Estimating Internal Consistency Reliability for Intensive Longitudinal Data

A Brief Tutorial

Sebastian Castro-Alvarez

University of Groningen

November 4th, 2025

university of groningen

faculty of behavioural and social sciences

MITNB
Measurement is the new black

tesc
Tilburg Experience Sampling Center

Groundwork and Acknowledgements

Castro-Alvarez, S., Bringmann, L. F., Back, J., & Liu, S. (2025). The many reliabilities of psychological dynamics: An overview of statistical approaches to estimate the internal consistency reliability of intensive longitudinal data. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000778

Castro-Alvarez, S., Zhou, D. J., Bringmann, L. F., Tutunji, R., Proppert, R. K. K., Rieble, C., . . . Liu, S. (In Press). Assessing the Internal Consistency Reliability of Ecological Momentary Assessment Measures: Insights from the WARN-D Study. https://doi.org/10.31234/osf.io/nrzsc_v1

Groundwork and Acknowledgements

Groundwork and Acknowledgements

Leertouwer, I., Keijsers, L., van Roekel, E., & Schuurman, N. K. (2025, June 18). A Practical Guide to Estimating Reliability of Intensive Longitudinal Data. https://doi.org/10.31234/osf.io/uq4sk_v1

Agenda

1. **Intensive Longitudinal Data and Reliability ($\approx$30 min)**
2. **Get to Work! R Practical ($\approx$40 min)**
3. **Final Thoughts - Take Home Messages ($\approx$20 min)**

Intensive Longitudinal Data and Reliability

- Emphasis on the individual.
- Reduced recall bias.
- Ecological validity

Intensive Longitudinal Data and Reliability

Are data reliable?

- Emphasis on the individual.
- Reduced recall bias.
- Ecological validity

Intensive Longitudinal Data and Reliability

What is Reliability and why do we care?

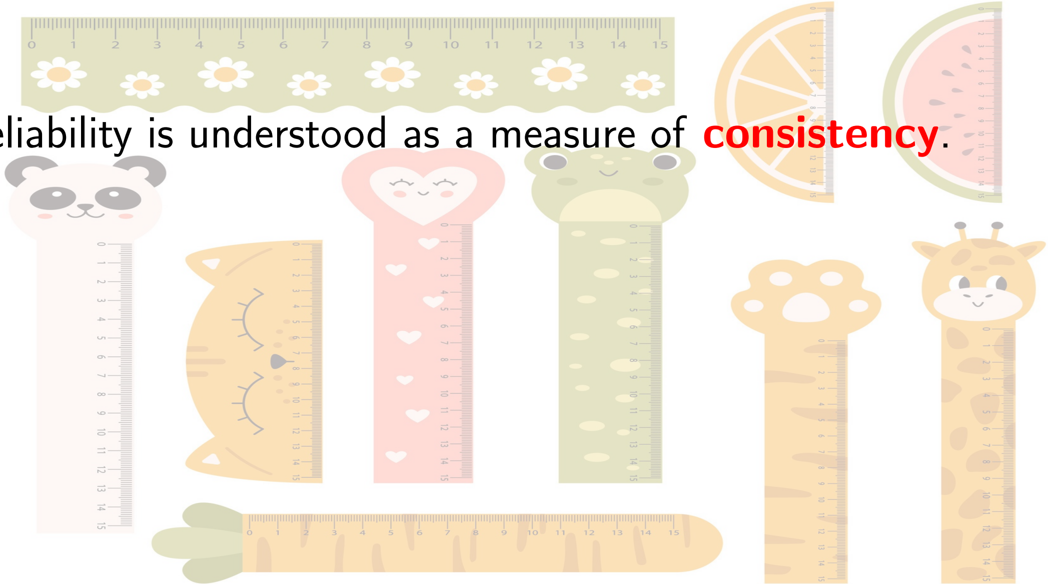
Intensive Longitudinal Data and Reliability

What is Reliability and why do we care?

