

Person- & Item-specific Reliability for Intensive Longitudinal Data

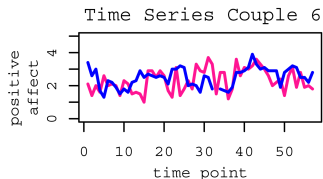
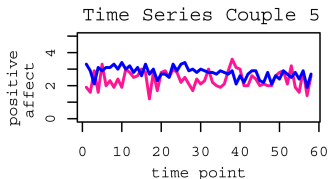
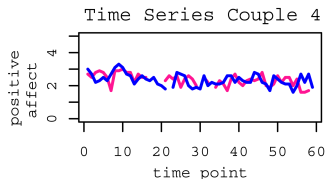
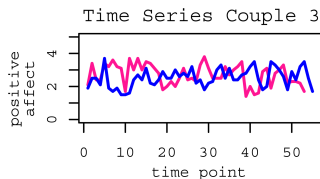
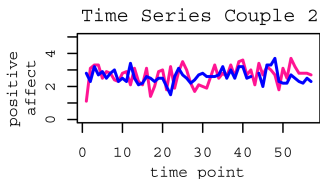
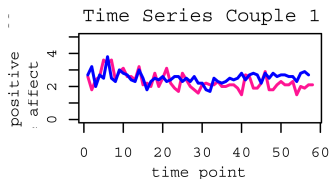
Noémi K. Schuurman

Utrecht University

ICPS March 2023

Intensive Longitudinal Data

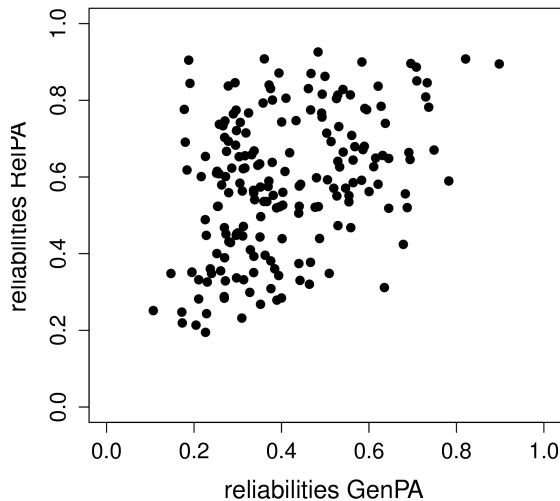
Many repeated measures, for many individuals.



Data collected by Ferrer, Steele & Hsieh (2012)

Data & Results

- ▶ Data from study by Ferrer, Steele, and Hsieh (2012)
- ▶ 190 women filled out a diary every evening
- ▶ about 60 to 90 repeated measures on daily General Positive Affect and Relationship Positive Affect



What is Measurement Error/Reliability?

Reliability concerns the consistency of measurements across "replications".

What is Measurement Error/Reliability?

Reliability concerns the consistency of measurements across "replications".



Reliability is the proportion of variance in a variable that is *not* due to *random* measurement error.

Image from: <https://www.floppycats.com/how-to-weigh-your-cat.html>

Classical Cross-sectional Approaches to Reliability

Types of Reliability Measures

- ▶ Test-Retest Reliability
- ▶ Parallel Test Reliability
- ▶ Internal Consistency Reliability

Classical Cross-sectional Approaches to Reliability

Types of Reliability Measures

- ▶ Test-Retest Reliability
- ▶ Parallel Test Reliability
- ▶ Internal Consistency Reliability

Context: Investigating stable between-person differences.

In all these classic reliability measures **within person fluctuations are considered measurement error/noise.**

Longitudinal Data and Reliability

- ▶ Separate between-person from within-person variation (if multiple participants)

Longitudinal Data and Reliability

- ▶ Separate between-person from within-person variation (if multiple participants)
- ▶ Separate out 'true' within-person fluctuations from random measurement errors.

Longitudinal Data and Reliability

- ▶ Separate between-person from within-person variation (if multiple participants)
- ▶ Separate out 'true' within-person fluctuations from random measurement errors.
 - Separate out what is 'systematic dynamic process' from 'random measurement error'.

