_i) + V(P_j) + V(T_k) + V(e_{ijk})$$

Total Variance	Item Variance	Person Variance	Time Variance	Error Variance

For now, we assume no interaction between random factors.

57

Comments

- Traditional measures of reliability are based on the extent to which items hang together across a sample of people
- Traditional reliability as measured by Cronbach's Alpha is high to the extent to which people who are high on one item tend to be high on the others, and people who are low on one item tend to be low on the others
- With ILD, at least three components of variance can be distinguished: Item, Person, & Time; when data are balanced across the three dimensions, the three components are orthogonal and sum to the total variance

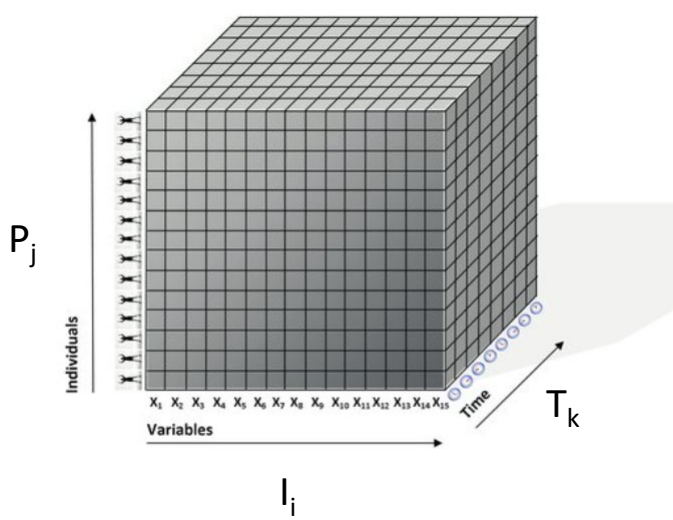
58

Objective

- Introduction to the Bar Examination and Grad Couples studies
- Random Effects ANOVA measurement model from Generalizability Theory (GT)

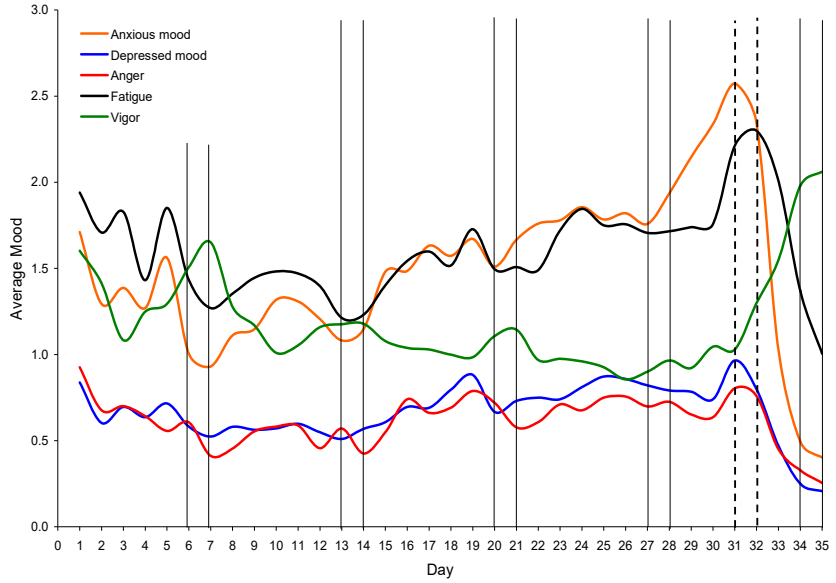
59

Design for Bar and Grad Couples Studies



60

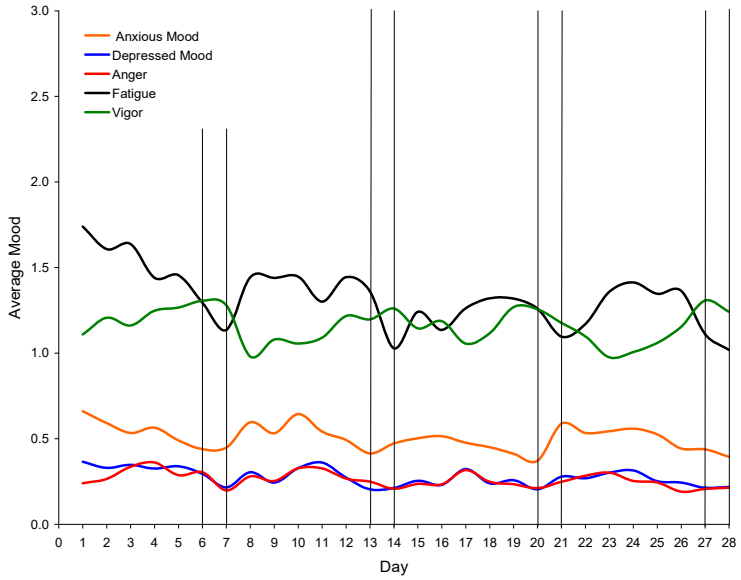
Mood Time Trajectories for in Bar Exam Study



Source: Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32, 917-929.*

61

Time Trajectories for Moods in Grad Couples



Source: Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32, 917-929.*

62

Statistical Model

$$M_{ijk} = \mu + I_i + P_j + T_k + IP_{ij} + IT_{ik} + PT_{jk} + [IPT_{ijk} + e_{ijk}]$$

I = Items: *i* = 1 … *m*

P = Persons: *j* = 1 … *n*

T = Time (days): *k* = 1 … *l*

In this application, we include all two-way interactions. The three-way is confounded with error.

63

Bar Exam Dataset: Variance Components

Source of Variance	Anxious Mood (%)	Depressed Mood (%)	Anger (%)	Fatigue (%)	Vigor (%)
$\hat{\sigma}^2_{PERSON}$	0.383 (27.1)	0.251 (24.3)	0.158 (18.1)	0.521 (36.8)	0.255 (26.9)
$\hat{\sigma}^2_{DAY}$	0.218 (15.4)	0.018 (01.8)	0.013 (1.5)	0.066 (4.7)	0.077 (8.1)
$\hat{\sigma}^2_{ITEM}$	0.011 (0.8)	0.141 (13.7)	0.056 (6.4)	0.004 (0.3)	0.004 (0.4)
$\hat{\sigma}^2_{PERSON*DAY}$	0.469 (33.1)	0.174 (16.9)	0.281 (32.3)	0.579 (40.8)	0.281 (29.6)
$\hat{\sigma}^2_{PERSON*ITEM}$	0.041 (2.9)	0.117 (11.3)	0.073 (8.5)	0.019 (1.3)	0.057 (6.1)
$\hat{\sigma}^2_{DAY*ITEM}$	0.002 (0.1)	0.013 (1.3)	0.004 (0.5)	0.00 (0.0)	0.006 (0.7)
$\hat{\sigma}^2_{ERROR}$	0.292 (20.6)	0.315 (30.6)	0.286 (32.8)	0.229 (16.1)	0.267 (28.2)
Total	1.416 (100)	1.029 (100)	0.871 (100)	1.418 (100)	0.947 (100)

Note. All variance component estimates were based on 35 diary days.

Source: Cranford et al., 1996

64

Grad Couples Dataset: Variance Components

Source of Variance	Anxious Mood (%)	Depressed Mood (%)	Anger (%)	Fatigue (%)	Vigor (%)
$\hat{\sigma}^2_{PERSON}$	0.229 (31.2)	0.072 (16.2)	0.053 (12.4)	0.430 (29.4)	0.356 (31.1)
$\hat{\sigma}^2_{DAY}$	0.003 (0.4)	0.001 (0.3)	0.001 (0.1)	0.023 (2.0)	0.006 (0.5)
$\hat{\sigma}^2_{ITEM}$	0.004 (0.5)	0.013 (2.9)	0.009 (2.2)	0.004 (0.2)	0.036 (3.2)
$\hat{\sigma}^2_{PERSON*DAY}$	0.281 (38.4)	0.162 (36.2)	0.207 (48.3)	0.663 (45.2)	0.388 (33.9)
$\hat{\sigma}^2_{PERSON*ITEM}$	0.024 (3.3)	0.034 (7.5)	0.025 (5.8)	0.080 (5.5)	0.076 (6.7)
$\hat{\sigma}^2_{DAY*ITEM}$	0.0 (0.0)	0.001 (0.2)	0.001 (0.1)	0.00 (0.0)	0.001 (0.1)
$\hat{\sigma}^2_{ERROR}$	0.191 (26.1)	0.165 (36.8)	0.133 (31.1)	0.261 (17.8)	0.280 (24.5)
Total	0.732 (100)	0.448 (100)	0.429 (100)	1.461 (100)	1.143 (100)

Note. All variance component estimates were based on 28 diary days.

Source: Cranford et al., 1996

65

Comments

- Although the Bar Exam and Grad Couples studies are both example of IL research, their research foci are very different; the differences can be seen in terms of variance components
- The Bar Exam study examines change as a function of a major stressor, and not surprisingly, its mood variance components due to time are large relative to the Grad Couples study, whose aim was to study couples in everyday life

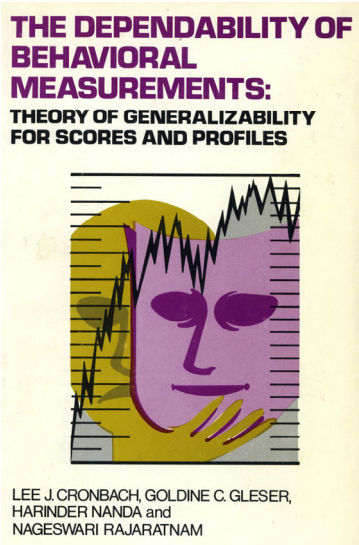
66

Objectives

- Introduction to Cronbach et al.'s Generalizability Theory (GT)
- Various types of reliability indices and their formulae
- Calculating Reliability of Change, R_c

67

Cronbach et al.'s Generalizability Theory (GT)



Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley.*

68

Reliability of Single-day Between-Person Differences

$$R_{1F}(between) = \frac{\sigma^2_{PERSON} + \left[\sigma^2_{PERSON*ITEM} / m\right]}{\sigma^2_{PERSON} + \left[\sigma^2_{PERSON*ITEM} / m\right] + \left[\sigma^2_{ERROR} / m\right]}$$

m = Number of Items

69

Reliability of Average Between-Person Differences

$$R_{MF}(between) = \frac{\sigma^2_{PERSON} + \left[\sigma^2_{PERSON*ITEM} / m\right]}{\sigma^2_{PERSON} + \left[\sigma^2_{PERSON*ITEM} / m\right] + \left[\sigma^2_{ERROR} / l * m\right]}$$

l = Number of Days

m = Number of Items

70

Reliability of Between-Person Differences in Change

$$R_C(change) = \frac{\left[\sigma^2_{PERSON*DAY}\right]}{\left[\sigma^2_{PERSON*DAY}\right] + \left[\sigma^2_{ERROR} / m\right]}$$

m = Number of Items

Calculation for Anxiety Change For Grad Couples Study → .82 = .281/ (.281 + .191/3)

71

Calculation of R_c for Anxiety in Grad Couples

.82 = .281/(.281 + .191/3)

Source of Variance	Anxious Mood (%)	Depressed Mood (%)	Anger (%)	Fatigue (%)	Vigor (%)
$\hat{\sigma}^2_{PERSON}$	0.229 (31.2)	0.072 (16.2)	0.053 (12.4)	0.430 (29.4)	0.356 (31.1)
$\hat{\sigma}^2_{DAY}$	0.003 (0.4)	0.001 (0.3)	0.001 (0.1)	0.023 (2.0)	0.006 (0.5)
$\hat{\sigma}^2_{ITEM}$	0.004 (0.5)	0.013 (2.9)	0.009 (2.2)	0.004 (0.2)	0.036 (3.2)
$\hat{\sigma}^2_{PERSON*DAY}$	0.281 (38.4)	0.162 (36.2)	0.207 (48.3)	0.663 (45.2)	0.388 (33.9)
$\hat{\sigma}^2_{PERSON*ITEM}$	0.024 (3.3)	0.034 (7.5)	0.025 (5.8)	0.080 (5.5)	0.076 (6.7)
$\hat{\sigma}^2_{DAY*ITEM}$	0.0 (0.0)	0.001 (0.2)	0.001 (0.1)	0.00 (0.0)	0.001 (0.1)
$\hat{\sigma}^2_{ERROR}$	0.191 (26.1)	0.165 (36.8)	0.133 (31.1)	0.261 (17.8)	0.280 (24.5)
Total	0.732 (100)	0.448 (100)	0.429 (100)	1.461 (100)	1.143 (100)

Note. All variance component estimates were based on 28 diary days.

72

Calculation of R_C for Anxiety in Grad Couples

TABLE 3: Generalizability Coefficients Computed from Variance Component Estimates

Interpretation	R_{1F} *(Between) Equation 2* — *Reliability (Between Persons) of Measures Taken on the Same Fixed Day*	R_{1R} *(Between) Equation 3* — *Reliability (Between Persons) of Measures When Persons Are Measured on Different Days*	R_{KF} *(Between) Equation 4* — *Reliability (Between Persons) of Average of Measures Taken Over K Fixed Days*	R_C *(Change) Equation 5* — *Reliability of Change (Within Person)*
Bar Exam Study				
Anxious mood	.80	.41	.99	.83
Depressed mood	.73	.51	.99	.62
Anger	.66	.33	.99	.75
Fatigue	.87	.45	.99	.88
Vigor	.76	.43	.99	.76
Graduate Couples Study				
Anxious mood	.79	.41	.99	.82
Depressed mood	.60	.28	.98	.85
Anger	.58	.20	.97	.82
Fatigue	.84	.38	.99	.88
Vigor	.80	.44	.99	.81

NOTE: For the Bar Exam Study, all variance component estimates were based on 35 diary days. For the Graduate Couples Study, all variance component estimates were based on 28 diary days.