Observed = True + Error

Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.



Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.

Consistency of the scores across:

- Items $\Rightarrow$ *Internal Consistency*.

Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.

Consistency of the scores across:

- Items ⇒ *Internal Consistency*.
- Time ⇒ *Test-retest*.

Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.

Consistency of the scores across:

- Items $\Rightarrow$ *Internal Consistency*.
- Time $\Rightarrow$ *Test-retest*.
- Equivalent tests $\Rightarrow$ *Parallel tests*.

Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.

Consistency of the scores across:

- Items $\Rightarrow$ *Internal Consistency*.
- Time $\Rightarrow$ *Test-retest*.
- Equivalent tests $\Rightarrow$ *Parallel tests*.
- Raters $\Rightarrow$ *Inter-rater*

Intensive Longitudinal Data and Reliability

Reliability is understood as a measure of **consistency**.

Consistency of the scores across:

- Items ⇒ *Internal Consistency.*
- Time ⇒ *Test-retest*
- Equivalent tests ⇒ *Parallel tests.*
- Raters ⇒ *Inter-rater*

How can we apply this to Intensive Longitudinal Data?

Intensive Longitudinal Data and Reliability

Thinking about reliability in intensive longitudinal data is challenging

Intensive Longitudinal Data and Reliability

Thinking about reliability in intensive longitudinal data is challenging . . .

but not impossible

Intensive Longitudinal Data and Reliability

In the last manuscript/paper you worked on:

- **Did you report the reliability of your ESM measures?**

Intensive Longitudinal Data and Reliability

In the last manuscript/paper you worked on:

- **Did you report the reliability of your ESM measures?**
- **Did you use multiple items to measure one construct (e.g., Positive Affect)?**

Intensive Longitudinal Data and Reliability

Reliability in ESM studies is dramatically underreported!

- Brose et al. (2020) 29 out of 50 (58%).
- Hall et al. (2021) 72 out of 234 (30.7%).
- Castro-Alvarez et al. (2024) 73 out of 124 (58.8%), and 94 computed sum scores!

Intensive Longitudinal Data and Reliability

Reliability in ESM studies is dramatically underreported!

- Brose et al. (2020) 29 out of 50 (58%).
- Hall et al. (2021) 72 out of 234 (30.7%).
- Castro-Alvarez et al. (2024) 73 out of 124 (58.8%), and 94 computed sum scores!

Reliability approach	Frequency
None	58
Cronbach's α (Whole)	33
Cronbach's α (One time point)	12
McDonald's ω (Whole)	1
Multilevel ω (Geldhof et al., 2014)	14
GT (Cranford et al., 2006)	12
Multilevel Modeling (Nezlek, 2017)	1

Intensive Longitudinal Data and Reliability

Approaches to study the internal consistency of intensive longitudinal data:

- Generalizability theory (Cranford et al., 2006).
- Multilevel modeling (Nezlek, 2017).
- Multilevel confirmatory factor analysis (Geldhof et al., 2014; Lai, 2021).
- P-technique and dynamic factor analysis (Fuller-Tyszkiewicz et al., 2017; Hu et al., 2016).
- DSEM: Multilevel dynamic factor analysis (Xiao et al., 2023).
- DSEM: Latent state trait theory (Castro-Alvarez et al., 2022).

Intensive Longitudinal Data and Reliability

- Daily diary data over 100 days from 193 participants (O'Laughlin et al., 2021; Williams et al., 2020).
- PANAS Positive affect items (10): *interested, excited, strong, enthusiastic, proud, inspired, determined, attentive, active, alert*.
- Compliance: Median $= 93$ (Min $= 14$, Max $= 100$).
- Cronbach's $\alpha = 0.95$

Intensive Longitudinal Data and Reliability

Generalizability Theory (Cranford et al., 2006)

$$y_{ijt} = \mu + P_i + I_j + O_t + (P \times I)_{ij} + (P \times O)_{it} + (I \times O)_{jt} + (P \times I \times O)_{ijt} + \varepsilon_{ijt}$$