Longitudinal Data and Reliability

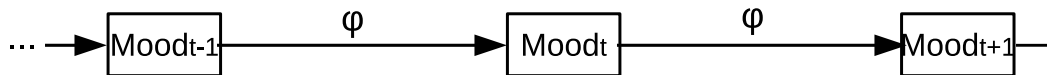
- ▶ Separate between-person from within-person variation (if multiple participants)
- ▶ Separate out 'true' within-person fluctuations from random measurement errors.
 - Separate out what is 'systematic dynamic process' from 'random measurement error'.

Important for Intensive Longitudinal Data:

- ▶ person-specific (personalized data, feedback, and analyses)
- ▶ item-specific (burden, theory concerns)

Autoregressive Model

- ▶ Model for the time series of a specific person ($N=1$, $T=\text{many}$)
- ▶ Variable is regressed on itself at (a) previous occasion(s)
- ▶ AR(1) model: on the nearest previous occasion

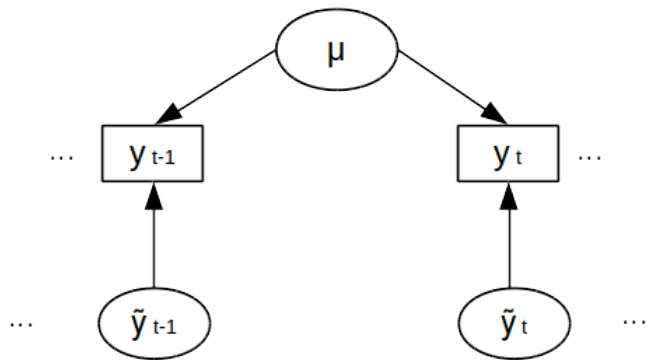


The N=1 AR(1) Model



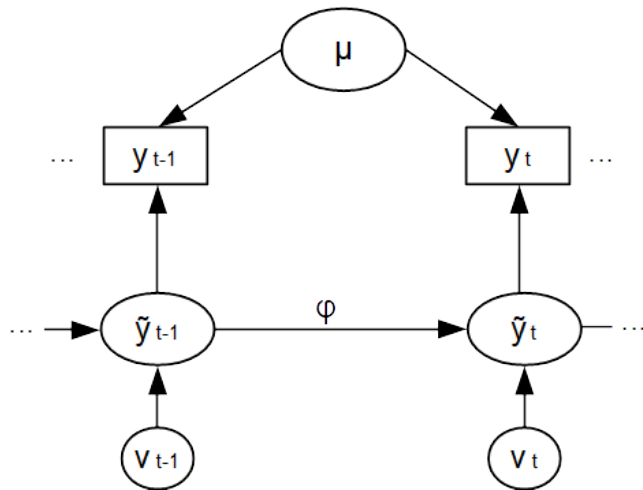
Mood t	Mood t-1
5	.
3	5
3	3
4	3
2	4
3	2
1	3
1	1
2	1
.	2