.82 = .281/(.281 + .191/3)

(Table is from Cranford et al., 2006)

73

Comments

- Despite their distinct research foci (a majors stressor vs. daily life), GT-based estimates of the reliability of change for daily mood, were in the acceptable range for the Bar and the Grad Couples studies

74

Objectives

- We will look at reliability of change in a new dataset, namely daily positive affect in breast cancer patients.
- We will show how this is done using SPSS VARCOMP and R's *lmer* multilevel modeling program.

75

Typical Data Structure for Multiple Items

person	time	item1	item2	item3	item4
301	1	2	2	3	4
301	2	2	3	3	2
301	3	3	3	2	2
301	4	4	3	3	3
301	5	2	2	2	2
301	6	2	2	1	2
301	7	2	2	1	2
301	8	2	2	2	2
301	9	2	3	2	2
301	10	3	3	3	3
309	1	3	3	–999	1
309	2	3	3	3	3
309	3	3	3	3	1
309	4	3	1	3	1
309	5	1	1	1	1
309	6	3	1	1	1
309	7	1	1	3	1
309	8	1	1	1	1
309	9	1	1	1	1
309	10	1	1	1	1
319	1	3	2	2	2

76

Restructured Dataset

person	time	item	y
301	1	1	2
301	1	2	2
301	1	3	3
301	1	4	4
301	2	1	2
301	2	2	3
301	2	3	3
301	2	4	2
.	.	.	.
.	.	.	.
.	.	.	.
309	1	1	3
309	1	2	3
309	1	3	–999
309	1	4	1
309	2	1	3
309	2	2	3
309	2	3	3
309	2	4	3
.	.	.	.
.	.	.	.

77

SPSS VARCOMP

```
VARCOMP Y BY Person Time Item
  /RANDOM=Person Time Item
  /METHOD= MINQUE(1)
  /DESIGN=Person Time Item Item*Person Item*Time Person*Time.
```

Note: MINQUE = Minimum norm quadratic unbiased estimator.

78

SPSS VARCOMP Output

Variance Estimates

Component	Estimate
Var(person)	.362
Var(time)	-.001[a]
Var(Item)	.049
Var(person * Item)	.190
Var(time * Item)	.005
Var(person * time)	.256
Var(Error)	.299

Dependent Variable: Y
Method: Minimum Norm Quadratic Unbiased Estimation (Weight = 1 for Random Effects and Residual)

a. For the ANOVA and MINQUE methods, negative variance component estimates may occur. Some possible reasons for their occurrence are: (a) the specified model is not the correct model, or (b) the true value of the variance equals zero.

R_C = Var(person*time) / (Var(person*time) + (Var(error)/m))

R_C = .256 / ((.256)+(.299/4)) = 0.773998

79

R: Get Variance Components Using *lmer*

```
##########################################################################################
#Read a csv file containing the data
psychometrics <- read.csv('psychometrics.csv')
##########################################################################################

##########################################################################################
#Create versions of person, time and item that R treats as factors
psychometrics$catitem <- as.factor(psychometrics$item)
psychometrics$catperson <- as.factor(psychometrics$person)
psychometrics$cattime <- as.factor(psychometrics$time)
##########################################################################################


##########################################################################################
#Estimate a model with crossed random effects of person, time, item and their interactions
mod1<-lmer(y ~  1 + (1|catperson) + (1|cattime) + (1|catitem)
     + (1|catperson:cattime) + (1|catperson:catitem) + (1|cattime:catitem), data=psychometrics)
summary(mod1)
##########################################################################################
```

80

R *lmer* Output: R_c Reliability of Change

```
> VarCorr(mod1)[[1]][1,1] #catperson:cattime
[1] 0.2552696
>
> VarCorr(mod1)[[2]][1,1] #catperson:catitem
[1] 0.1901111
>
> VarCorr(mod1)[[3]][1,1] #catperson
[1] 0.36178
>
> VarCorr(mod1)[[4]][1,1] #cattime:catitem
[1] 0.004982553
>
> VarCorr(mod1)[[5]][1,1] #cattime
[1] 0
>
> VarCorr(mod1)[[6]][1,1] #catitem
[1] 0.04859581
>
> sigma(mod1)^2            #Residual
[1] 0.2995423
```

```
> #Use specific components to calculate Rc, the reliability of change
> #Formula: Var(person by time)/(Var(person by time) + (Var(person by item by time) + Var(error))/4)
> Rc<-(VarCorr(mod1)[[1]][1,1])/((VarCorr(mod1)[[1]][1,1]) + (sigma(mod1)^2)/4)
> Rc
[1] 0.7731807
```

81

Comments

- A scale based on four positive affect items for breast cancer patients in daily life was shown to be reliable for the assessment of change using SPSS VARCOMP and R's *lmer*

82

Objective

- Showing how reliability of Change in IL data can also be assessed using an SEM-based approach

83

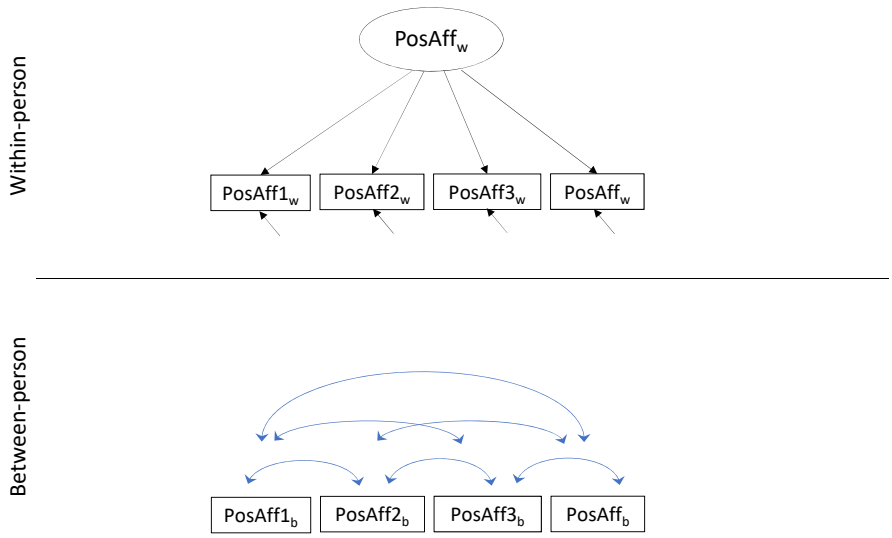
Multilevel Confirmatory Factor Analysis (CFA)

$$\omega = \frac{(\Sigma\lambda_k)^2}{(\Sigma\lambda_k)^2 + \Sigma\psi_k^2}$$

- The numerator of omega is the squared sum of the estimated factor loadings (lambda) of each item *k*.
- The denominator adds to this quantity the sum of the error variances (ψ^2) of each item.
- In essence, this formula reflects the classical test theory conceptualization of reliability where the numerator reflects true score variance and the denominator reflects the total variance of all items.

84

Multilevel CFA Approach: Between-Saturated



85

Multilevel CFA and Omega Syntax in Mplus

```
TITLE:    Psychometrics example (Breast Cancer patients and Daily PA);
DATA:     FILE IS psychometrics.dat;
VARIABLE: NAMES ARE personid time item1 item2 item3 item4; !Positive affect items
          USEVAR = item1 item2 item3 item4 time;
          WITHIN = time;
          CLUSTER = personid;
          missing are all (-999);
ANALYSIS: TYPE = TWOLEVEL;
          PROCESSORS=4;
MODEL:    %WITHIN%
            PAw by
            item1* (a)
            item2 item3 item4 (b-d);
            item1 item2 item3 item4 (e-h);
            PAw@1;
            item1 item2 item3 item4 on time;
          %BETWEEN%
            item1 item2 item3 item4 with item1 item2 item3 item4; !Saturated at between
          MODEL CONSTRAINT:
            new (omegaw);
            omegaw=((a+b+c+d)**2)/(((a+b+c+d)**2) + (e+f+g+h));
OUTPUT:   sampstat stdyx;
```

86

Multilevel CFA and Omega: Mplus Output

```
MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
Within Level

 PAW      BY
    ITEM1              0.467      0.039     11.914      0.000
    ITEM2              0.468      0.051      9.091      0.000
    ITEM3              0.588      0.044     13.217      0.000
    ITEM4              0.502      0.049     10.265      0.000

 ITEM1      ON
    TIME              -0.027      0.016     -1.660      0.097

 ITEM2      ON
    TIME              -0.008      0.020     -0.405      0.686

 ITEM3      ON
    TIME              -0.038      0.020     -1.964      0.050

 ITEM4      ON
    TIME               0.019      0.018      1.080      0.280

 Variances
    PAW                1.000      0.000    999.000    999.000

 Residual Variances
    ITEM1              0.278      0.029      9.442      0.000
    ITEM2              0.359      0.051      7.072      0.000
    ITEM3              0.251      0.032      7.844      0.000
    ITEM4              0.293      0.034      8.730      0.000
```

```
New/Additional Parameters
    OMEGAW             0.776      0.023     33.948      0.000
```

87

Comment

- Although it is very different in approach, a multilevel SEM-based analysis of the four daily positive mood items produces a very similar reliability of change estimate to that obtained from Generalizability Theory.

88

Chapter Summary

- When considering psychometrics for IL data, the focus should be on within-person change versus between-person differences (Coefficient Alpha only focuses on the latter)
- We demonstrated two comparable ways of obtaining an index of reliability of change

89

Recommended Readings

- Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32*, 917-929.
- Shrout, P. E., & Lane, S. P. (2012a). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302–320). New York: Guilford Press.
- Shrout, P. E., & Lane, S. P. (2012b). Reliability. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology: Vol. 1. Foundations, planning, measures and psychometrics* (pp. 643–660). Washington, DC: American Psychological Association.

90

Chapter 7

Modeling of Dyadic Intensive Longitudinal Data

Objectives

- Provide some motivation for collecting and analyzing dyadic intensive longitudinal data
- Define what a dyadic design is
- Review the processes that lead to interdependence in dyadic data

92

Why Would You Want to Collect Dyadic Data?

- So far, we have focused on how IL data can be used to trace changes over time and to model a putative underlying causal process
- The lives of individuals are rarely spent in isolation and an important context of daily life are dyadic relationships (e.g., intimate partners, parents, friends, mentors)
- As we hope to show, dyadic IL studies allow researchers to ask new questions deriving from the nonindependence of the outcomes of dyad members that cannot be answered with IL studies of individuals

93

Dyadic Design

- A dyadic design is where each person is linked to only one other person in a sample and both people are measured on the same variable(s)
- Different from each person rating every other person in a group (round-robin) or one spouse rating the social support received from the other spouse

94

What Makes Dyadic Data Unique?

Dyadic processes that create (statistical) dependencies in a pair of partners' data:

- **Selection** into a relationship as a function of similarity (e.g., in political beliefs) or complementarity (e.g., in dominance vs. submissiveness)
- Direct causal influence by partners on one another's thoughts, feelings, and behaviors (a process known as **interdependence** (e.g., demand-withdrawal)
- **Shared experiences** that create similarity (e.g., friends are part of the same cultural milieu and develop similar tastes in music)

95

What Makes Dyadic Data Unique?

- These processes typically lead dyad members to be more similar to (or more different from) one another than two randomly chosen people from the population
- In statistical terms, the scores from dyad members are said to be **nonindependent** or **interdependent**
- In a manner of speaking, once you know "something" about one partner in a dyad, you tend to know "something" about the other

96

Summary

- For the purposes of this workshop chapter, a dyadic design is one in which both members of a dyad have been measured using the same variables
- Dyadic processes that members engage in will lead to outcomes that tend to be correlated within couple, reflecting nonindependence or interdependence
- Modeling data from dyads needs to be able to account for nonindependence in their data

97

Objectives

- Consider the two broad types of dyads that can exist
- This chapter will be focus on modeling interdependence distinguishable dyads where members can be distinguished on some important characteristic such as gender or role
- Cover two “dyadic” analysis approaches we definitely think you should avoid

98

Types of Dyads

- Two types of dyads
  1. Distinguishable (heterosexual couples, mother-child)
  2. Exchangeable or interchangeable (friendship pairs, identical twins)
- Nonindependence for can be measured:
  1. Pearson correlation for distinguishable dyads
  2. Intraclass correlation for exchangeable dyads
     - ANOVA and “double entry” methods

99

“Dyadic” Approaches to Avoid

- Analyze the data from each partner separately in two models/analyses
  - Results in a loss of information about the very interdependencies that make dyadic diary data unique
  - You really don’t want to do this and claim you are analyzing dyadic data
- Creating a composite “dyad” score by taking the average of the two partner scores
  - Can cover up effects that may be very different for each partner
  - Doing this essentially assumes that you are working with indistinguishable dyads which would need to be tested first

100

Summary

- When analyzing dyadic data, there are distinguishable dyads and exchangeable dyads; while both are interesting to study, we will focus this workshop chapter on analyzing IL data from distinguishable dyads
- If you have IL data from dyads, try to focus on modeling the sources of dyadic interdependence using the approach we describe in this chapter and avoiding doing analyses of each partner separately or averaging the two partners scores

101

Objectives

- The issue of using a 2-level versus a 3-level model for analyzing (intensive) longitudinal dyadic data is a tricky one that we hope to clarify in this section
- Compare the multilevel structures of longitudinal data from partners in dyads to the canonical case of kids in classrooms
- Encourage the intuition of a dyad having a stream of repeatedly measured outcomes for Partner A and a similar stream for Partner B

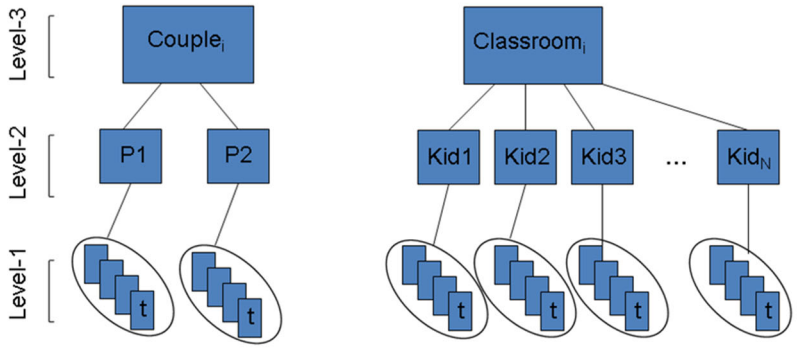
102

Levels of Analysis

- Longitudinal data for distinguishable dyads such as
  - husband-wife
  - parent-child
  - therapist-patient
- Conceptualized as having a 3-level nested structure but a 2-level structure in terms of random effects