Intensive Longitudinal Data and Reliability

Generalizability Theory (Cranford et al., 2006)

$$y_{ijt} = \mu + P_i + I_j + O_t + (P \times I)_{ij} + (P \times O)_{it} + (I \times O)_{jt} + (P \times I \times O)_{ijt} + \varepsilon_{ijt}$$

$$R_{kr} = \frac{\sigma^2_P + \sigma^2_{P \times I}/M}{\sigma^2_P + \sigma^2_{P \times I}/M + \sigma^2_O/K + \sigma^2_{P \times O}/K + \sigma^2_\epsilon/MK} \qquad R_c = \frac{\sigma^2_{P \times O}}{\sigma^2_{P \times O} + \sigma^2_\epsilon/M}$$

Also proposed the following coefficients: R_{kf}, R_{1f}, and R_{1r}.

Intensive Longitudinal Data and Reliability

	Variance Component		Reliability Coefficient
σ^2_P	3.27	R_{1f}	0.95
σ^2_I	0.04	R_{kf}	1.00
σ^2_O	0.04	R_{1r}	0.65
$\sigma^2_{P\times I}$	0.47	R_{kr}	0.99
$\sigma^2_{P\times O}$	1.59	R_c	0.89
$\sigma^2_{I\times O}$	0.01		
σ^2_ϵ	1.89		
Total	7.30		

Intensive Longitudinal Data and Reliability

Multilevel modeling (Nezlek, 2017)

$$
\begin{aligned}
&\text{Level 1 (Item):} && y_{ijt} = \pi_{0it} + \varepsilon_{ijt} && \varepsilon_{ijt} \sim N(0, \sigma_{\varepsilon}^2) \\
&\text{Level 2 (Time):} && \pi_{0it} = \beta_{00i} + r_{0it} && r_{0it} \sim N(0, \sigma_{r}^2) \\
&\text{Level 3 (Persons):} && \beta_{00i} = \gamma_{000} + u_{00i} && u_{00i} \sim N(0, \sigma_{u}^2)
\end{aligned}
$$

Intensive Longitudinal Data and Reliability

Multilevel modeling (Nezlek, 2017)

Level 1 (Item):	$y_{ijt} = \pi_{0it} + \varepsilon_{ijt}$	$\varepsilon_{ijt} \sim N(0, \sigma^2_\varepsilon)$
Level 2 (Time):	$\pi_{0it} = \beta_{00i} + r_{0it}$	$r_{0it} \sim N(0, \sigma^2_r)$
Level 3 (Persons):	$\beta_{00i} = \gamma_{000} + u_{00i}$	$u_{00i} \sim N(0, \sigma^2_u)$

$$R_{krn} = \frac{\sigma^2_u}{\sigma^2_u + \sigma^2_r/K + \sigma^2_\epsilon/MK}$$

$$R_{cn} = \frac{\sigma^2_r}{\sigma^2_r + \sigma^2_\epsilon/M}$$

Intensive Longitudinal Data and Reliability

	Variance Component		Reliability Coefficient
$\sigma^2_{u_{00}}$	3.32	R_{krn}	0.99
$\sigma^2_{r_0}$	1.57	R_{cn}	0.87
σ^2_{ϵ}	2.42		
Total	7.31		

Intensive Longitudinal Data and Reliability

Multilevel CFA (Geldhof et al., 2014; Lai, 2021)

$$y_{ijt} = \lambda_j^b \xi_i + \vartheta_{ij} + \lambda_j^w \zeta_{it} + \upsilon_{ijt} \text{ with } \lambda_j^b = \lambda_j^w$$