Autoregressive Model



$$y_t = \mu + \tilde{y}_t$$

Autoregressive Model

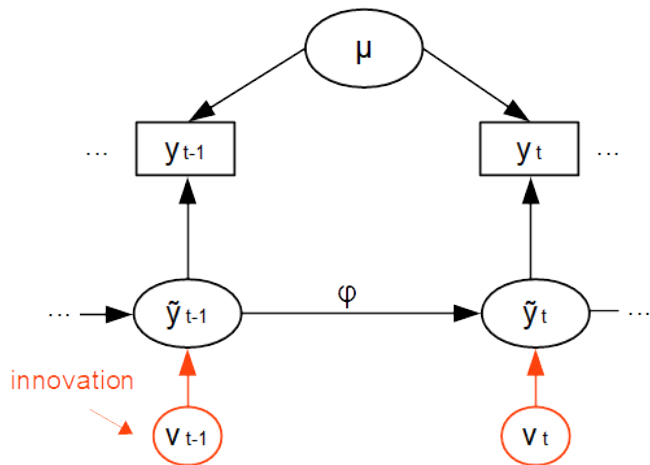


$$y_t = \mu + \tilde{y}_t$$

$$\tilde{y}_t = \phi \tilde{y}_{t-1} + v_t$$

$$v_t \sim N(0, \omega^2)$$

Autoregressive Model

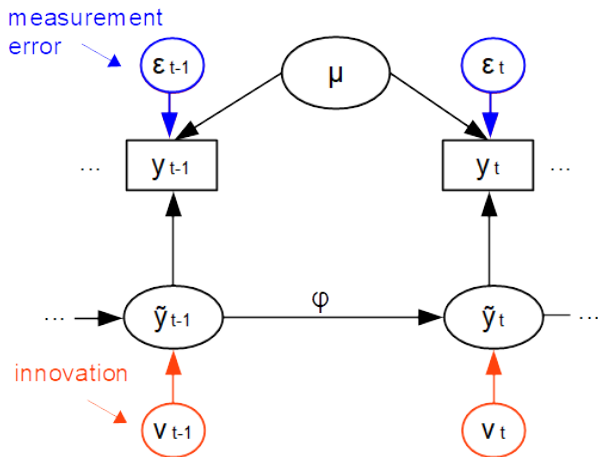


$$y_t = \mu + \tilde{y}_t$$

$$\tilde{y}_t = \phi_i \tilde{y}_{t-1} + v_t$$

$$v_t \sim N(0, \omega^2)$$

Autoregressive Model with Measurement Error



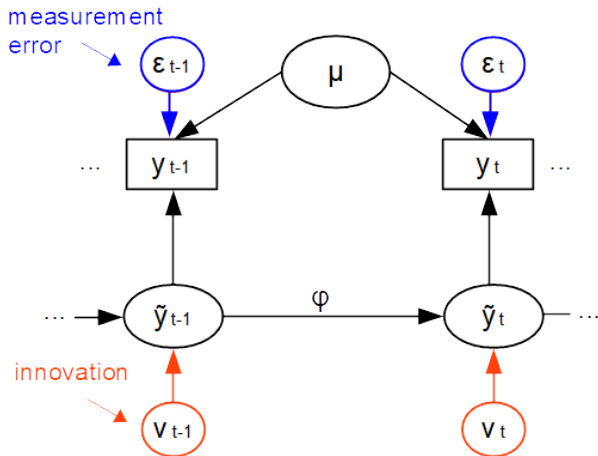
$$y_t = \mu + \tilde{y}_t + \epsilon_t$$

$$\tilde{y}_t = \phi \tilde{y}_{t-1} + v_t$$

$$v_t \sim N(0, \omega^2)$$

$$\epsilon_t \sim N(0, \sigma^2)$$

Autoregressive Model with Measurement Error



$$y_t = \mu + \tilde{y}_t + \epsilon_t$$

$$\tilde{y}_t = \phi \tilde{y}_{t-1} + v_t$$

$$v_t \sim N(0, \omega^2)$$

$$\epsilon_t \sim N(0, \sigma^2)$$

$$Reliability = 1 - \frac{\sigma^2}{\text{var}(y)}$$

Dynamic Model

Caveats

- ▶ Requires Bayesian estimation.
- ▶ Requires many repeated measures (multilevel modeling helps).
- ▶ Model assumptions (fix with extensions).
- ▶ Requires equal time intervals between repeated measures (fix with missing observations).

Dynamic Model

Caveats

- ▶ Requires Bayesian estimation.
- ▶ Requires many repeated measures (multilevel modeling helps).
- ▶ Model assumptions (fix with extensions).
- ▶ Requires equal time intervals between repeated measures (fix with missing observations).

Extensions

- ▶ Multilevel for multiple participants (Schoorman, 2016; Schoorman & Hamaker, 2019)
- ▶ Multivariate with correlated measurement errors (Schoorman, 2016; Schoorman & Hamaker, 2019)
- ▶ Other dynamic models...
- ▶ Etc!

Dynamic Model

Caveats

- ▶ Requires Bayesian estimation.
- ▶ Requires many repeated measures (multilevel modeling helps).
- ▶ Model assumptions (fix with extensions).
- ▶ Requires equal time intervals between repeated measures (fix with missing observations).