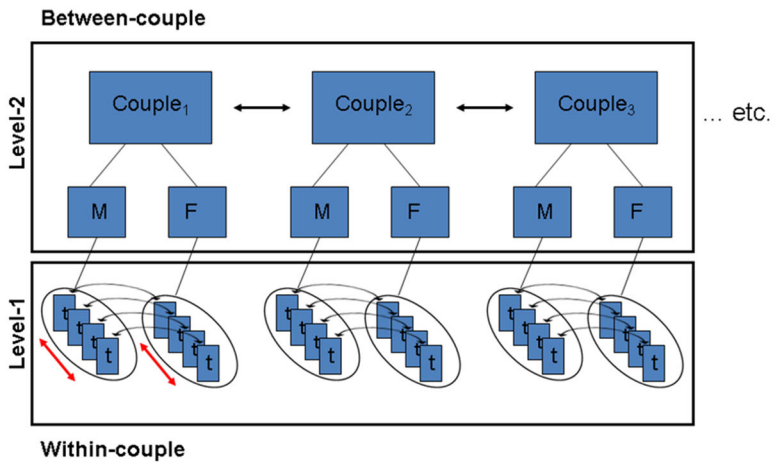
103

Dyads as 3-Level Data Structures



104

A Two-Level Statistical Dyadic Model

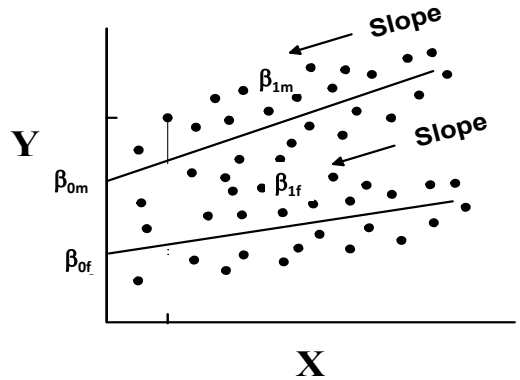


105

Dyad Partners' Data in Parallel

Level 1 Equation (for one couple):

$$Y_i = \beta_{0m} + \beta_{1m}X_i + \beta_{0f} + \beta_{1f}X_i + r_i$$



106

Comments

- While at some level it is very intuitive to approach the analysis of dyadic longitudinal data using a 3-level model ("timepoints nested within individuals nested with couples"), we think that in this case "2" gives you more than "3"
- From our perspective, a 2-level model treating the members' data as multiple (multivariate) outcomes allows you to pick up the lower-level (i.e., within-day across partner residual covariance) and upper-level (i.e., covariances between the partner random effects) sources of dyadic interdependence
- Can be helpful to think about each dyad member's stream of data unfolding in parallel

107

Objectives

- Introduce the example Work-to-Home Stress Spillover dataset
- Review key research questions
- Apply a 2-level dyadic multilevel model and estimate in Mplus and R
- Interpret results and cover what could go in an APA-style write-up

108

Example Dataset

- Work-to-Home Stress Spillover Study
- N = 100 Dual income couples; N = 21 Days
- End of workday & bedtime assessments
- 2 sets of individual stacked per couple
- 100 x 21 x 2 = 4200 data lines

109

Example Dataset

Daily measures for male and female partners separately:

- Number of work stressors (by end of work day)
  - Male mean = 3.01, male sd = 1.03, range: 0-7
  - Female mean = 2.97, male sd = 1.03, range: 0-6
- Relationship dissatisfaction (at bedtime)
  - Male mean = 5.08, male sd = 1.35, range: 0-10
  - Female mean = 4.64, female sd = 1.40, range: 0-10

110

Research Questions

- How much work-to-home stress carryover for the typical wife? For the typical husband?
- How much do husbands and wives differ in this effect?
- To what degree are husbands and wives linked in their work-to-home stress carryover?

111

Time-Varying Predictors: Within- vs. Between

- This applies to dyadic as well as non-dyadic data
- Need to separate of within-person and between-person effects in longitudinal predictors
- An individual's daily number of work stressors fluctuates around his/her average number of work stressors and this average differs from person to person

112

Time-Varying Predictors: Within- vs. Between

- Split the original predictor variable (*wrkstrs*) into two predictor variables:
  - the stable between-persons mean for each person across their diary days (mean*wrkstrs*)
  - the deviations from this mean representing within-person change (*wrkstrsw*)
- This allows the coefficient for *wrkstrsw* to be interpreted as a "pure" within-persons effect

113

Centering Level-1 Predictors

Raw X	X_{ij}	**wrkstrs**
Grand Centered X	$\breve{X}_{ij}$	**wrkstrsc**
Group Mean of Grand-Centered X	$\breve{X}_{.j}$	**wrkstrscb**
Grand Centered & Group Centered X	$\breve{X}_{ij} - \breve{X}_{.j}$	**wrkstrscw**

114

Variables: Daily work-to-home stress spillover

- id: Couple id
- male: Male dummy variable (0/1)
- female: Female dummy variable (0/1)
- time7c: Centered time coding
- reldis: Relationship dissatisfaction (0-10 Daily score)
- wrkstrs : Work stressors (count)

115

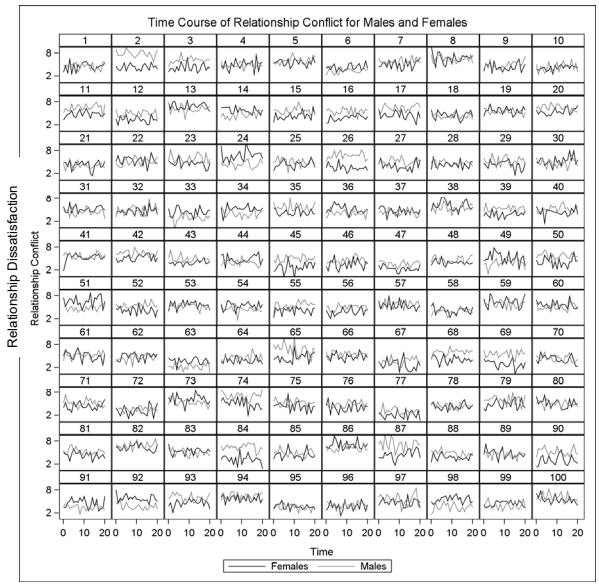
Data Structure (Stacked)

Data structure (first 7 days for 1 couple) for intensive longitudinal dataset on dyads

couple	Personid	time	time7c	gender	female	male	reldis	wrkstrs	wrkstrsc	wrkstrscb	wrkstrscw
1	1	0	-1.50	1	1	0	3.03	3	.010	-.324	.333
1	1	1	-1.36	1	1	0	4.62	3	.010	-.324	.333
1	1	2	-1.21	1	1	0	2.85	3	.010	-.324	.333
1	1	3	-1.07	1	1	0	6.40	4	1.010	-.324	1.333
1	1	4	-.93	1	1	0	2.54	1	-1.990	-.324	-1.667
1	1	5	-.79	1	1	0	5.16	2	-.990	-.324	-.667
1	1	6	-.64	1	1	0	2.70	3	.010	-.324	.333
…	…	…	…	…	…	…	…	…	…	…	…
1	2	0	-1.50	0	0	1	4.46	3	.010	-.133	.143
1	2	1	-1.36	0	0	1	4.88	3	.010	-.133	.143
1	2	2	-1.21	0	0	1	4.58	3	.010	-.133	.143
1	2	3	-1.07	0	0	1	4.49	1	-1.990	-.133	-1.857
1	2	4	-.93	0	0	1	5.04	3	.010	-.133	.143
1	2	5	-.79	0	0	1	4.87	3	.010	-.133	.143
1	2	6	-.64	0	0	1	5.63	2	-.990	-.133	.857
…	…	…	…	…	…	…	…	…	…	…	…

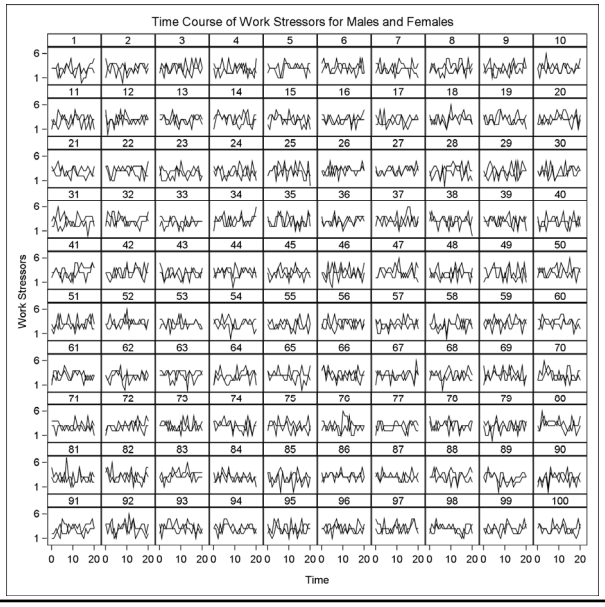
116

First Look: Time Course of DV



117

First Look: Time Course of IV



118

Basic Multilevel Model for Dyadic Process

$$YM_{ij} = \beta_{0jM} + \beta_{1jM}(\check{X}M_{ij} - \check{X}M_{.j}) + \beta_{2jM}\check{T}_{ij} + \varepsilon M_{ij}$$

$$YF_{ij} = \beta_{0jF} + \beta_{1jF}(\check{X}F_{ij} - \check{X}F_{.j}) + \beta_{2jF}\check{T}_{ij} + \varepsilon F_{ij}$$

$$\beta_{0jM} = \gamma_{00M} + \gamma_{01M}\check{X}M_{.jM} + u_{0jM}$$

$$\beta_{0jF} = \gamma_{00F} + \gamma_{01F}\check{X}F_{.jF} + u_{0jF}$$

$$\beta_{1jM} = \gamma_{10M} + u_{1jM}$$

$$\beta_{1jF} = \gamma_{10F} + u_{1jF}$$

$$\beta_{2jM} = \gamma_{20M}$$

$$\beta_{2jF} = \gamma_{20F}$$

119

Basic Multilevel Model for Dyadic Process Data

$$YM_{ij} = \gamma_{00M} + \gamma_{01M}\check{X}M_{.j} + \gamma_{10M}(\check{X}M_{ij} - \check{X}M_{.j}) + \gamma_{20M}\check{T}_{ij}$$

$$+ u_{0jM} + u_{1jM}(\check{X}M_{ij} - \check{X}M_{.j}) + \varepsilon M_{ij}$$

$$YF_{ij} = \gamma_{00F} + \gamma_{01F}\check{X}F_{.j} + \gamma_{10F}(\check{X}F_{ij} - \check{X}F_{.j}) + \gamma_{20F}\check{T}_{ij}$$

$$+ u_{0jF} + u_{1jF}(\check{X}F_{ij} - \check{X}F_{.j}) + \varepsilon F_{ij}$$

$$U = \begin{bmatrix} u_{0jM} \\ u_{0jF} \\ u_{1jM} \\ u_{1jF} \end{bmatrix}_{4\times 1} \quad Cov(U) = T = \begin{bmatrix} \tau_{00MM} & \tau_{00MF} & \tau_{01MM} & \tau_{01MF} \\ \tau_{00FM} & \tau_{00FF} & \tau_{01FM} & \tau_{01FF} \\ \tau_{10MM} & \tau_{10MF} & \tau_{11MM} & \tau_{11FM} \\ \tau_{10FM} & \tau_{10FF} & \tau_{11FM} & \tau_{11FF} \end{bmatrix}_{4\times 4}$$

120

Lower-level Residual Structure for Each Dyad

$$Cov\left[\begin{bmatrix} M_1 \\ \varepsilon M_2 \\ \vdots \\ \varepsilon M_{21} \\ \\ \varepsilon F_1 \\ \varepsilon F_2 \\ \vdots \\ \varepsilon F_{21} \end{bmatrix}_{42\times 1}\right] = \begin{bmatrix} \sigma_M^2 & \sigma_{MF} \\ \sigma_{FM} & \sigma_F^2 \end{bmatrix}_{2\times 2} \otimes \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{20} & \rho^{21} \\ \rho & 1 & \rho & \ddots & & \rho^{20} \\ \rho^2 & \rho & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \rho & \rho^2 \\ \rho^{20} & & \ddots & \rho & 1 & \rho \\ \rho^{21} & \rho^{20} & \cdots & \rho^2 & \rho & 1 \end{bmatrix}_{21\times 21}$$

121

Dyadic SPSS MIXED with AR1

MIXED relconf WITH male female time7c wrkstrscw wrkstrscb

/FIXED=female male male*time7c female*time7c male*wrkstrscw female*wrkstrscw male*wrkstrscb female*wrkstrscb | NOINT SSTYPE(3)

/METHOD=REML

/PRINT=G SOLUTION TESTCOV

/RANDOM=female male male*wrkstrscw female*wrkstrscw | SUBJECT(coupleid) COVTYPE(UN)

/REPEATED=time | SUBJECT(coupleid*gender) COVTYPE(ar1).