Intensive Longitudinal Data and Reliability

Multilevel CFA (Geldhof et al., 2014; Lai, 2021)

$$y_{ijt} = \lambda_j^b \xi_i + \vartheta_{ij} + \lambda_j^w \zeta_{it} + \upsilon_{ijt} \text{ with } \lambda_j^b = \lambda_j^w$$

$$\omega^b = \frac{(\sum_{j=1}^M \lambda_j)^2 \sigma_\xi^2}{(\sum_{j=1}^M \lambda_j)^2 (\sigma_\xi^2 + \sigma_\zeta^2 / \tilde{K}) + \sum_{j=1}^M \sigma_{\vartheta_j}^2 + \sum_{j=1}^M \sigma_{\upsilon_j}^2 / \tilde{K}},$$

$$\omega^w = \frac{(\sum_{j=1}^M \lambda_j)^2 \sigma_\zeta^2}{(\sum_{j=1}^M \lambda_j)^2 \sigma_\zeta^2 + \sum_{j=1}^M \sigma_{\upsilon_j}^2}.$$

Intensive Longitudinal Data and Reliability

	Reliability Coefficient	C.I.
ω^{w}	0.896	(0.894, 0.898)
ω^{b}	0.977	(0.971, 0.981)
ω^{2l}	0.953	(0.951 0.954)

Intensive Longitudinal Data and Reliability

P-technique and dynamic factor analysis (Fuller-Tyszkiewicz et al., 2017; Hu et al., 2016).

P-technique FA

$$y_{ijt} = \lambda_{ij}^{w} \zeta_{it} + \upsilon_{ijt}$$

$$\omega_i^w = \frac{(\sum_{j=1}^{M} \lambda_{ij}^{w})^2 \sigma_{\zeta_i}^2}{(\sum_{j=1}^{M} \lambda_{ij}^{w})^2 \sigma_{\zeta_i}^2 + \sum_{j=1}^{M} \sigma_{\upsilon_{ij}}^2}$$

Intensive Longitudinal Data and Reliability

P-technique and dynamic factor analysis (Fuller-Tyszkiewicz et al., 2017; Hu et al., 2016).

P-technique FA

$$y_{ijt} = \lambda^{w}_{ij}\zeta_{it} + \upsilon_{ijt}$$

Dynamic FA

$$y_{ijt} = \lambda^{w}_{ij}\zeta_{it} + \upsilon_{ijt}$$

$$\zeta_{it} = \phi_i \zeta_{i,t-1} + \varepsilon_{it}$$

$$\omega^{w}_{i} = \frac{(\sum_{j=1}^{M} \lambda^{w}_{ij})^2 \sigma^2_{\zeta_i}}{(\sum_{j=1}^{M} \lambda^{w}_{ij})^2 \sigma^2_{\zeta_i} + \sum_{j=1}^{M} \sigma^2_{\upsilon_{ij}}}$$

$$\omega^{w}_{i} = \frac{(\sum_{j=1}^{M} \lambda^{w}_{ij})^2 \frac{\sigma^2_{\varepsilon_i}}{1-\phi^2_i}}{(\sum_{j=1}^{M} \lambda^{w}_{ij})^2 \frac{\sigma^2_{\varepsilon_i}}{1-\phi^2_i} + \sum_{j=1}^{M} \sigma^2_{\upsilon_{ij}}}$$

Intensive Longitudinal Data and Reliability

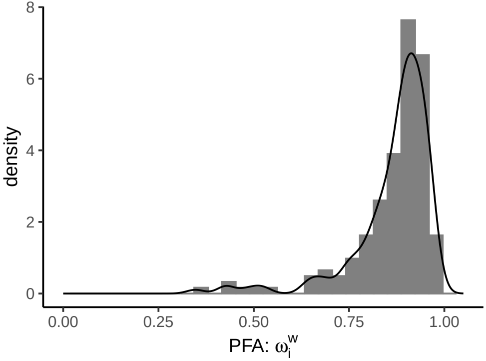
Estimated for participants with 50 or more timepoints (N $=$ 172)

Intensive Longitudinal Data and Reliability

Estimated for participants with 50 or more timepoints (N = 172)

