# Reliability of ESM Assessments of Mood and Mood Sensitivity

Wolfgang Viechtbauer
Maastricht University, The Netherlands
KU Leuven, Belgium
http://www.wvbauer.com

## Reliability of ESM Data

- discuss two aspects:
  - reliability of the measurements themselves
  - reliability of (within-person) relationships
- explain some approaches to assess these
- use measurement of mood (positive and negative affect) and mood reactivity as an illustrate example



## Illustrative Data

- merged dataset from 9 different ESM studies
- 10 semi-random beeps in 90-min intervals between 7:30am and 10:30pm for 4-6 days
- only using data from day 1 to 4



## Illustrative Data

| | controls | at risk | psychotic | depressed |
|---|---|---|---|---|
| n | 693 | 214 | 235 | 125 |
| % females | 85% | 55% | 34% | 76% |
| mean age | 30.6 | 37.2 | 35.3 | 44.0 |
| mean number of responses (SD) | 27.3 (6.3) | 29.7 (5.8) | 27.5 (6.4) | 30.9 (5.3) |
| range | 14-40 | 15-40 | 14-40 | 16-40 |
| total responses | 18,945 | 6,357 | 6,452 | 3,866 |

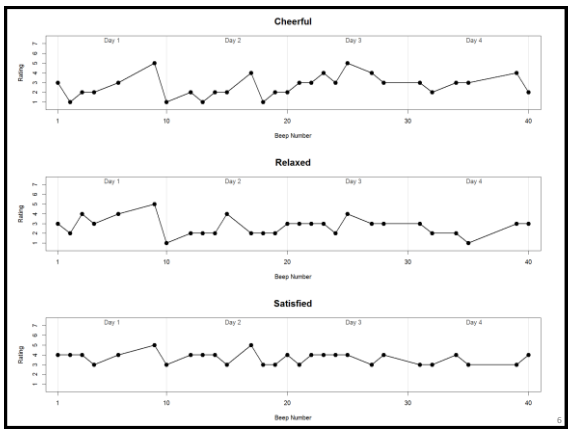


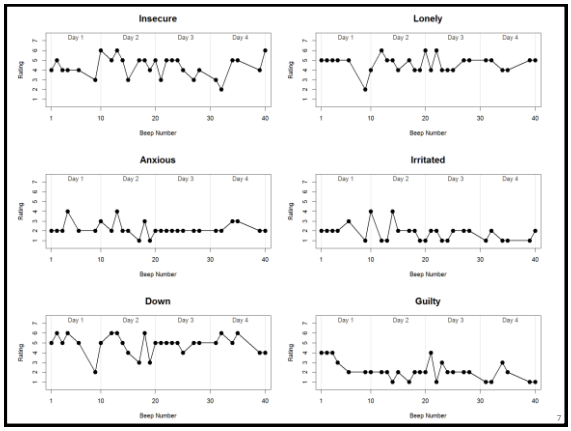
## Mood Items

| Positive Affect | Negative Affect |
|---|---|
| cheerful | insecure |
| relaxed | lonely |
| satisfied | anxious |
| | irritated |
| | down |
| | guilty |

(each item rated on 1 – 7 scale)








## Reliability of PA and NA

- cannot examine reliability by examining consistency of measurements over time
- but can examine consistency of multiple items measuring the same construct
- analysis approach: multilevel factor analysis

| subject | day | beep | cheerful | relaxed | satisfied | insecure | ... |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 3 | 3 | 4 | 4 | ... |
| 1 | 1 | 2 | 1 | 2 | 4 | 5 | ... |
| ... | ... | ... | ... | ... | ... | ... | ... |

| subject | day | beep | item | y |
|---|---|---|---|---|
| 1 | 1 | 1 | cheerful | 3 |
| 1 | 1 | 1 | relaxed | 3 |
| 1 | 1 | 1 | satisfied | 4 |
| 1 | 1 | 1 | insecure | 4 |
| ... | ... | ... | ... | ... |
| 1 | 1 | 2 | cheerful | 1 |
| 1 | 1 | 2 | relaxed | 2 |
| 1 | 1 | 2 | satisfied | 4 |
| 1 | 1 | 2 | insecure | 5 |
| ... | ... | ... | ... | ... |