Note: Does not allow for within-couple residual covariance or male & female AR(1)

122

Dyads in Mplus with AR1 residual structure

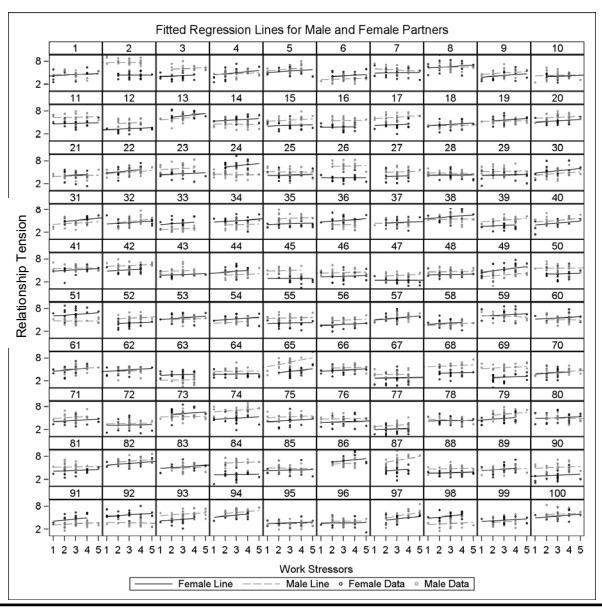
Mplus - [v8-dyads-ar1-residual.inp]

File Edit View Mplus Plot Diagram Window Help

```
TITLE:    Extension of Dyads (Work-to-Home Stress Example);
DATA:     FILE IS dyadsdata.dat;
VARIABLE: NAMES ARE coupleid personid time time7c freldis fwrkstrs fwrkstrsc
          fwrkstrscb fwrkstrscw mreldis mwrkstrs mwrkstrsc mwrkstrscb mwrkstrscw;
          USEVAR =  time7c freldis mreldis fwrkstrs mwrkstrs;
          WITHIN =  time7c;
          CLUSTER = coupleid;
          LAGGED = freldis(1) mreldis(1);
DEFINE:   CENTER fwrkstrs mwrkstrs (GRANDMEAN);
ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR=bayes;
          PROCESSORS = 4;
          BITERATIONS = (5000);
MODEL:    %WITHIN%
           cf | freldis ON fwrkstrs; ! Latent separation of fwrkstrs
           cm | mreldis ON mwrkstrs; ! Latent separation of mwrkstrs
           freldis^ ON freldis^1; !This is to specify AR(1) error structure
           mreldis^ ON mreldis^1; !This is to specify AR(1) error structure
           freldis mreldis ON time7c;
           freldis (freldisv1); mreldis (mreldisv1);
           freldis WITH mreldis (covlev1);
          %BETWEEN%
```

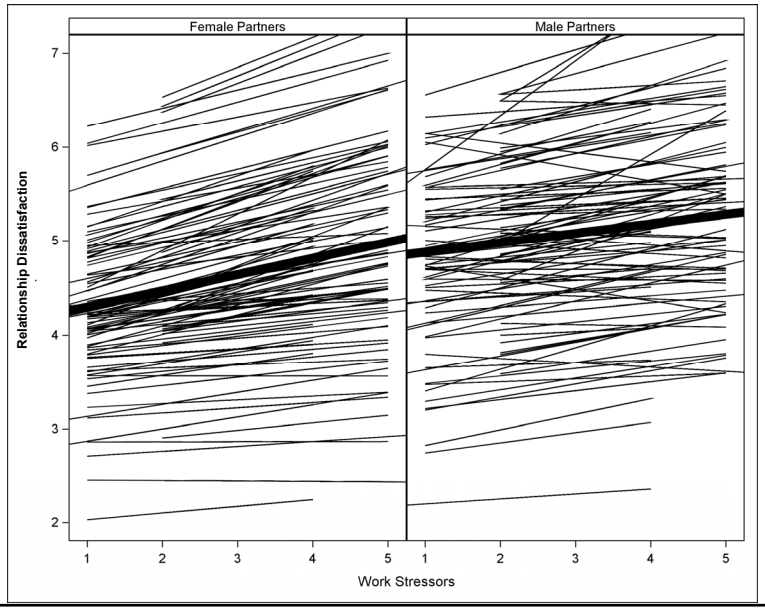
V8-dyads-ar1-residual.inp

Scatterplot and Fitted Lines



124

Spaghetti Plot of Dyadic Example



125

Table of Dyadic Results in APA Format

Table 1. Parameter Estimates for Dyadic Multilevel Model of Evening Relationship Dissatisfaction as a Function of Number of Work Stressors for Male and Female Dyad Partners

					CI_{95}	
Fixed effects (intercepts, slopes)	Estimate	(*SE*)	t^a	p^b	Lower	Upper
M_Intercept	5.09	(0.10)	48.98	<.001	4.88	5.29
F_Intercept	4.65	(0.10)	46.98	<.001	4.45	4.84
M_Work stressor Slope	0.11	(0.03)	4.27	<.001	0.06	0.16
F_Work stressor Slope	0.16	(0.03)	6.37	<.001	0.11	0.21
M_Time Slope	0.01	(0.02)	0.51	.61	–0.03	0.06
F_Time Slope	–0.02	(0.03)	–0.97	.33	–0.08	0.03
Mean M_Work Stressor Slope	–0.14	(0.44)	–0.33	.75	–1.01	0.73
Mean F_Work Stressor Slope	0.62	(0.43)	1.45	.15	–0.23	1.47
					$CI_{95}{}^c$	
Random effects ([co-]variances)	Estimate	(*SE*)	z	p^b	Lower	Upper
Level-2 (between-couples)[d]						
M_Intercept	1.03	(0.15)	6.76	<.001	0.79	1.41
F_Intercept	0.92	(0.14)	6.66	<.001	0.70	1.27
M_Work Stressor Slope	0.028	(0.009)	3.10	.001	0.02	0.06
F_Work Stressor Slope	0.016	(0.009)	1.79	.04	0.01	0.07
M-F Intercept covariance	0.26	(0.11)	2.39	.02	0.05	0.47
M-F Slope covariance	0.011	(0.007)	1.57	.12	-0.003	0.024
Level-1 (within-couples)						
M_Residual	0.761	(0.025)	30.81	<.001	0.72	0.81
F_Residual	1.00	(0.03)	30.81	<.001	0.94	1.06
M–F Residual covariance	0.064	(0.020)	3.21	.001	0.02	0.10
Autocorrelation	0.010	(0.017)	0.60	.55	-0.02	0.04

Note: N = 100 couples, 21 days. M, male partner; F, female partner.

[a]Degrees of freedom are 98 for tests of intercepts and mean work stressors and 99 for work stressors and time slopes.

[b]All p-values are two-tailed except in the case of variances, where one-tailed p-values are used (because variances are constrained to be non-negative).

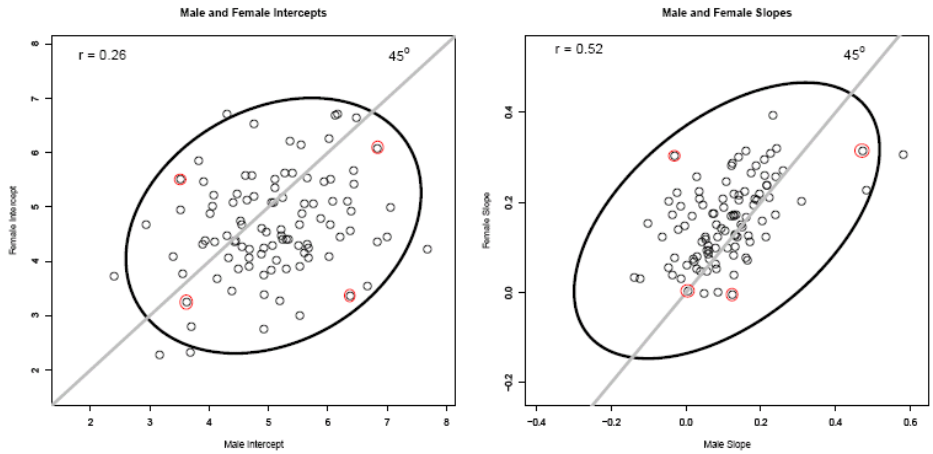
[c]Confidence intervals for variances were computed using the Satterthwaite method (see Littell, Milliken, Stroup, Wolfinger, & Schabenberger, 2006).

[d]Covariances among male and female partner intercepts and slopes were estimated but, for the sake of brevity, are not shown.

126

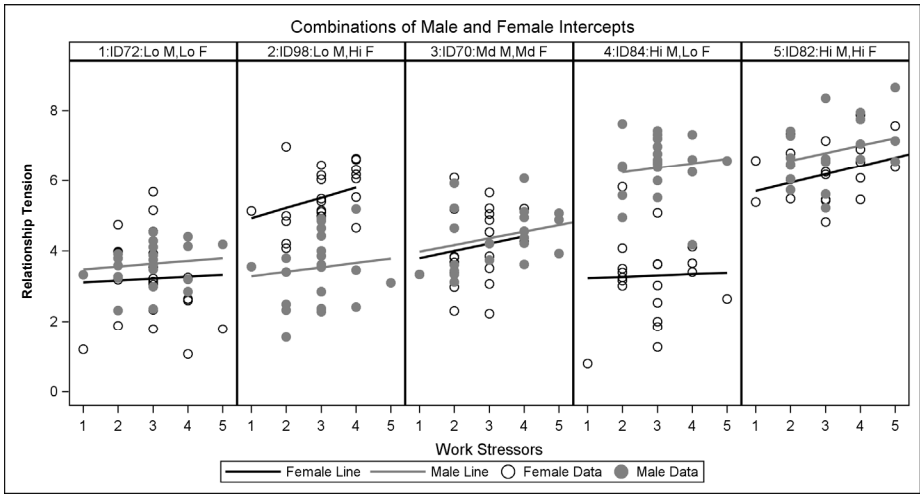
Visualizing the Random Effects

Male-female random effect correlations: EB estimates for the sample; ML estimates of population ellipses



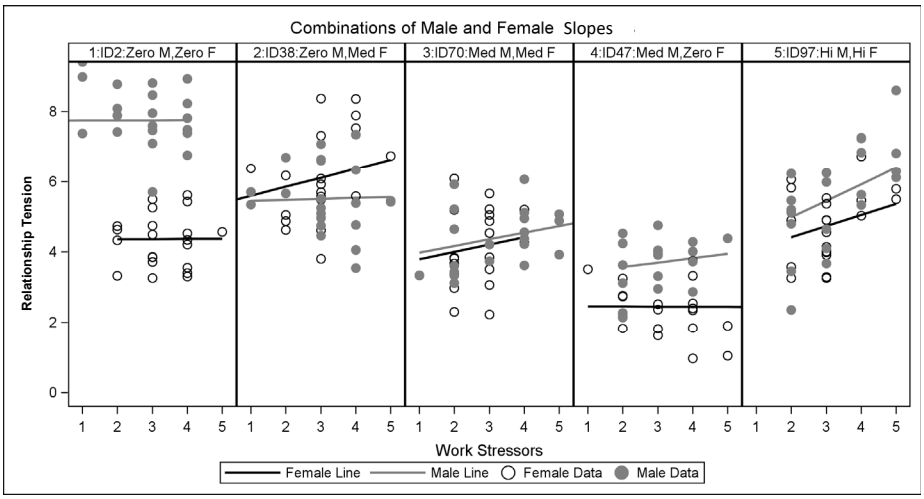
127

Combinations of Zero, Medium, and High Male and Female Intercepts



128

Combinations of Zero, Medium, and High Male and Female Intercepts



129

Example Write-up

In summary, the major hypothesis to be tested is whether the number of daily stressors at work predicts degree of end-of-the-day relationship dissatisfaction for both husbands and wives.

METHOD

Design

Couples were recruited using advertisements in local newspapers. Each partner completed an initial set of background questionnaires and then completed two online daily diaries for each of 21 consecutive days. The first diary was completed within an hour of the end of the workday and assessed experiences in the workplace, including the number of work stressors experienced that day. The second diary was completed later that day, within an hour of going to bed, and assessed feelings of dissatisfaction and tension in the relationship with the partner.

Sample

The sample consisted of male and female partners from 100 dual-career married or cohabiting heterosexual couples from a metropolitan city who volunteered to participate in a study of "work and family experiences." Male partners were, on average, 30.9 years old (*SD* = 5.1) and female partners were, on average, 29.4 years old (*SD* = 4.7). Sample composition: 54.2% were European American, 28.3% were Hispanic American, 9% were African American, 5.4% were listed in other categories (e.g., Caribbean descent), and 2.1% were Asian American and Pacific Islanders. Couples were cohabiting for an average of 5.2 years (*SD* = 4.4). The median number of children in the home was two.

130

Additional Issues

- Getting the full AR1 residual error structure (UN@AR1) with a two equation system approach (Mplus and brms in R)
- Actor and Partner effects

131

Comments

- Dyadic process modeling is like doing multilevel process modeling with two correlated outcomes
- Dyadic nonindependence is modeled by 1) covariances of the husband and wife between-couple random effects and 2) daily within-couple residual covariation
- Single outcome multilevel modeling packages (e.g., SPSS, SAS, lme4 in R) require the dyadic data to be stacked and this can become unwieldy so we focus on analysis in Mplus and brms in R that can handle multiple outcomes
- Can be extended to full actor-partner interdependence modeling framework (which we turn to next)

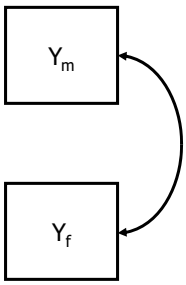
132

Objectives

- Introduce the basic actor-partner interdependence modeling (APIM) framework as applied to longitudinal data
- Define actor and partner effects
- Demonstrate the importance of modeling both outcomes simultaneously to take full advantage of APIM

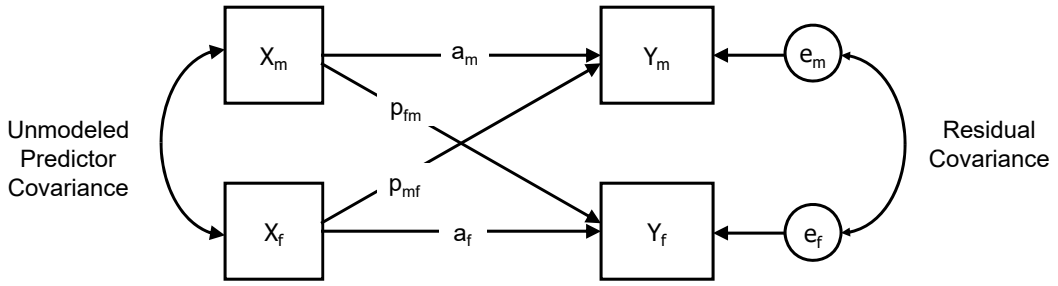
133

Initial Dyadic Covariance



134

Actor-Partner Interdependence Model (APIM)



a = Actor effect (Intrapersonal)

p = Partner effect (Interpersonal)