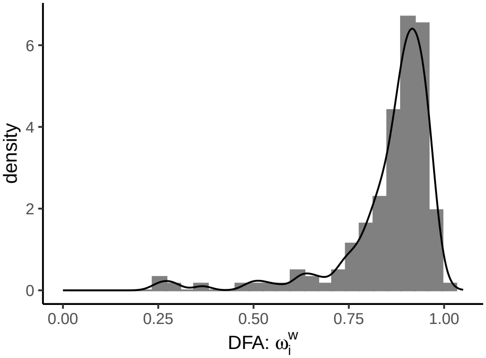
P-Technique FA

Median $\omega_i^w = 0.89$ (Min = 0.35, Max = 0.98).

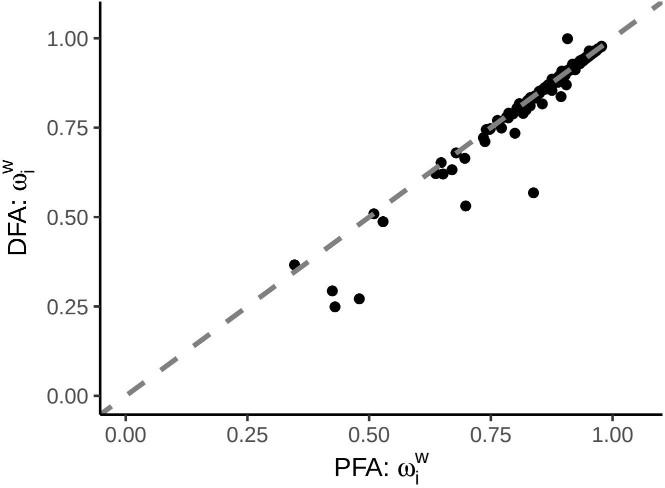


Dynamic FA

Median $\omega_i^w = 0.89$ (Min = 0.25, Max = 0.99).



Intensive Longitudinal Data and Reliability



Intensive Longitudinal Data and Reliability

Two-Level Random Dynamic Model-Based Approach (Xiao et al., 2023)

Latent Decomposition: $\boldsymbol{y}_{it} = \boldsymbol{y}_i^b + \boldsymbol{y}_{it}^w$

Between-level Measurement: $\boldsymbol{y}_i^b = \boldsymbol{\mu} + \boldsymbol{\Lambda}^b \boldsymbol{\xi}_i + \boldsymbol{\vartheta}_i$

Within-level Measurement: $\boldsymbol{y}_{it}^w = \boldsymbol{\Lambda}_i^w \boldsymbol{\zeta}_{it} + \boldsymbol{\upsilon}_{it}$

Within-level Structure: $\boldsymbol{\zeta}_{it} = \boldsymbol{\Phi}_i \boldsymbol{\zeta}_{i,t-1} + \boldsymbol{\varepsilon}_{it}$

Intensive Longitudinal Data and Reliability

Two-Level Random Dynamic Model-Based Approach (Xiao et al., 2023)

Latent Decomposition:	$\boldsymbol{y}_{it} = \boldsymbol{y}_i^b + \boldsymbol{y}_{it}^w$
Between-level Measurement:	$\boldsymbol{y}_i^b = \boldsymbol{\mu} + \boldsymbol{\Lambda}^b \boldsymbol{\xi}_i + \boldsymbol{\vartheta}_i$
Within-level Measurement:	$\boldsymbol{y}_{it}^w = \boldsymbol{\Lambda}_i^w \boldsymbol{\zeta}_{it} + \boldsymbol{\upsilon}_{it}$
Within-level Structure:	$\boldsymbol{\zeta}_{it} = \boldsymbol{\Phi}_i \boldsymbol{\zeta}_{i,t-1} + \boldsymbol{\varepsilon}_{it}$