Extensions

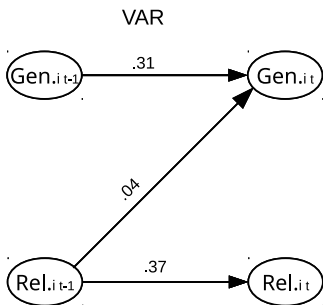
- ▶ Multilevel for multiple participants (Schuurman, 2016; Schuurman & Hamaker, 2019)
- ▶ Multivariate with correlated measurement errors (Schuurman, 2016; Schuurman & Hamaker, 2019)
- ▶ Other dynamic models...
- ▶ Etc!

Usage

- ▶ As a tool to estimate reliability for single items & persons.
- ▶ As a measurement model - that corrects your dynamic model for measurement error.

Correcting for ME Matters

Empirical Example: General PA and Relationship PA

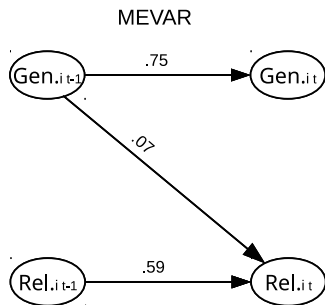


mean ϕ_{geni} : .31 (.28, .34)

mean ϕ_{reli} : .37 (.34, .40)

mean $\phi_{gen \rightarrow reli}$: .04 (.02, .07)

mean $\phi_{rel \rightarrow geni}$: .02 (.00, .04)



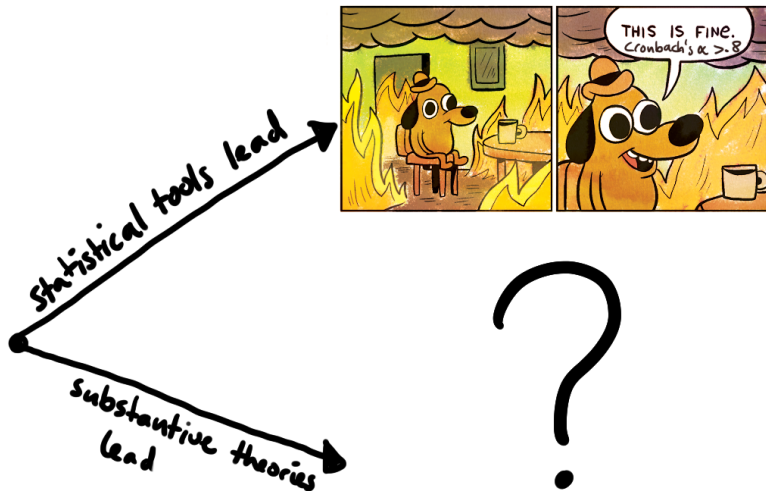
mean ϕ_{geni} : .75 (.69, .80)

mean ϕ_{reli} : .59 (.53, .64)

mean $\phi_{gen \rightarrow reli}$: -.03 (-.07, .00)

mean $\phi_{rel \rightarrow geni}$: .07 (.02, .13)

Substantive & Measurement Theory should Lead

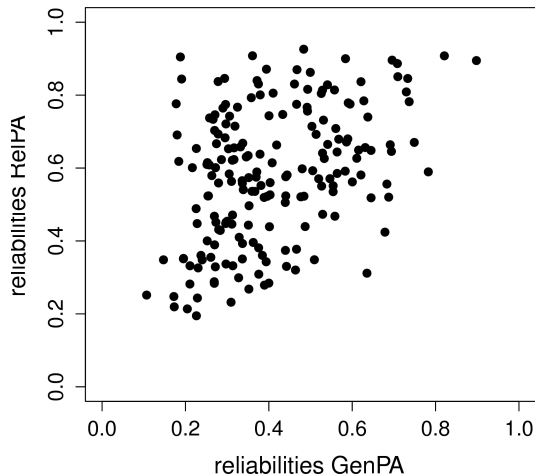


www.nkschuurman.com

- ▶ **Single subject:** Schuurman, N.K., Houtveen, J.H., Hamaker, E.L. (2015). Incorporating measurement error in $n=1$ psychological autoregressive modeling. *Frontiers in Psychology*.
- ▶ **Multilevel, open software:** Schuurman, N.K. (2016). Multilevel Autoregressive Modeling in Psychology: Snags and Solutions. Chapter 5. (pdf on website)
- ▶ **Multilevel, Mplus/DSEM:** Schuurman, N.K. Hamaker, E.L. (2019). Measurement error and person-specific reliability in multilevel autoregressive models. *Psychological Methods*.