A basic video introduction to APIM can be found here:
https://cancercontrol.cancer.gov/brp/hbrb/flashe-webinar.html

135

Actor and Partner Effects (Kenny, 1996)

- Actor effect: When a person's score on a predictor variable has an influence on that same person's score on an outcome variable
- Partner effect: When person's score on a predictor has an influence on his/her partner's score on an outcome
- Note that actor effects and partner effects are typically interpreted controlling for the other

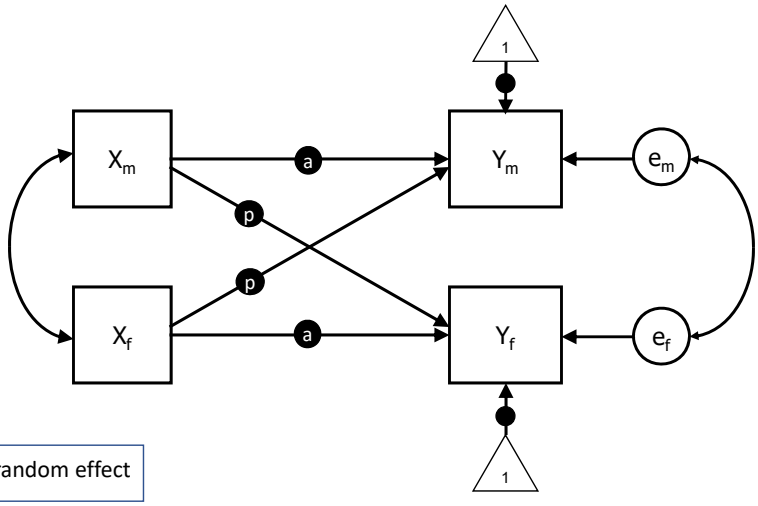
136

Modeling Both Dyadic Outcomes Simultaneously

- We suggest modeling both partners' outcomes simultaneously, otherwise we cannot…
  - Test whether both partner's actor effects or both partner's partner effects are different
  - Test whether actor and partner effects are different from each other
  - Constrain actor or partner paths to obtain a pooled estimate (which also has greater power)
  - Represent unmodeled sources of interdependence at the lower-and upper-levels

137

Multilevel Dyadic Process APIM



● represents a random effect

138

Comments

- The APIM is a way of explaining the dyadic covariation between dyad members' outcomes
- The dyadic data example from this chapter can be extended to be a "fully" dyadic model by adding partner effects (a la APIM)
- The partner effect is by convention labeled based on the outcome, so the male partner effect is the male outcome regressed on the female predictor (p_{mf}) (See Garcia, Kenny, & Ledermann, 2015)

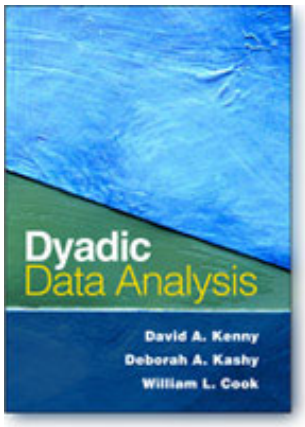
139

Chapter Summary

- Dyadic intensive longitudinal data allow investigators to ask important new questions that cannot be addressed in intensive longitudinal studies of individuals
- There are challenges of modeling dyadic interdependence at the within- and between-dyad levels
- What is new about the analysis model is that it treats the data from dyad partners as a paired longitudinal series where the pairing is modeled
  - At each time point in the series, such that, for example, at time points where one partner is especially dissatisfied, the other partner is also
  - At the level of the statistical parameters for each series, such that, for example, if one partner has a relatively steep work-stress-to-dissatisfaction slope, the other partner is likely to have one also

140

"Bible" of Dyadic Data Analysis



2006 Guilford Press
Methodology in the Social Sciences Series

141

Recommended Readings

- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York: Guilford Press.
- Laurenceau, J.-P., & Bolger, N. (2012). Analyzing diary and intensive longitudinal data from dyads. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 407–422). New York: Guilford Press.
- Gonzalez, R., & Griffin, D. (2012). *Dyadic data analysis.* In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbooks in psychology®. APA handbook of research methods in psychology, Vol. 3. Data analysis and research publication* (p. 439–450). American Psychological Association.

142

Chapter 8

Multilevel Within-Person Mediation

Objectives

- Show why mediation analysis is a central concern of scientific investigation
- Show the importance of Sewall Wright's method of path analysis in allowing one to assess mediational effects

144

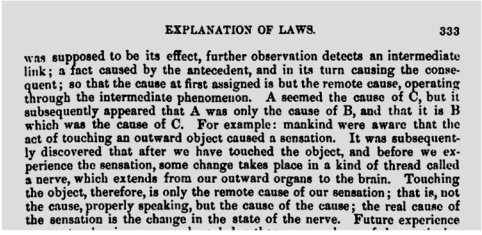
Assessing Mediation: A Central Aim of Investigation

- Identify some interesting phenomenon or effect
  - Social loafing
  - Positive illusions
  - Self-fulfilling prophecies
- Then identify intervening processes that can explain the effect
  - Diffusion of responsibility
  - Self-esteem maintenance
  - Behavioral confirmation

145

Explaining a Phenomenon: Postulate a Causal Chain

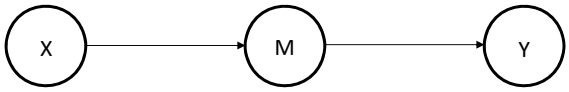
- Discussed in J. S. Mill, *A System of Logic*, 1843



EXPLANATION OF LAWS. 333

was supposed to be its effect, further observation detects an intermediate link; a fact caused by the antecedent, and in its turn causing the consequent; so that the cause at first assigned is but the remote cause, operating through the intermediate phenomenon. A seemed the cause of C, but it subsequently appeared that A was only the cause of B, and that it is B which was the cause of C. For example: mankind were aware that the act of touching an outward object caused a sensation. It was subsequently discovered that after we have touched the object, and before we experience the sensation, some change takes place in a kind of thread called a nerve, which extends from our outward organs to the brain. Touching the object, therefore, is only the remote cause of our sensation; that is, not the cause, properly speaking, but the cause of the cause; the real cause of the sensation is the change in the state of the nerve. Future experience

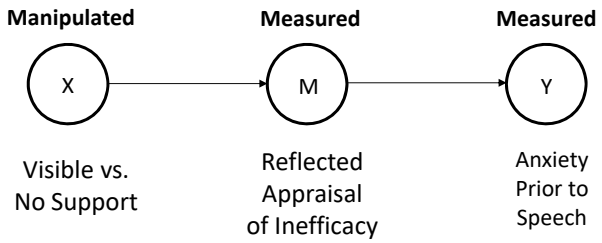
- A change in X causes a change in M
- A change in M causes a change in Y.



146

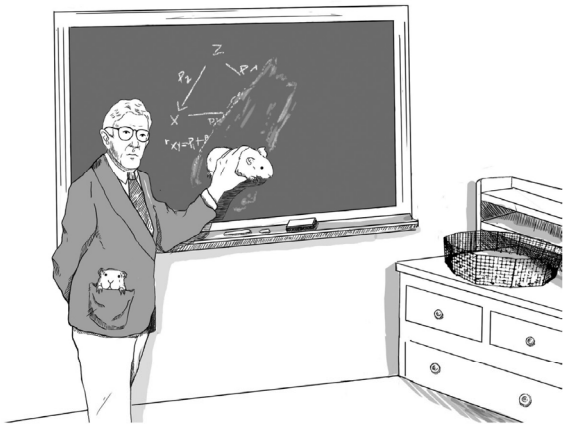
Manipulate or Measure the Mediator?

- **Bolger, N., & Amarel, D. (2007).** Effects of support visibility on adjustment to stress: Experimental evidence. *Journal of Personality and Social Psychology, 92*, 458-475.



147

Sewall Wright: Inventor of Path Analysis



Taken from Pearl and Mackenzie (2018): *The Book of Why*. Basic Books

FIGURE 2.6. Sewall Wright was the first person to develop a mathematical method for answering causal questions from data, known as path diagrams. His love of mathematics surrendered only to his passion for guinea pigs. (*Source:* Drawing by Dakota Harr.)

148

The Very First Path Diagram

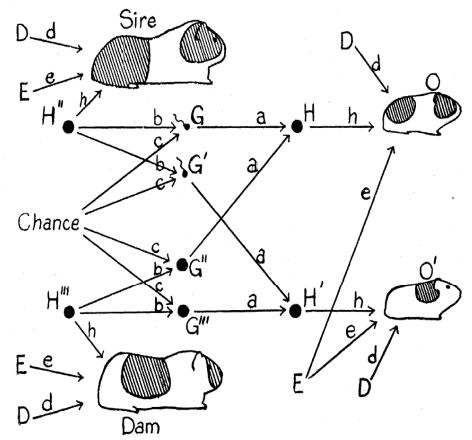


FIG. 5.

Diagram illustrating the casual relations between litter mates (O, O′) and between each of them and their parents. H, H′, H″, H‴ represent the genetic constitutions of the four individuals, G, G′, G″, and G‴ that of four germ cells. E represents such environmental factors as are common to litter mates. D represents other factors, largely ontogenetic irregularity. The small letters stand for the various path coefficients.

Wright, Sewall (1920). The relative importance of heredity and environment in determining the piebald pattern of guinea-pigs. *Proceedings of the National Academy of Sciences of the United States of America, 6, 320-332.*

149

Comments

- The identification of mediating processes has long been understood to be a central part of theory specification and testing
- Path modeling, essentially systems of linked linear equations, was originally invented to study intergenerational genetic transmission
- It is now by far the most common statistical approach to assessing mediation

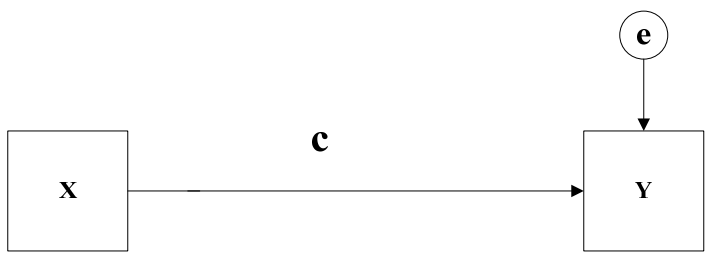
150

Objectives

- Review linear models of single-level mediation
- Show the key equations and assumptions.

151

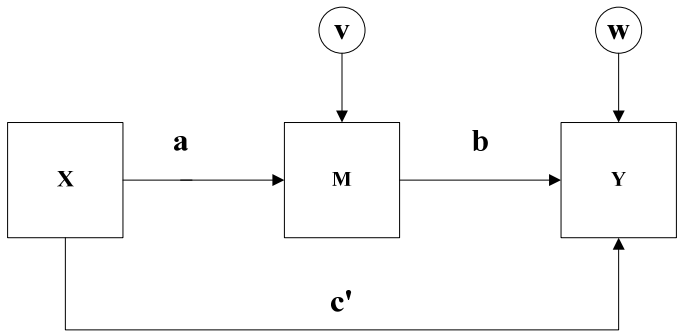
Total Effect



$$Y_i = d_0 + cX_i + \varepsilon_i$$

152

Direct (c') and Indirect (ab) Effects



$$M_i = d_1 + aX_i + v_i$$

$$Y_i = d_2 + c'X_i + bM_i + w_i$$

$$c = c' + ab$$

153

Additional Assumptions

- Y does not affect M (no reciprocal effect)
- No third variable causing both M and Y
- X and M are free of measurement error

154

Objectives

- Introduce within-person mediation, an idea that requires the use of multilevel structural equation modeling
- Show how random effects enter into the key equations and assumptions.
- Emphasize the ways in which within-person mediation differs from single-level mediation, especially with respect to estimating intervening pathways.

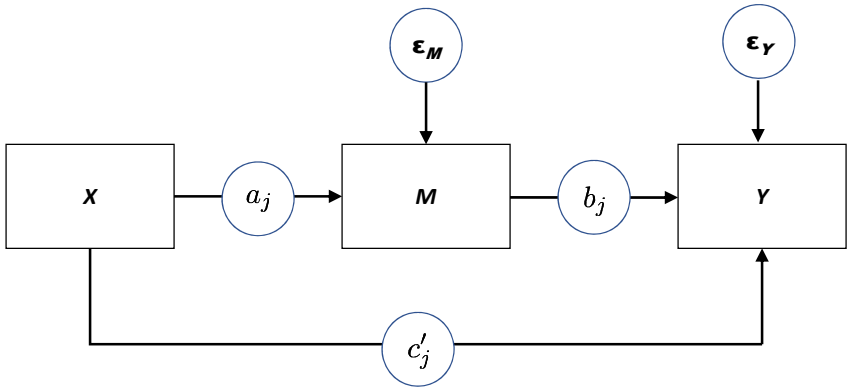
155

Within-person Mediation

- Each participant is intensively assessed on a set of variables (e.g., X, M, and Y) over time
- We can conceive of each participant having his/her own mediation effect
- This permits us to estimate an
  - (i) **average within-person mediated effect** as well as
  - (ii) **between-person heterogeneity** around that average

156

Within-person Mediation: Path Diagram



157

Mediation in Multilevel Models: Level-1

$$Y_{ij} = d_{0j} + c_j X_{ij} + \varepsilon_{ij}$$

$$M_{ij} = d_{1j} + a_j X_{ij} + v_{ij}$$

$$Y_{ij} = d_{2j} + c_{j'} X_{ij} + b_j M_{ij} + w_{ij}$$

158

Level 2 Equations

$$c_j = c + u_{0j}$$

$$a_j = a + u_{1j}$$

$$b_j = b + u_{2j}$$

$$c'_j = c' + u_{3j}$$

159

Kenny, Korchmaros, & Bolger (2003)

Lower Level Mediation in Multilevel Models

David A. Kenny and
Josephine D. Korchmaros
University of Connecticut

Niall Bolger
New York University

Multilevel models are increasingly used to estimate models for hierarchical and repeated measures data. The authors discuss a model in which there is mediation at the lower level and the mediational links vary randomly across upper level units. One repeated measures example is a case in which a person's daily stressors affect his or her coping efforts, which affect his or her mood, and both links vary randomly across persons. Where there is mediation at the lower level and the mediational links vary randomly across upper level units, the formulas for the indirect effect and its standard error must be modified to include the covariance between the random effects. Because no standard method can estimate such a model, the authors developed an ad hoc method that is illustrated with real and simulated data. Limitations of this method and characteristics of an ideal method are discussed.