$$\omega_p^b = \frac{(\sum_{j=1}^{M_p} \lambda_j^b)^2 \sigma_{\xi_p}^2}{(\sum_{j=1}^{M_p} \lambda_j^b)^2 \sigma_{\xi_p}^2 + \sum_{j=1}^{M_p} \sigma_{\vartheta_j}^2}$$

$$\omega_{ip}^w = \frac{(\sum_{j=1}^{M_p} \lambda_{ij}^w)^2 \sigma_{\zeta_{ip}}^2}{(\sum_{j=1}^{M_p} \lambda_{ij}^w)^2 \sigma_{\zeta_{ip}}^2 + \sum_{j=1}^{M_p} \sigma_{\upsilon_{ij}}^2}$$

Intensive Longitudinal Data and Reliability

How do we get $\sigma^2_{\zeta_{ip}}$?

Intensive Longitudinal Data and Reliability

How do we get $\sigma^2_{\zeta_{ip}}$?

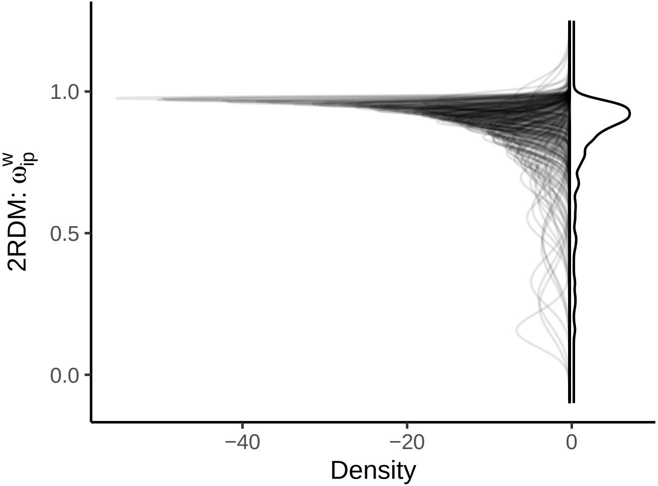
$$\boldsymbol{\Sigma}_{\zeta_{\boldsymbol{i}}} = \text{mat}((\boldsymbol{I} - \boldsymbol{\Phi}_{\boldsymbol{i}} \otimes \boldsymbol{\Phi}_{\boldsymbol{i}})^{-1} \text{vec}(\boldsymbol{\Sigma}_{\varepsilon_{\boldsymbol{i}}}))$$

Intensive Longitudinal Data and Reliability

How do we get $\sigma^2_{\zeta_{ip}}$?

$$\boldsymbol{\Sigma}_{\zeta_{\boldsymbol{i}}} = \text{mat}((\boldsymbol{I} - \boldsymbol{\Phi}_{\boldsymbol{i}} \otimes \boldsymbol{\Phi}_{\boldsymbol{i}})^{-1} \text{vec}(\boldsymbol{\Sigma}_{\varepsilon_{\boldsymbol{i}}})) \quad \Rightarrow \qquad \sigma^2_{\zeta_{pi}} = \frac{\sigma^2_{\varepsilon_i}}{1 - \phi^2_i}$$

Intensive Longitudinal Data and Reliability



- $\omega_p^b = 0.99$ $(0.989, 0.993)$
- Median $\omega_{ip}^w = 0.90$ (Min $=$ 0.16, Max $= 0.98$).

Intensive Longitudinal Data and Reliability

LST Theory: Mixed-effect Trait-State-Occasion Model (Castro-Alvarez et al., 2022)

Observed Score Decomposition:	$y_{ijt} = \theta_{ijt} + \upsilon_{ijt}$
Latent State Decomposition:	$\theta_{ijt} = \xi_{ij} + \lambda_j^w \zeta_{it}$
Latent Occasion Residual Dynamics:	$\zeta_{it} = \phi_i \zeta_{i,t-1} + \varepsilon_{it}$