## Multivariate Multilevel Model

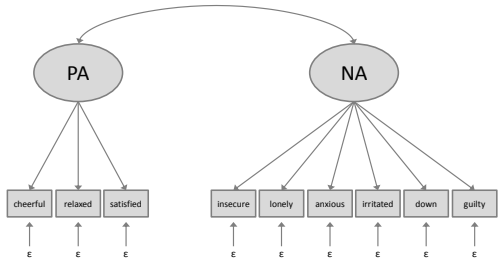
$$y_{ijk} = \mu_k + u_{ik} + u_{ijk}$$

$y_{ijk}$ = response of person $i$ at beep $j$ to item $k$
$\mu_k$ = mean of item $k$ (averaged over persons and beeps)
$u_{ik}$ = random effects for items at the person level
$u_{ijk}$ = random effects for items at the beep level

$$\text{Var}\begin{bmatrix} u_{i1} \\ u_{i2} \\ u_{i3} \\ \dots \end{bmatrix} = \begin{bmatrix} \tau_1^2 & \rho_{12}\tau_1\tau_2 & \rho_{13}\tau_1\tau_3 & \dots \\ \dots & \tau_2^2 & \rho_{23}\tau_2\tau_3 & \dots \\ \dots & \dots & \tau_3^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix} \quad \text{Var}\begin{bmatrix} u_{ij1} \\ u_{ij2} \\ u_{ij3} \\ \dots \end{bmatrix} = \begin{bmatrix} v_1^2 & \phi_{12}v_1v_2 & \phi_{13}v_1v_3 & \dots \\ \dots & v_2^2 & \phi_{23}v_2v_3 & \dots \\ \dots & \dots & v_3^2 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix}$$

## Two-Factor CFA Model

based on estimated correlations at the person and beep level, fit two-factor CFA model




## Estimate Reliability

then estimate reliability from CFA model using McDonald's omega

$$\omega = \frac{\left(\sum_{i=1}^{p} \hat{\lambda}_i\right)^2}{\left(\sum_{i=1}^{p} \hat{\lambda}_i\right)^2 + \sum_{i=1}^{p} \hat{\sigma}_i^2}$$

$p$ = number of items loading on factor
$\hat{\lambda}_i$ = loading of item $i$
$\hat{\sigma}_i^2$ = error variance of item $i$

## Results

**– Person Level –**

| | controls | at risk | psychotic | depressed |
|---|---|---|---|---|
| PA | 0.93 | 0.95 | 0.89 | 0.94 |
| NA | 0.91 | 0.94 | 0.93 | 0.91 |

**– Beep Level –**

| | controls | at risk | psychotic | depressed |
|---|---|---|---|---|
| PA | 0.75 | 0.73 | 0.71 | 0.83 |
| NA | 0.66 | 0.65 | 0.70 | 0.76 |

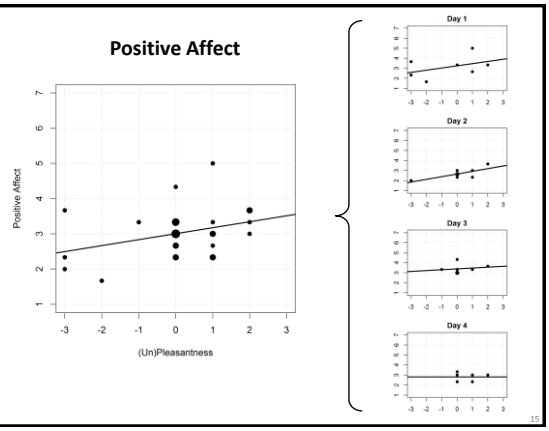
13

## Reliability of Relationships

- often interested in within-person relationships
- example: (un)pleasantness of most important event since previous beep (rated -3 to +3) and PA/NA (rated on 1 to 7 scale)

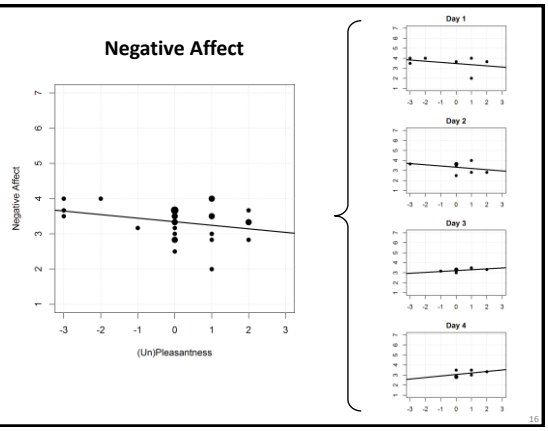
14

## Positive Affect




15

## Negative Affect




16

## Model with Random Effects per Day

$$y_{ijk} = (\beta_{0j}+u_{0ij}) + (\beta_{1j}+u_{1ij})x_{ijk} + \epsilon_{ijk}$$

$y_{ijk}$ = response (PA/NA) for person $i$, day $j$, beep $k$
$x_{ijk}$ = predictor value (pleasantness) of person $i$, day $j$, beep $k$
$\beta_{0j}$ = average intercept on day $j$
$\beta_{1j}$ = average slope on day $j$
$u_{0ij}$ = random intercept effect for person $i$ on day $j$
$u_{1ij}$ = random slope effect for person $i$ on day $j$
$\epsilon_{ijk}$ = error