160

Decomposition of Total Effect for Average Person

- We need a crucial extra term for the effect decomposition: $\sigma_{a_j b_j}$

$$c = c' + ab + \sigma_{a_j b_j}$$

161

Comments

- Within-person mediation allow for individual-specific mediating processes: There is mediation for the typical person and there is (the potential for) heterogeneity across persons
- Decomposition of average causal effects becomes more complex in the sense that a new intervening effect is called for, namely, the covariance of individual differences in the X-to-M and M-to-Y paths

162

Objectives

- Introduce the example research question and dataset on work stress spillover into home
- Estimate a multilevel mediation model using the data
- Interpret the average mediation and between-person heterogeneity
- Show that the additional covariance terms in the effect decomposition equation identifies a previously unknown type of indirect effect

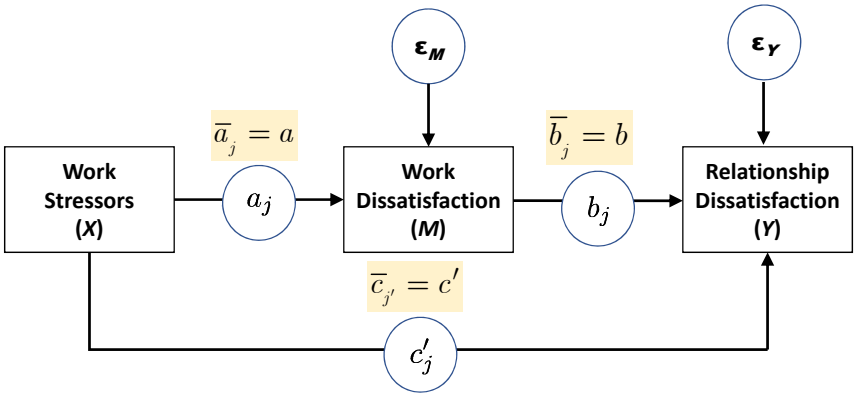
163

Example Dataset: Work Stress Spillover

- N = 100 Female partners
- N = 21 Days
- Daily measures:
  - Number of work stressors (at end of work day)
  - Work dissatisfaction (at end of work day; 0-10 scale)
  - Relationship dissatisfaction (at bedtime; 0-10 scale)

164

Work Stress Spillover: Fixed & Random Effects



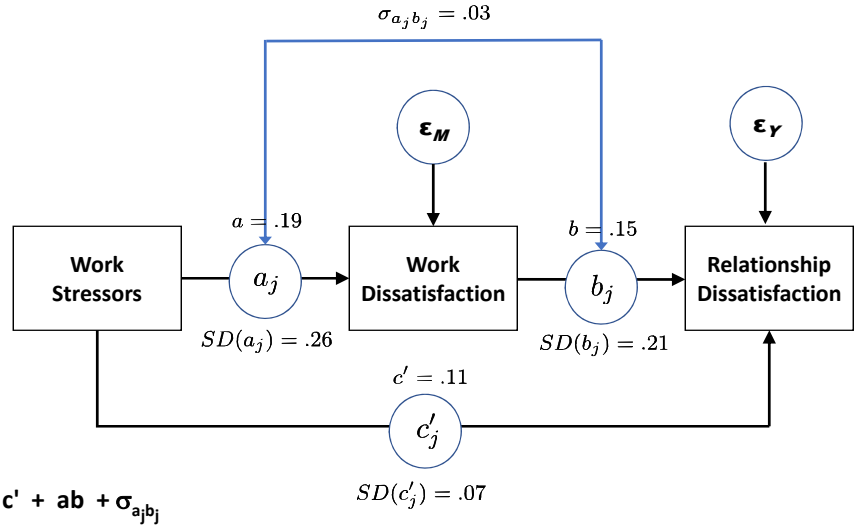
$$M_{ij} = \beta_{00j} + a_j X_{ij} + \varepsilon_M$$

$$Y_{ij} = \beta_{10j} + c'_j X_{ij} + b_j M_{ij} + \varepsilon_Y$$

$$c = c' + ab + \sigma_{a_j b_j}$$

165

Work Stress Spillover: Parameter Estimates



Total effect = c' + ab + $\sigma_{a_j b_j}$

0.16 = 0.11 + (0.19*0.15) + 0.03

166

Effect Decomposition (for the Average Person)

- 36% (0.06/0.16) of the relationship between work stressors and marital dissatisfaction for the average person **is explained** by work dissatisfaction
- Half of that amount (0.03) is accounted for by a conventional indirect effect (0.19*0.15)

Total effect = c' + ab + $\sigma_{a_j b_j}$

0.16 = 0.11 + (0.19*0.15) + 0.03

- The other half is accounted for by the $a_j b_j$ covariance (.031/0.059 = .52)
- This is a form of moderation that helps explain the effect of work stressors for the average person

167

Conducting 1 → 1 → 1 Multilevel Mediation

- Restructuring of the data involving stacking is necessary in these kinds of multilevel regression programs
  - Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods, 11, 142-163.*

168

Data Restructuring for Multilevel Analysis (N=4, T=4)

ID	Day	X	M	Y
1	1	3	7	11
1	2	5	9	12
1	3	4	8	11
1	4	3	7	11
2	1	5	8	8
2	2	5	7	8
2	3	5	7	9
2	4	5	7	9
3	1	4	7	9
3	2	4	7	9
3	3	4	8	9
3	4	4	8	9
4	1	7	9	11
4	2	5	8	10
4	3	6	9	11
4	4	7	11	11

ID	Day	X	M	DV	YVAR	MVAR	DVCODE
1	1	3	7	7	0	1	m
1	1	3	7	11	1	0	y
1	2	5	9	9	0	1	m
1	2	5	9	12	1	0	y
1	3	4	8	8	0	1	m
1	3	4	8	11	1	0	y
1	4	3	7	7	0	1	m
1	4	3	7	11	1	0	y
2	1	5	8	8	0	1	m
2	1	5	8	8	1	0	y
2	2	5	7	7	0	1	m
2	2	5	7	8	1	0	y
2	3	5	7	7	0	1	m
2	3	5	7	9	1	0	y
2	4	5	7	7	0	1	m
2	4	5	7	9	1	0	y
3	1	4	7	7	0	1	m
3	1	4	7	9	1	0	y
3	2	4	7	7	0	1	m
3	2	4	7	9	1	0	y
3	3	4	8	8	0	1	m
3	3	4	8	9	1	0	y
3	4	4	8	8	0	1	m
3	4	4	8	9	1	0	y
4	1	7	9	9	0	1	m
4	1	7	9	11	1	0	y
4	2	5	8	8	0	1	m
4	2	5	8	10	1	0	y
4	3	6	9	9	0	1	m
4	3	6	9	11	1	0	y
4	4	7	11	11	0	1	m
4	4	7	11	11	1	0	y

169

Comments

- Using traditional multilevel programs requires restructuring (stacking) the data
- Because this is cumbersome to do, we recommend the approach in the next section

170

Objectives

- Use multilevel SEM to represent a within-person mediation model as a 2 equation system

171

1→1→1 Multilevel Mediation in Mplus

```
TITLE:   Within-subject mediation(all between-subject deviated variables);
DATA:    FILE IS mediatedata.dat;
VARIABLE: NAMES ARE id time timec freldis fwkdis fwkstr
fwkstrc fwkdisc freldisc fwkstrcb fwkdiscb freldiscb
fwkstrcw fwkdiscw freldiscw x m y;
        USEVAR = id timec x m y;
        WITHIN = timec x m y;
        CLUSTER = id;
ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR=ml;
```

```
MODEL:   %WITHIN%
        cp | y ON x;
        a | m ON x;
        b | y ON m;
        y m ON timec;
        [m@0 y@0];
       %BETWEEN%
        cp WITH a b;
        [a] (ma); a (vara); [b] (mb); b (varb);
        [cp] (mcp);
        a WITH b (covab);
MODEL CONSTRAINT:
        NEW MED TE PME CORR;
        MED=ma*mb+covab;
        TE=ma*mb+covab+mcp;
        PME=(ma*mb+covab)/(ma*mb+covab+mcp);
        CORR=covab/sqrt(vara*varb);
OUTPUT:    sampstat cinterval;
```

172

Possible Extensions

- Explaining between person variation in the **a**, **b**, and **c'** random effects
- Dealing with measurement error: Incorporating measurement models for X, M, and Y
- Mediating processes taking dyadic dependence into account
- Handling temporal separation of X, M, and Y
- Modern literature on causal inference: Rubin, Pearl, Robbins & Greenland

173

Chapter Summary

- In this lecture we introduced statistical theory behind systems of equations approaches to assessing mediation
- We showed how the theory works for conventional, single level mediation
- Then we showed the additional complexities that come when you use it in a multilevel context with intensive longitudinal data
- We showed you how it works for an empirical example of stress spillover from work to home

174

Recommended Readings

- Cole & Maxwell (2003; Journal of Abnormal Psychology)
- Krull & MacKinnon (2001; Multivariate Behavioral Research)
- Kenny, Korchmaros, & Bolger (2003; Psychological Methods)
- Bauer, Preacher, & Gil (2006; Psychological Methods)

175

Chapter 9

Modeling Extensions and Future Directions for Intensive Longitudinal Data Analysis

Objectives

- Review the additional features and extensions of multilevel modeling of IL data that have been incorporated into what has been broadly called Dynamic Structural Equation Modeling

177

Dynamic Structural Equation Modeling (DSEM)

- DSEM is a morphing of three previously separate traditions and captures the "best" of all these different analytic worlds:
  - SEM: Multiple (multivariate) outcomes; latent variable measurement modeling
  - Multilevel Modeling: Nested and crossed data structures (kids in classrooms, longitudinal panel design); random effects capturing between-person heterogeneity
  - Time Series Analysis: High-T repeated measures on a single unit; dynamical modeling with autoregressive effects

178

Dynamic Structural Equation Modeling (DSEM)

- Intensive longitudinal data (coming in particular from EMA and ESM designs) have interesting and challenging features that can benefit from this combined analytic framework
  - Numerous, dense, and unbalanced assessments; unequal measurement intervals; multiple outcomes of interest; within- and between-person variability on Ys and Xs; desire to estimate lagged effects as well as handle measurement error
- While some of these features can be handled in some R packages, Mplus v8 with its DSEM capabilities has been developed with a focus on intensive longitudinal data

179

DSEM Features

- Autoregressive effects and residual autocorrelation can
  - be specified and estimated
  - be allowed to vary across subjects
  - can be predicted by between subjects variables
- Residual variances can
  - be allowed to vary across subjects (i.e., random)
  - can be predicted by between subjects variables

180

DSEM Features

- Latent centering
  - Of predictors that vary both within- and between-persons
  - Of lagged Y as a predictor (addresses Nickell's bias; see Asparouhov & Muthén, 2018)
- Standardization of slopes
  - Using individual within-person SDs
  - As well as using between-person SDs of random effects

181

DSEM Features

- Handling unequal intervals (for EMA & ESM designs) when estimating lagged effects
  - Uses TINTERVAL
- All new intensive longitudinal features now done using Bayesian analysis (ESTIMATOR = BAYES)
  - Default uses "non-informative priors" which is essentially an ML solution
  - Dramatically reduces model non-convergence due to modeling multiple random effects

182

Objectives

- Describe what random within-person residual variance (or variance heterogeneity) is, why we would want to model it, and what it means conceptually
- Demonstrate multilevel modeling with variance heterogeneity with a real world example involving continuous glucose monitoring in persons with type 2 diabetes

183

Supporting Paper

Multivariate Behavioral Research, 50:334–349, 2015
Copyright © Taylor & Francis Group, LLC
ISSN: 0027-3171 print / 1532-7906 online
DOI: 10.1080/00273171.2014.1003772

Routledge
Taylor & Francis Group

A Multilevel AR(1) Model: Allowing for Inter-Individual Differences in Trait-Scores, Inertia, and Innovation Variance

Joran Jongerling
Department of Methodology and Statistics, Utrecht University

Jean-Philippe Laurenceau
Department of Psychological and Brain Sciences, University of Delaware

Ellen L. Hamaker
Department of Methodology and Statistics, Utrecht University

In this article we consider a multilevel first-order autoregressive [AR(1)] model with random intercepts, random autoregression, and random innovation variance (i.e., the level 1 residual variance). Including random innovation variance is an important extension of the multilevel AR(1) model for two reasons. First, between-person differences in innovation variance are important from a substantive point of view, in that they capture differences in sensitivity and/or exposure to unmeasured internal and external factors that influence the process. Second, using simulation methods we show that modeling the innovation variance as fixed across individuals, when it should be modeled as a random effect, leads to biased parameter estimates. Additionally, we use simulation methods to compare maximum likelihood estimation to Bayesian estimation of the multilevel AR(1) model and investigate the trade-off between the number of individuals and the number of time points. We provide an empirical illustration by applying the extended multilevel AR(1) model to daily positive affect ratings from 89 married women over the course of 42 consecutive days.

184

What is error variance?

- "One person's error is another person's treasure."
- Error = measurement error +
  true random noise +
  unobserved causal influences

185

Multilevel Modeling with Variance Heterogeneity

- In the standard multilevel model, we can have intercepts and slopes that vary across subjects (i.e., random effects) and assumes homogeneous σ^2
- In a somewhat newer extension of multilevel modeling, residual error variances can:
  - be allowed to vary across subjects
  - can be predicted by between-subjects variables
  - can be used to predict between-subjects distal outcomes
- We saw a limited version of this in the dyadic example where we had male and female residuals

186

Why Want Random Residual Variances?