Intensive Longitudinal Data and Reliability

LST Theory: Mixed-effect Trait-State-Occasion Model (Castro-Alvarez et al., 2022)

Observed Score Decomposition: $y_{ijt} = \theta_{ijt} + \upsilon_{ijt}$

Latent State Decomposition: $\theta_{ijt} = \xi_{ij} + \lambda_j^w \zeta_{it}$

Latent Occasion Residual Dynamics: $\zeta_{it} = \phi_i \zeta_{i,t-1} + \varepsilon_{it}$

Model implied variance of the responses to an item per person:

$$\sigma^2_{Y_{ij}} = \sigma^2_{\xi_j} + (\lambda_j^w)^2 \frac{\phi_i^2}{1-\phi_i^2} \sigma^2_\varepsilon + (\lambda_j^w)^2 \sigma^2_\varepsilon + \sigma^2_{\upsilon_j} \qquad Rel(Y_{ij}) = \frac{\sigma^2_{\xi_j} + (\lambda_j^w)^2 \frac{\phi_i^2}{1-\phi_i^2} \sigma^2_\varepsilon + (\lambda_j^w)^2 \sigma^2_\varepsilon}{\sigma^2_{Y_{ij}}}$$

Intensive Longitudinal Data and Reliability

Summaries of the estimated reliability across participants and items.

	Median	Min	Max
Interested	0.79	0.78	0.91
Excited	0.78	0.77	0.90
Strong	0.79	0.78	0.89
Enthusiastic	**0.83**	0.82	0.93
Proud	0.78	0.78	0.89
Inspired	0.80	0.79	0.91
Determined	0.73	0.72	0.88
Attentive	0.77	0.76	0.89
Active	**0.66**	0.65	0.83
Alert	0.71	0.70	0.84

Get to Work! R Practical

Time for you to try these approaches!

Get to Work! R Practical

Find all the materials in the MITNB Drive Folder.

Data:

- Your own data.
- Simulated dataset (*data2rdm.dat*).
- Careless Response data from yesterday (*ELISA_CR.rds*).

Follow the practical guide: *practical.pdf*.

Get to Work! R Practical

Do not expect to work through all the approaches today!
Realistic goals:

1. Estimate the between- and within person reliability with GT, MLM, and ML-CFA.
2. (Optional) Estimate the reliability estimates base on P-technique and dynamic factor analysis. (Running this can take a couple of minutes)
3. (Optional) If you have Mplus, prepare the Mplus syntax of the 2RDM for your data. Test it with maximum 10000 iteration and no thinning.
4. (Optional) If you have Mplus, prepare the Mplus syntax of the ME-TSO for your data. Test it with maximum 5000 iterations and no thinning.

Final Thoughts - Take Home Messages

After applying these approaches to the data:

- Did you learn something interesting about your data? Something that surprised you?
- What did you find challenging?
- Do you have a preference for a certain approach?

Final Thoughts - Take Home Messages

From my experience with these approaches:

- GT and MLM estimates tend to be quite similar, with the within-person reliability usually smaller in MLM.
- It is common to find convergence issues when using GT.
- The between-person reliability in ML-CFA tends to be lower when compared to GT or MLM.
- Person-specific reliability coefficients are almost identical among P-technique, DFA, and 2RDM.
- P-technique fails more often than DFA.
- The reliability coefficients in the ME-TSO show very little variability.

Final Thoughts - Take Home Messages

Take home **messages**:

1. Reliability of ESM data should not be taken for granted.
2. When having multiple items measuring one construct, I strongly recommend using at least one of these approaches than none.
3. Which methods to use might depend on the amount of data, the design, and the goals of the study.
4. Approaches that consider the individual should be preferred.

Final Thoughts - Take Home Messages

Limitations

- Only looked at internal consistency. However, test-retest reliability coefficients for ESM have been proposed by Schuurman and Hamaker (2019) and Dejonckheere et al. (2022).
- The discussed methods are essentially for continuous, stationary, discrete-time ESM data.
- These approaches also assume that measurement invariance across time points holds.