Person-specific reliabilities

But why are those reliabilities so low?



Person-specific reliabilities

But why are those reliabilities so low?

- ▶ The 'measurement error' term is not only capturing truly random measurement error.

Person-specific reliabilities

But why are those reliabilities so low?

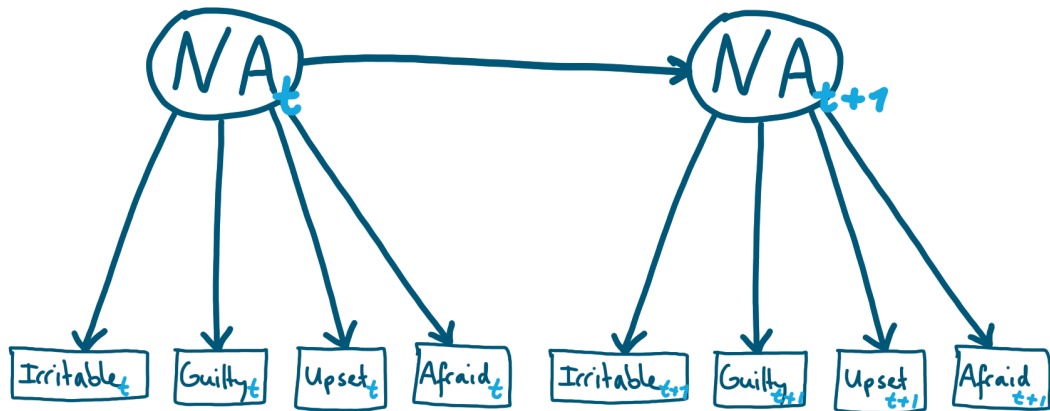
- ▶ The 'measurement error' term is not only capturing truly random measurement error.
- ▶ It captures ANY within person fluctuations that are unique to one measurement occasion.

Person-specific reliabilities

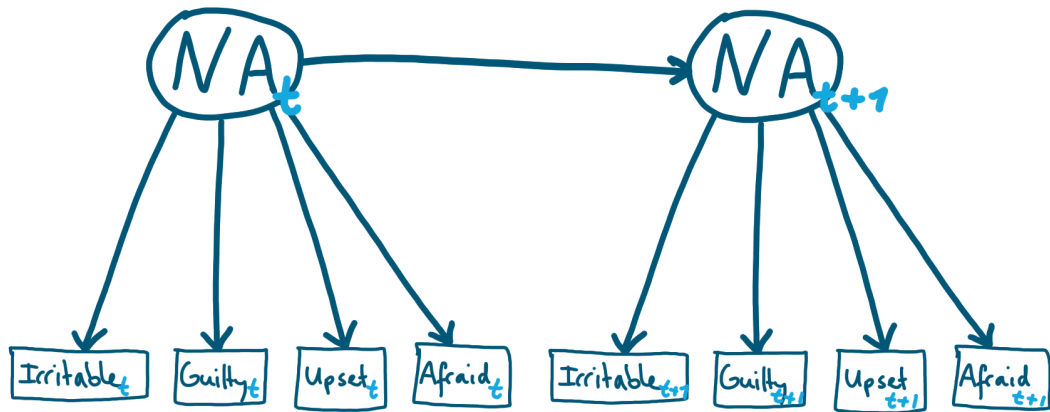
But why are those reliabilities so low?

- ▶ The 'measurement error' term is not only capturing truly random measurement error.
- ▶ It captures ANY within person fluctuations that are unique to one measurement occasion.
- ▶ Notable: I allowed the measurement errors of RelPA and GenPA to be correlated - and they were correlated ~ 0.25 on average across persons.
- ▶ Our reliability estimate is a lower bound reliability (all reliability estimates are).