- If there really are random residual variances, then ignoring this source of individual differences leads to bias in some multilevel modeling parameters (e.g., autoregressive effects)
- Between-person differences in residual variance are important from a substantive point of view, in that they capture differences in sensitivity and/or exposure to unmeasured internal and external factors that influence the process

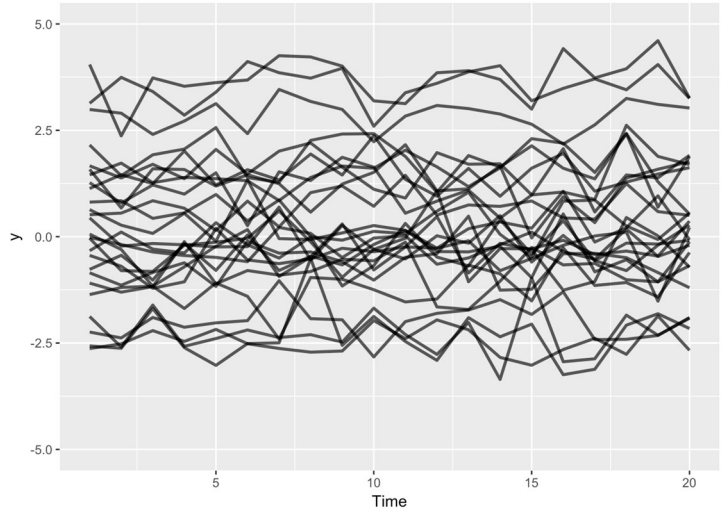
187

What Do Random Residual Variances Mean?

- The residual can be conceptualized as all **unobserved temporal factors** that influence the process under investigation, both internally (e.g., hormonal levels, alcohol intake, cognitions, associations, appraisal of events) and externally (e.g., social obligations, personal interactions, weather, political developments).
- Individual differences in the residual variance implies we allow not only for individual differences in **sensitivity and/or responsiveness** to these factors, but also for individual differences in **exposure** to these factors (or, more specifically, individual differences in the variability of these factors).

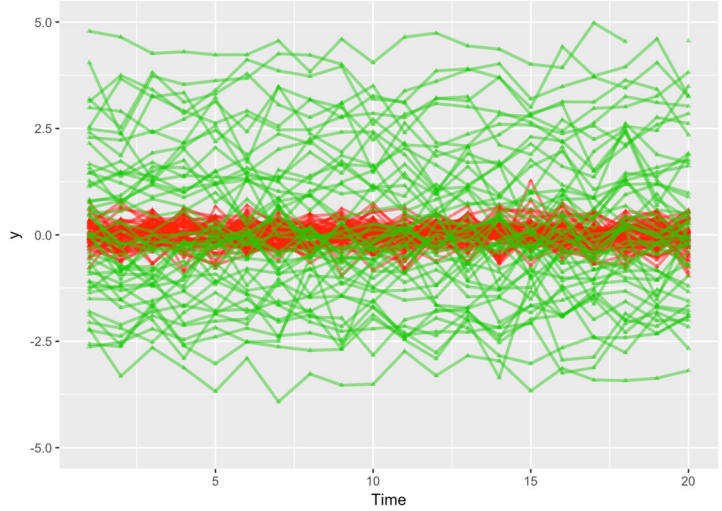
188

Homogeneity of Within-Person Variance



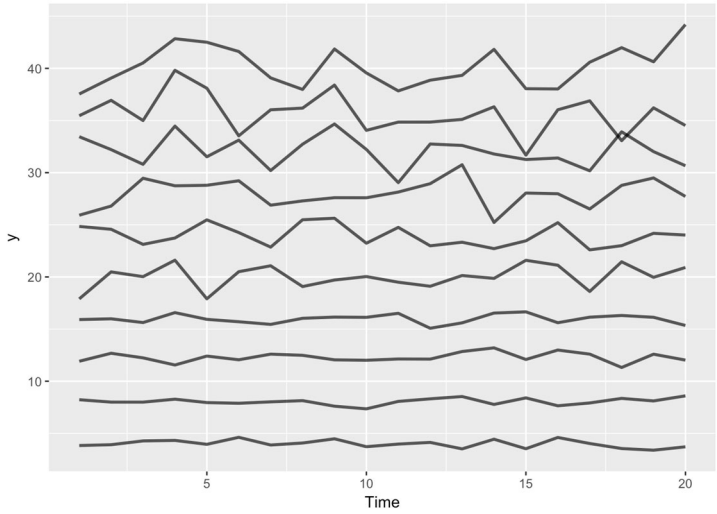
189

Heterogeneous Within-Person Variance (by group)



190

Heterogeneous Within-Person Variance (by person)



191

Data Example Aims

- Data example: Glucose Regulation in People with Type 2 Diabetes
- To explore individual differences in patient glycemic control (i.e., level and variability)
- Explore whether or not patient-level characteristics of 1) taking insulin or 2) duration of T2DM diagnosis are related to levels of glycemic control

192

Recruitment Brochure

Daily Diabetes Management
in the Marital Context

A collaborative project supported by the National Institute for Diabetes and Digestive and Kidney Diseases and conducted by:

CHRISTIANA CARE HEALTH SYSTEM
UNIVERSITY OF DELAWARE
NIH National Institute of Diabetes and Digestive and Kidney Diseases

Our Goal

This project was designed to examine social and behavioral factors that may influence glycemic control (blood sugar control) in people diagnosed with type 2 diabetes mellitus, such as stress and related interpersonal factors in the family (e.g., stress levels experienced by you and your spouse/ partner, the ways that you and your spouse/ partner handle stress, etc.). We hope that the information gained from this study will eventually help improve programs designed for couples who are coping with diabetes together.

Who can participate?

This study is for type 2 diabetes mellitus patients and their partners (living together; either married or in a committed relationship).

- Both patient and partner must:
  - Be comfortable speaking, reading, and answering survey questions in English.
  - Have their own smartphone.
  - Not have any major psychiatric disorders.
- Patients must:
  - Have been diagnosed with type 2 diabetes mellitus for at least 6 months.
  - Not be pregnant (if female).
  - Not have another serious or life-threatening disease (e.g., coronary artery disease, cancer, etc.) or an endogenous bleeding disorder (e.g., hemophilia).
  - Not be taking medications like Coumadin or Pradaxa.
- Partners/Spouses must not have diabetes.

What does participation involve?

Over the course of about a year, ***both you and your partner*** will be asked to make 3 visits to our research office at the Diabetes and Metabolic DIseases Center (DMDC) at Christiana Care Health System. You will also be asked to complete several surveys online, using your own smart-phone in your daily life and your computer at home.

- ***Baseline Assessment:*** Participating couples will come to the DMDC for an initial visit, which will take about 2.5 hours. At this visit, the patient will be equipped with a continuous glucose monitor (CGM), which he/she will wear for the next 7 consecutive days. Both partners will also be asked to use their smartphones to take pictures of the food and drinks they consume each day. The patient will also provide blood and urine samples. Both the patient and partner will com-plete a clinical interview and several questionnaires, and have their height, weight, and waist circumfer-ence measured.
- ***7-Day "Diary" Period:*** Immediately following the baseline assessment, both partners will complete 7 brief surveys per day for 7 consecutive days. Partici-pants will be reminded when it is time to complete a survey via SMS text messages to their smartphones. During this 7-day period, patients will continue to wear the CGM, and both patients and partners will be asked to wear an accelerometer device on their wrists to measure physical activity.
- ***Return Visit:*** Participating couples will come to the DMDC to return the CGM and accelerometer devices. Couples will also be asked to engage in brief, video-taped conversation tasks and answer a few short questionnaires.
- ***Follow-Up Surveys:*** A set of online questionnaires completed at home by both partners 3 months and 6 months after the Baseline Assessment.

193

Patient Demographics

- ***N*** = 63 patients with T2D and their spouses/partners
- **Gender**: 2/3 male
- **Age**: *M* = 61 (*SD* = 10)
- **Race & ethnicity**: 97% non-Hispanic/Latino, 70% Caucasian, 24% African American, 4% other/multiracial, 2% Asian
- **Time since T2D diagnosis:** *M* = 6.5 (*SD* = 5.8)
- **Insulin-treated** (vs. oral only): 56%

194

Glycemic Control: Beyond HbA1c

- Limitations of HbA1c, the gold standard for assessing glycemic control
  - Only yields average glucose over past ~3 months
  - Does not reflect variability or excursions
- **Glucose variability** has been linked to long-term complications of T2D, independent of HbA1c (i.e., mean blood sugar)
  - Cardiovascular disease (i.e., oxidative stress) and cognitive functioning

195

Lots on glucose variability in diabetes literature

Acta Diabetol (2012) 49 (Suppl 1):S153–S160
DOI 10.1007/s00592-012-0391-4

ORIGINAL ARTICLE

The importance of HbA1c and glucose variability in patients with type 1 and type 2 diabetes: outcome of continuous glucose monitoring (CGM)

Giovanni Sartore · Nino Cristiano Chilelli · Silvia Burlina · Paola Di Stefano · Francesco Piarulli · Domenico Fedele · Andrea Mosca · Annunziata Lapolla

Review
Complications
Diabetes Metab J 2015;39:273-282
http://dx.doi.org/10.4093/dmj.2015.39.4.273
pISSN 2233-6079 · eISSN 2233-6087

dmj
DIABETES & METABOLISM JOURNAL

Glycemic Variability: How Do We Measure It and Why Is It Important?

Sunghwan Suh', Jae Hyeon Kim'

DIABETES TECHNOLOGY & THERAPEUTICS
Volume 7, Number 6, 2005
© Mary Ann Liebert, Inc.

Quantifying Temporal Glucose Variability in Diabetes via Continuous Glucose Monitoring: Mathematical Methods and Clinical Application

BORIS P. KOVATCHEV, Ph.D., WILLIAM L. CLARKE, M.D., MARC BRETON, Ph.D., KENNETH BRAYMAN, M.D., and ANTHONY MCCALL, M.D., Ph.D.

POINT-COUNTERPOINT

Glucose Variability: Where It Is Important and How to Measure It

J. Hans DeVries

Toward Defining the Threshold Between Low and High Glucose Variability in Diabetes

Contents available at ScienceDirect
Diabetes Research and Clinical Practice
journal homepage: www.elsevier.com/locate/diabres
ELSEVIER
International Diabetes Federation

Review
Glucose variability: An emerging target for the treatment of diabetes mellitus

Louis Monnier,[1] Claude Colette,[1] Anne Wojtusciszyn,[2] Sylvie Dejager,[2] Eric Renard,[3] Nicolas Molinari,[4] and David R. Owens[5]

Simona Frontoni[a], Paolo Di Bartolo[b], Angelo Avogaro[c], Emanuele Bosi[d], Giuseppe Paolisso[e], Antonio Ceriello[f,*]

Diabetes Care 2017;40:832–838 | https://doi.org/10.2337/dc16-1769

Glycemic Variability

Should we and can we prevent it?

LOUIS MONNIER, MD[1]
CLAUDE COLETTE, PHD[2]

DIABETES TECHNOLOGY & THERA
Volume 11, Supplement 1, 2009
© Mary Ann Liebert, Inc.
DOI: 10.1089/dia.2008.0132

Glucose Variability; Does It Matter?

Sarah E. Siegelaar, Frits Holleman, Joost B. L. Hoekstra, and J. Hans DeVries

Glycemic Variability: Both Sides of the Story

ANTONIO CERIELLO, MD[1]
ERIC S. KILPATRICK, MD[2]

taken over a p
showed that w

Interpretation of Continuous Glucose Monitoring Data: Glycemic Variability and Quality of Glycemic Control

David Rodbard, M.D.

196

Continuous Glucose Monitoring (CGM)

dexcom®
CONTINUOUS GLUCOSE MONITORING

197

A CGM Spokesperson J-P's Daughter Likes

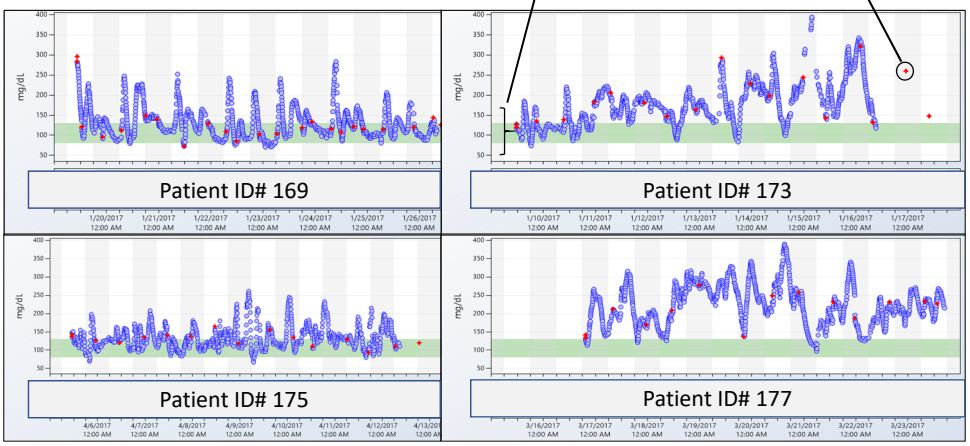
198

Example Patient CGM Data

Max. N = 2,016 readings

Target range:
70 – 180 ml/dL

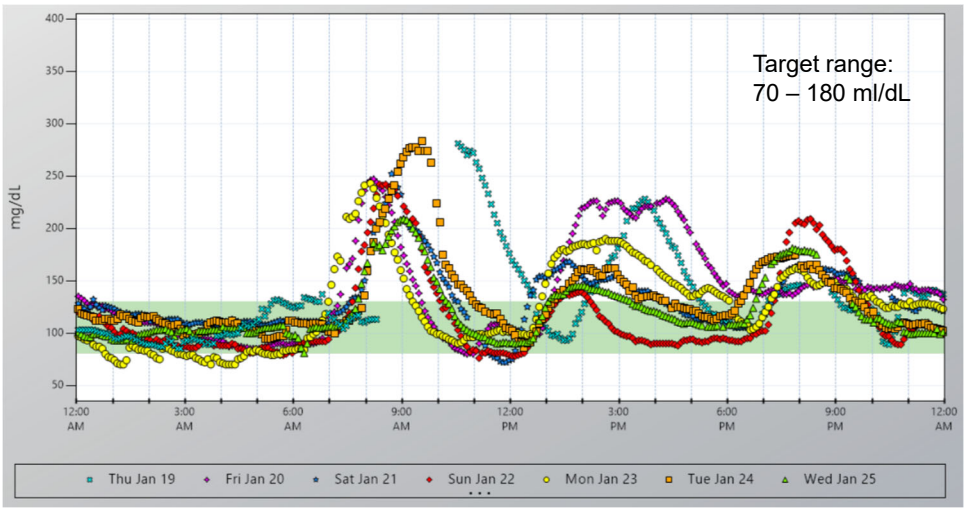
Calibrations
(glucometer readings)