Thank you

Sebastian Castro-Alvarez, PhD
s.castro.alvarez@rug.nl
/secastroal

References I

Brose, A., Schmiedek, F., Gerstorf, D., & Voelkle, M. C. (2020).The measurement of within-person affect variation. *Emotion*, *20*(4), 677–699. https://doi.org/10.1037/emo0000583

Castro-Alvarez, S., Bringmann, L., Back, J., & Liu, S. (2024, March). The Many Reliabilities of Psychological Dynamics: An Overview of Statistical Approaches to Estimate the Internal Consistency Reliability of Intensive Longitudinal Data. https://doi.org/10.31234/osf.io/qyk2r

Castro-Alvarez, S., Tendeiro, J. N., de Jonge, P., Meijer, R. R., & Bringmann, L. F. (2022).Mixed-Effects Trait-State-Occasion Model: Studying the Psychometric Properties and the Person–Situation Interactions of Psychological Dynamics. *Structural Equation Modeling*, *29*(3), 438–451. https://doi.org/10.1080/10705511.2021.1961587

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*(7), 917–929. https://doi.org/10.1177/0146167206287721

Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., & Mestdagh, M. (2022).Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological Assessment*, *34*(12), 1138–1154. https://doi.org/10.1037/pas0001178

Fuller-Tyszkiewicz, M., Hartley-Clark, L., Cummins, R. A., Tomyn, A. J., Weinberg, M. K., & Richardson, B. (2017).Using dynamic factor analysis to provide insights into data reliability in experience sampling studies. *Psychological Assessment*, *29*(9), 1120–1128. https://doi.org/10.1037/pas0000411

Geldhof, G. J., Preacher, K. J., & Zyphur, M. J. (2014).Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods*, *19*(1), 72–91. https://doi.org/10.1037/a0032138

Hall, M., Scherner, P. V., Kreidel, Y., & Rubel, J. A. (2021).A Systematic Review of Momentary Assessment Designs for Mood and Anxiety Symptoms. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.642044

References II

Hu, Y., Nesselroade, J. R., Erbacher, M. K., Boker, S. M., Burt, S. A., Keel, P. K., Neale, M. C., Sisk, C. L., & Klump, K. (2016).Test Reliability at the Individual Level. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(4), 532–543. https://doi.org/10.1080/10705511.2016.1148605

Lai, M. H. C. (2021).Composite reliability of multilevel data: It's about observed scores and construct meanings. *Psychological Methods*, *26*(1), 90–102. https://doi.org/10.1037/met0000287

Nezlek, J. B. (2017).A practical guide to understanding reliability in studies of within-person variability. *Journal of Research in Personality*, *69*, 149–155. https://doi.org/10.1016/j.jrp.2016.06.020

O'Laughlin, K. D., Liu, S., & Ferrer, E. (2021).Use of Composites in Analysis of Individual Time Series: Implications for Person-Specific Dynamic Parameters. *Multivariate Behavioral Research*, *56*(3), 408–425. https://doi.org/10.1080/00273171.2020.1716673

Schuurman, N. K., & Hamaker, E. L. (2019).Measurement error and person-specific reliability in multilevel autoregressive modeling. *Psychological Methods*, *24*(1), 70–91. https://doi.org/10.1037/met0000188

Williams, D. R., Martin, S. R., Liu, S., & Rast, P. (2020).Bayesian Multivariate Mixed-Effects Location Scale Modeling of Longitudinal Relations Among Affective Traits, States, and Physical Activity. *European Journal of Psychological Assessment*, *36*(6), 981–997. https://doi.org/10.1027/1015-5759/a000624

Xiao, Y., Wang, P., & Liu, H. (2023).Assessing intra- and inter-individual reliabilities in intensive longitudinal studies: A two-level random dynamic model-based approach. *Psychological Methods*. https://doi.org/10.1037/met0000608