Factor modeling as a tool for filtering out ME

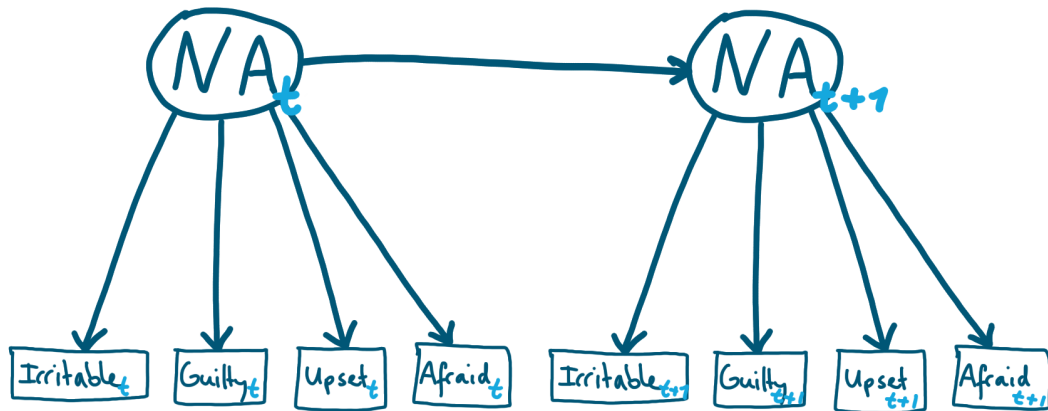


Factor modeling as a tool for filtering out ME



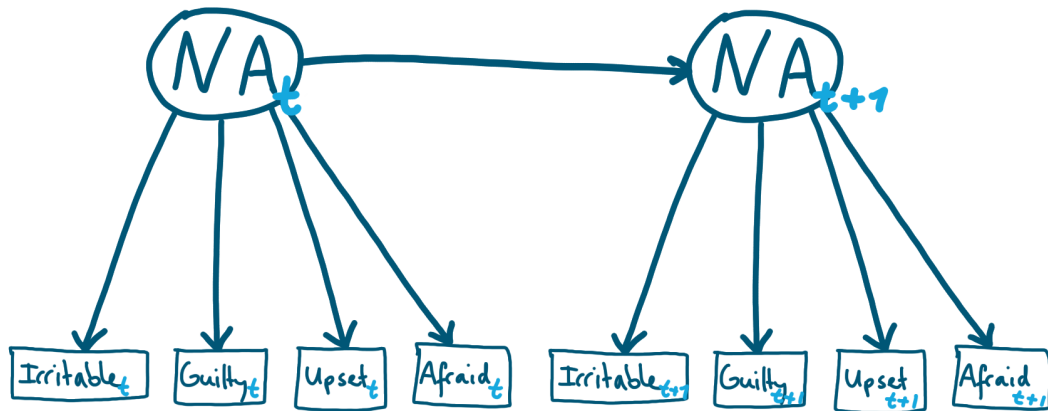
- Are they really exchangeable, parallel items, that measure the same thing?

Factor modeling as a tool for filtering out ME



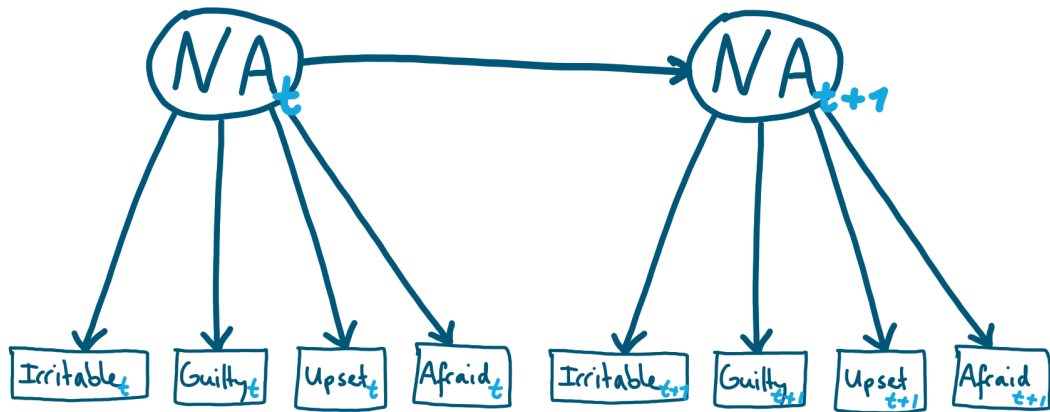
- ▶ Are they really exchangeable, parallel items, that measure the same thing?
- ▶ Do we really expect these items to all usually increase and decrease together at each occasion?