Correlation between patients' mean CGM value and baseline HbA1c: $r = .70$, $p < .001$

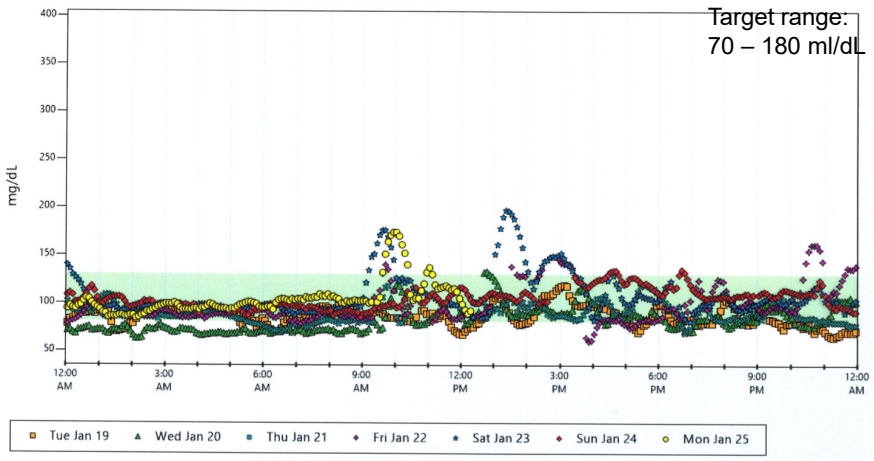
199

Example Raw CGM Data (ID# 169)



200

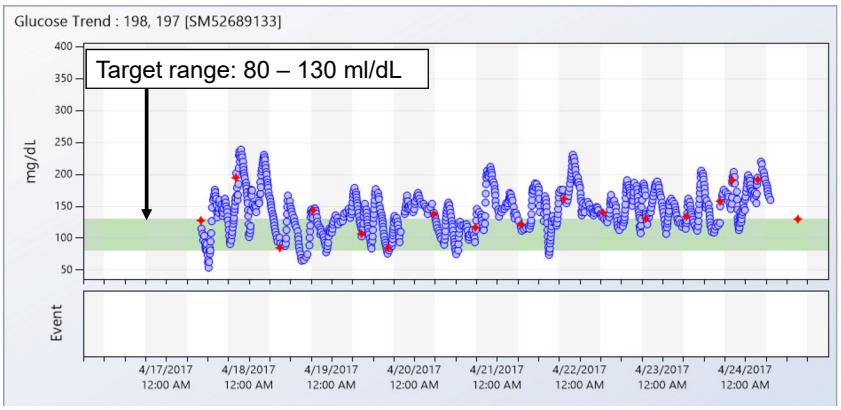
Example of Emily's (Non-Diabetic) Raw CGM Data



201

Participant 197

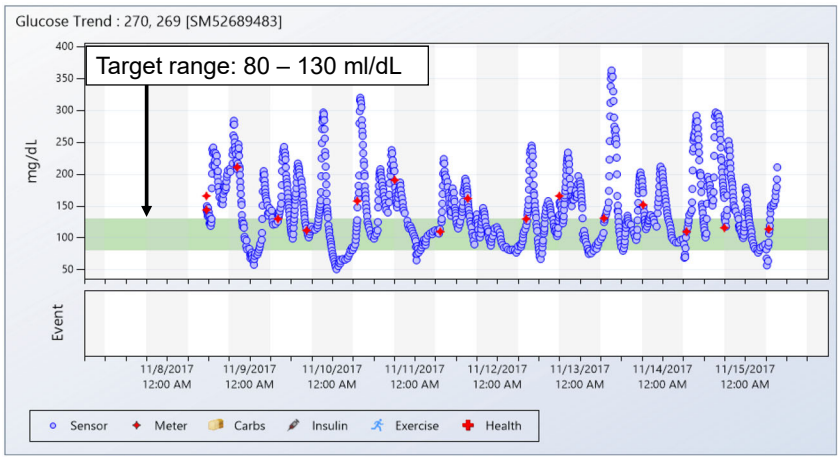
Mean glucose: 142 mg/dL



202

Participant 269

Mean glucose: 140 mg/dL



203

Data Analytic Considerations

- Multilevel location-scale model with random intercept variance, random linear time slope, and random residual variance parameters (e.g., Hedeker, Mermelstein, & Demirtas, 2012)
- Exponential function of residual variance used to ensure positive variances
- Bayesian estimation (with non-informative priors) implemented in Mplus 8
- Linear time included as fixed and random effect (detrending)

204

Results

- 63 patients providing 122,950 every-5 minutes blood glucose readings (~97% compliance)
- Both **taking insulin** and **length of TDM2 diagnosis** predicted higher mean CGM blood glucose levels
- Both **taking insulin** and **length of TDM2 diagnosis** predicted greater residual variability

205

Mplus Syntax

```
ANALYSIS: TYPE = TWOLEVEL RANDOM;
          ESTIMATOR = BAYES;
          PROCESSORS = 4;
          BITERATIONS = (2000);
MODEL:
%WITHIN%
slope|CGMv on time; ! regresses CGM value on centered linear time (detrending)
logv | CGMv; ! specifies random residual variance for CGM value outcome

%BETWEEN%
CGMv on p1insulin T2DdurationES; ! regress CGM mean on insulin and T2 duration
logv on p1insulin T2DdurationES; ! regress latent random residual on insulin and T2 duration
CGMv logv slope with CGMv logv slope; ! allow for random effect covariances
p1insulin T2DdurationES; ! makes predictors "outcomes" that can have missing
```

206

Excerpt of Results

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
Between Level						
LOGV ON						
P1INSULIN	0.514	0.177	0.003	0.164	0.866	*
T2DDURATIO	0.027	0.012	0.013	0.003	0.050	*
CGMV ON						
P1INSULIN	23.484	7.758	0.002	7.330	37.912	*
T2DDURATIO	1.082	0.502	0.018	0.077	2.080	*
CGMV WITH						
LOGV	11.890	3.401	0.000	6.681	19.912	*
SLOPE	60.848	30.027	0.015	7.934	128.369	*
LOGV WITH						
SLOPE	0.492	0.630	0.206	-0.722	1.766	
Means						
P1INSULIN	-0.006	0.066	0.473	-0.128	0.127	
T2DDURATIO	0.038	1.046	0.481	-2.030	2.023	
SLOPE	0.235	0.833	0.388	-1.351	1.800	
Intercepts						
CGMV	157.320	3.853	0.000	149.800	164.921	*
LOGV	7.442	0.085	0.000	7.276	7.619	*
Variances						
P1INSULIN	0.261	0.052	0.000	0.185	0.394	*
T2DDURATIO	66.430	12.798	0.000	47.981	97.829	*
SLOPE	40.676	8.336	0.000	28.746	61.002	*
Residual Variances						
CGMV	915.739	190.234	0.000	644.913	1398.595	*
LOGV	0.442	0.090	0.000	0.309	0.653	*

207

Excerpt of Results

R-SQUARE

Within-Level R-Square Averaged Across Clusters

Variable	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%
CGMV	0.046	0.001	0.000	0.044	0.048

Between Level

Variable	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%
CGMV	0.208	0.075	0.000	0.075	0.364

Variable	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%
LOGV	0.222	0.080	0.000	0.080	0.390

208

Resources

- Mplus Time Series Analysis: Dynamic Structural Equation Modeling (DSEM) https://www.statmodel.com/TimeSeries.shtml
- Hedeker papers and package: Mixed Effects Location Scale Modeling https://voices.uchicago.edu/hedeker/mixwild_mixregls/
- Data example on Dr. Nilam Ram's page: https://quantdev.ssri.psu.edu/tutorials/analysis-experience-sampling-ema-data-chapter-6-multilevel-model-heterogeneous-variance

209

Summary

- Random within-person residual variance is a way of allowing each individual to have their own residual reflecting greater fluctuations and temporal reactivity in the person's outcome to the environment
- Estimating and modeling of random within-person residual variance can be very interesting extensions of multilevel modeling of IL data to answer new questions

210

Objectives

- Describe two biases that are addressed by using latent mean centering of predictors versus observed person-mean centering

211

Lüdtke's Bias

- Recall the standard person-mean centering of a predictor that varies both within- and between-person:

$$(X_{ij} - X_{.j})$$

- Lüdtke's bias takes place when observed person-mean centering doesn't estimate the between effect correctly
- This occurs because of measurement error in in $X_{.j}$ that is not accounted for

212

A Fix for Lüdtke's Bias

- Lüdtke et al. (2008) have demonstrated that latent centering based on multilevel SEM approach eliminates this bias
- This approach also accommodates missing data and a random effect on the predictor involved in the latent centering

213

Nickell's Bias

- Comes from a known issue in dynamic panel and time series models from econometrics (i.e., autoregressive models involving the lagged outcome (Y_{t-1}) as a predictor of current outcome (Y_t)
- If observed person-mean centering of the lagged predictor is used, then this leads to a bias in the autoregressive regression coefficient (due to using the sample instead of the true mean for centering)

214

A Fix for Nickell's bias

- Asparouhov & Muthén (2018) have demonstrated that latent centering in multilevel SEM approach eliminates Nickell's bias
- Latent centering is implemented (with a confirmatory note near the top of the output) when doing TYPE=TWOLEVEL modeling with ESTIMATOR=BAYES and by entering the same daily predictor in both %WITHIN% and %BETWEEN% parts of the model

215

Comments

- Latent mean centering appears to solve these and other biases beyond the scope of this workshop
- Currently, we are not aware of any other package than Mplus that implements latent mean centering

216

Objectives

- Describe the basic challenge with trying to decide how to standardize an average within-subjects regression coefficient
- Demonstrate a recommended approach to standardization called within-person standardization

217

Standardizing Regression Coefficients

- Unstandardized (Raw) Coefficient: Tells you the shift in outcome in its measurement units for a one unit shift in predictor in its measurement unit
- Standardized Coefficient: Tells you the shift in outcome in standard deviation units for a one standard deviation shift in predictor

$$\beta = B * (SD_X/SD_Y)$$

218

Which Variance do You Choose to Use?

- You use focus on the within-person SD_X and within-person SD_Y for the sample
- You can use the between-person SD_X and within-person SD_Y for the sample
- You can use the overall (combined) SD_X and within-person SD_Y for the sample

- The Problem: You will get different standardized coefficients depending on which you choose

219

A Principled Solution (Schuurman et al., 2016)

- Because of a desire to focus on and respect within-person effects and between-person heterogeneity in these effects, there is an argument to standardize each person's within-person coefficient using their own individual standard deviations (iSD_X and iSD_Y)
- For the standardized average within-person effect, you would then report the average of the individual standardized within-person effects
- This is known as within-person (person-specific) standardization

220

Process Example in DSEM (Raw Coefficients)

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
Between Level						
CONFSLP ON						
RELQUAL	1.029	0.520	0.024	0.009	2.047	*
INTIMACY ON						
CONFLICTB	-0.815	1.143	0.241	-2.999	1.445	
CONFBRQ	2.526	1.769	0.076	-1.016	5.956	
RELQUAL	0.101	0.496	0.416	-0.879	1.075	
INTIMACY WITH						
CONFSLP	0.427	0.287	0.051	-0.100	1.049	
Intercepts						
INTIMACY	4.533	0.227	0.000	4.079	4.981	*
CONFSLP	-2.024	0.392	0.000	-2.771	-1.234	*
Residual Variances						
INTIMACY	0.892	0.200	0.000	0.599	1.369	*
CONFSLP	3.017	0.793	0.000	1.836	4.932	*

221

Process Example in DSEM (Standardized)

STDYX Standardization

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	Upper 2.5%	Significance
Within-Level Standardized Estimates Averaged Over Clusters						
CONFSLP | INTIMACY ON						
CONFLICTW	-0.232	0.018	0.000	-0.267	-0.196	*
INTIMACY^ ON						
INTIMACY^1	-0.046	0.026	0.041	-0.096	0.006	
INTIMACY ON						
TIME7C	-0.017	0.021	0.223	-0.059	0.025	
Residual Variances						
INTIMACY	0.856	0.012	0.000	0.832	0.879	*
Between Level						
CONFSLP ON						
RELQUAL	0.202	0.098	0.024	0.002	0.382	*

222

Comments

- Within-person standardization using each person's estimated within-person effect and standard deviations of X and Y is a defensible way to report standardized effects
- This is implemented in Mplus by adding the "standardized (cluster)" option to the OUTPUT command, but can also be done by hand using results from other packages

223

Updated Workshop Examples

Updated Workshop Examples Syntax for Mplus v8

- Chapter 2: Intimacy Tx example (Time Course)
- Chapter 3: Conflict reactivity (Process)
- Chapter 6: Dichotomous outcome
- Chapter 7: Dyads

224

Chapter Summary

- We covered several interesting and exciting extensions of the standard multilevel model approach to analyzing IL data
- DSEM in Mplus is an analytic framework for implementing them
- Hopefully, packages in R will be developed in the near future to be able to conduct and reproduce DSEM results

225

Recommended Readings

- McNeish, D. & Hamaker, E.L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods, 25*, 610–635.
- Asparouhov, T. & Muthén, B. (2020). Comparison of models for the analysis of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 27*, 275-297.
- Asparouhov, T. & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. *Structural Equation Modeling: A Multidisciplinary Journal, 26*, 119-142.

226