17

## Model with Random Effects per Day

$$y_{ijk} = (\beta_{0j}+u_{0ij}) + (\beta_{1j}+u_{1ij})x_{ijk} + \epsilon_{ijk}$$

$$\text{Var}\begin{bmatrix} u_{0i1} \\ u_{0i2} \\ u_{0i3} \\ u_{0i4} \\ u_{1i1} \\ u_{1i2} \\ u_{1i3} \\ u_{1i4} \end{bmatrix} = \left[\begin{array}{cccc|cccc} \tau_{01}^2 & & \text{intercept} & & & \text{intercept-} & & \\ & \tau_{02}^2 & \text{covariances} & & & \text{slope} & & \\ & & \tau_{03}^2 & & & \text{covariances} & & \\ & & & \tau_{04}^2 & & & & \\ \hline & & & & \tau_{11}^2 & & \text{slope} & \\ & & & & & \tau_{12}^2 & \text{covariances} & \\ & & & & & & \tau_{13}^2 & \\ & & & & & & & \tau_{14}^2 \end{array}\right]$$

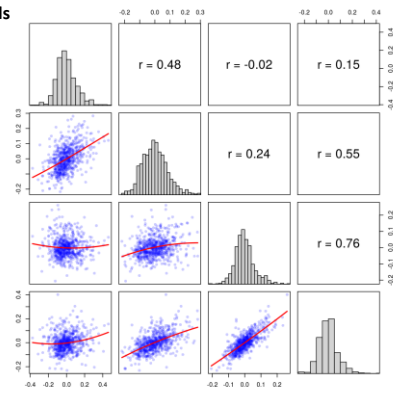
18

## Slope Reliability

- compute Cronbach's alpha (or some similar measure) based on var-cov / correlation matrix of the random slope effects
- visualize by plotting predicted values of random effects (BLUPs) against each other

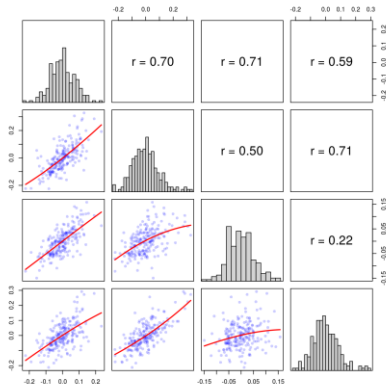
19

**controls**



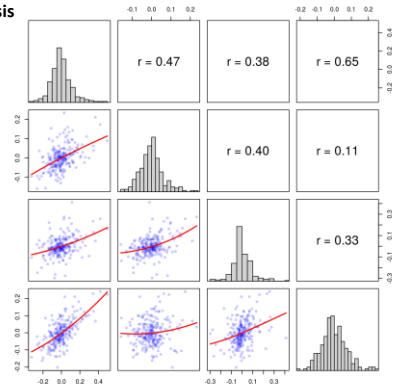

20

**at risk**



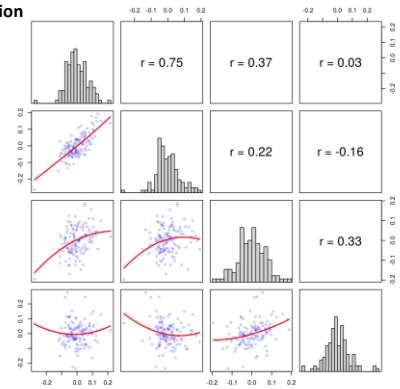

21

**psychosis**




22

**depression**




23

## Results

- estimated reliability of the mood sensitivity relationship based on 4 days:

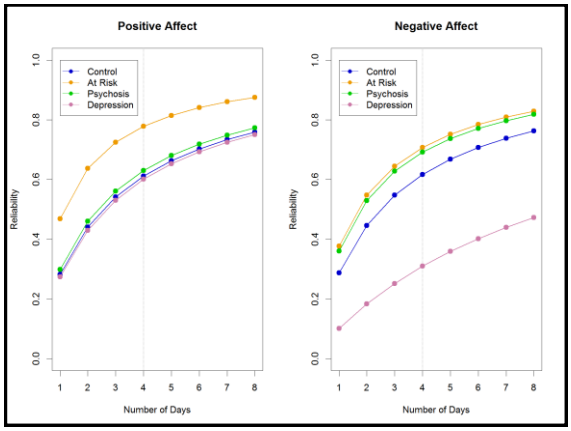
| | controls | at risk | psychotic | depressed |
|---|---|---|---|---|
| PA | 0.61 | 0.78 | 0.63 | 0.60 |
| NA | 0.62 | 0.71 | 0.69 | 0.31 |

24

## Projected Reliabilities

- use Spearman-Brown equation to estimate reliability for a different number of days

25




## Final Points

- reliability of measurements ≠ reliability of (within-person) associations
- if association is not a stable trait (over the days of the study), underestimate reliability

27