Factor modeling as a tool for filtering out ME



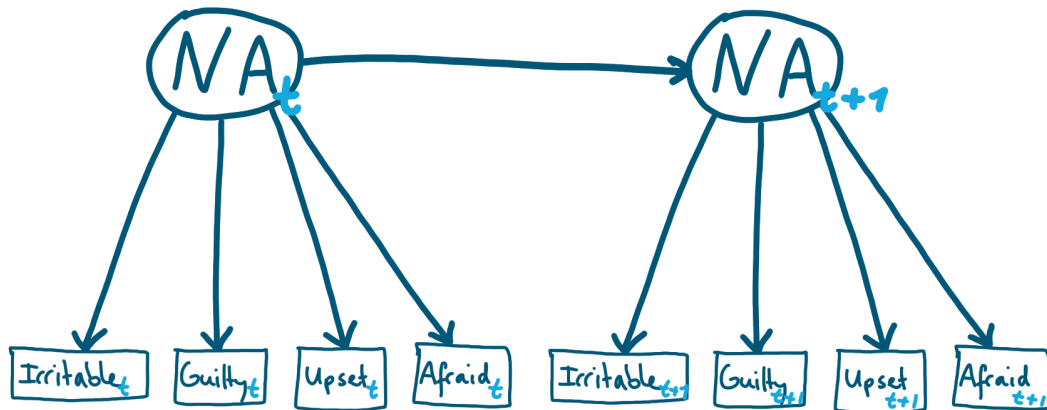
- ▶ Are they really exchangeable, parallel items, that measure the same thing?
- ▶ Do we really expect these items to all usually increase and decrease together at each occasion?
- ▶ Is that the case for all persons?

Factor modeling as a tool for filtering out ME



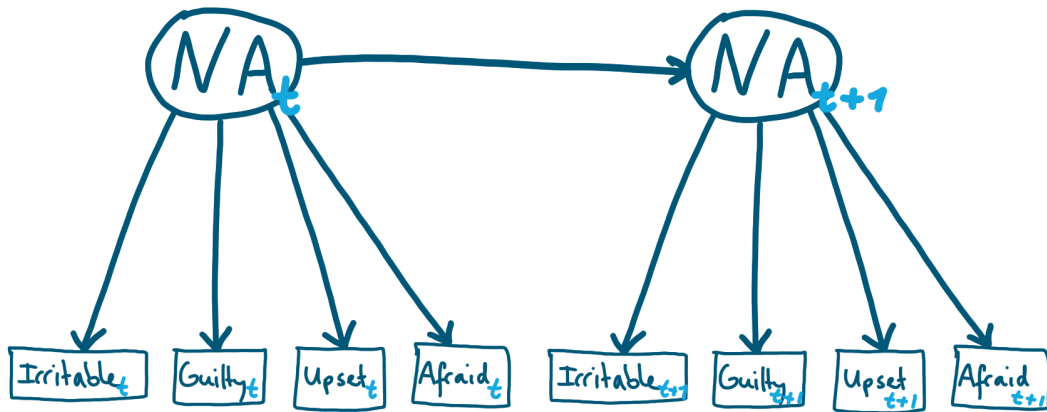
- Could we intervene directly on the latent variable for any particular person?

Factor modeling as a tool for filtering out ME



- ▶ Could we intervene directly on the latent variable for any particular person?
- ▶ Is the latent variable something real that acts within a person?

Factor modeling as a tool for filtering out ME



- ▶ Could we intervene directly on the latent variable for any particular person?
- ▶ Is the latent variable something real that acts within a person?
- ▶ How do changes in the latent variable result exactly in the observed scores?