

Analyzing ESM Time Series

jonashaslbeck@gmail.com

May 10th, 2022

Discussion of Time Series Analysis

Questions:

1. Characterize Emotion Dynamics
2. Heterogeneity between people?
3. Effectiveness of intervention?

Three Sessions:

1. Looking at data, descriptives, basic modeling (Jonas)
2. Modeling and model selection (Jerry)
3. Non-stationarity / trends and causality (Lourens)

... with Eiko chiming in here and there!

All analyses shown in this session are fully reproducible and available at <http://github.com/jmbh/NSMDworkshopMay22>.

The Example Dataset

Open data available from Rowland, Wenzel, and Kubiak (2020).

- Goal: study effect of mindfulness on affect network
- 125 undergraduate students
- 4 positive, 4 negative emotions
- 40 consecutive days
- six measures / day
- so, maximum of 240 measurements
- response scale: 0-100 analogue slider
- Two groups: Mindfulness intervention vs. Control

The Example Dataset

```
data <- readRDS("ESMdata.RDS")
head(data)
```

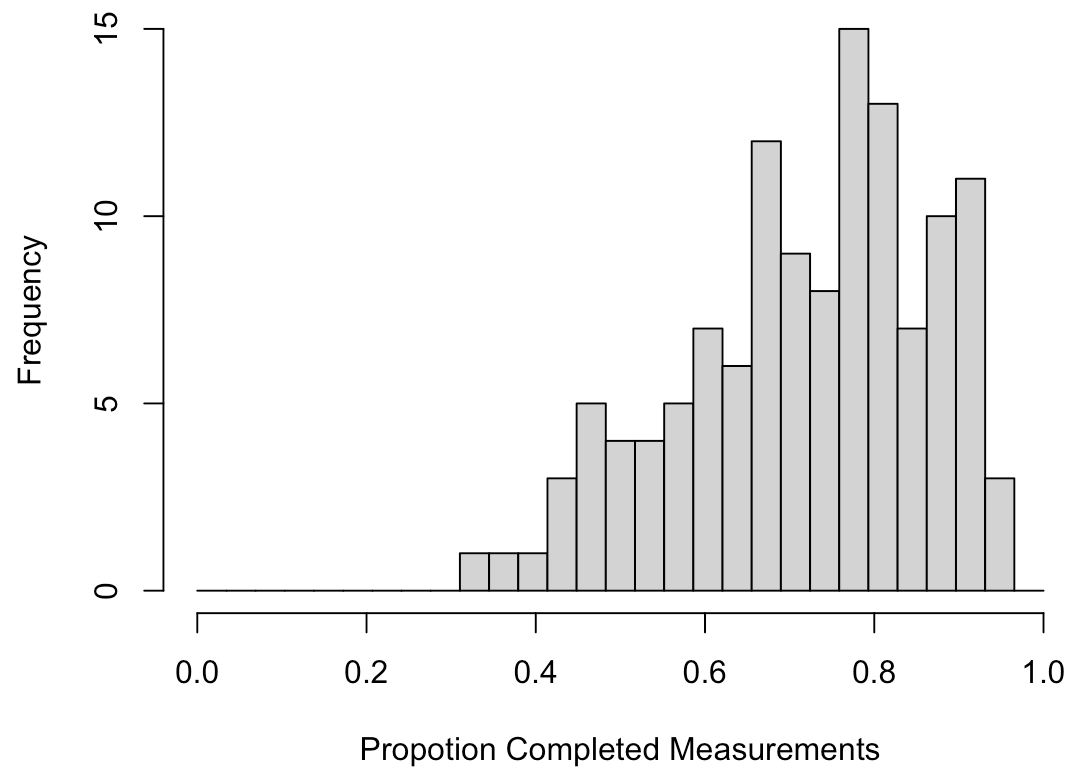
```
##   subj_id dayno beep group happy excited relaxed satisfied angry anxious
## 1      1     1   1     2    79      64      30        74    19      7
## 2      1     1   2     2    NA      NA      NA        NA    NA     NA
## 3      1     1   3     2    20      24      37        36    62     41
## 4      1     1   4     2    25      24      33        29    42     68
## 5      1     1   5     2    37      24      40        17    35     36
## 6      1     1   6     2    NA      NA      NA        NA    NA     NA
##   depressed sad
## 1         8  58
## 2        NA  NA
## 3        73  67
## 4        77  81
## 5        12  76
## 6        NA  NA
```

```
length(unique(data$subj_id))
```

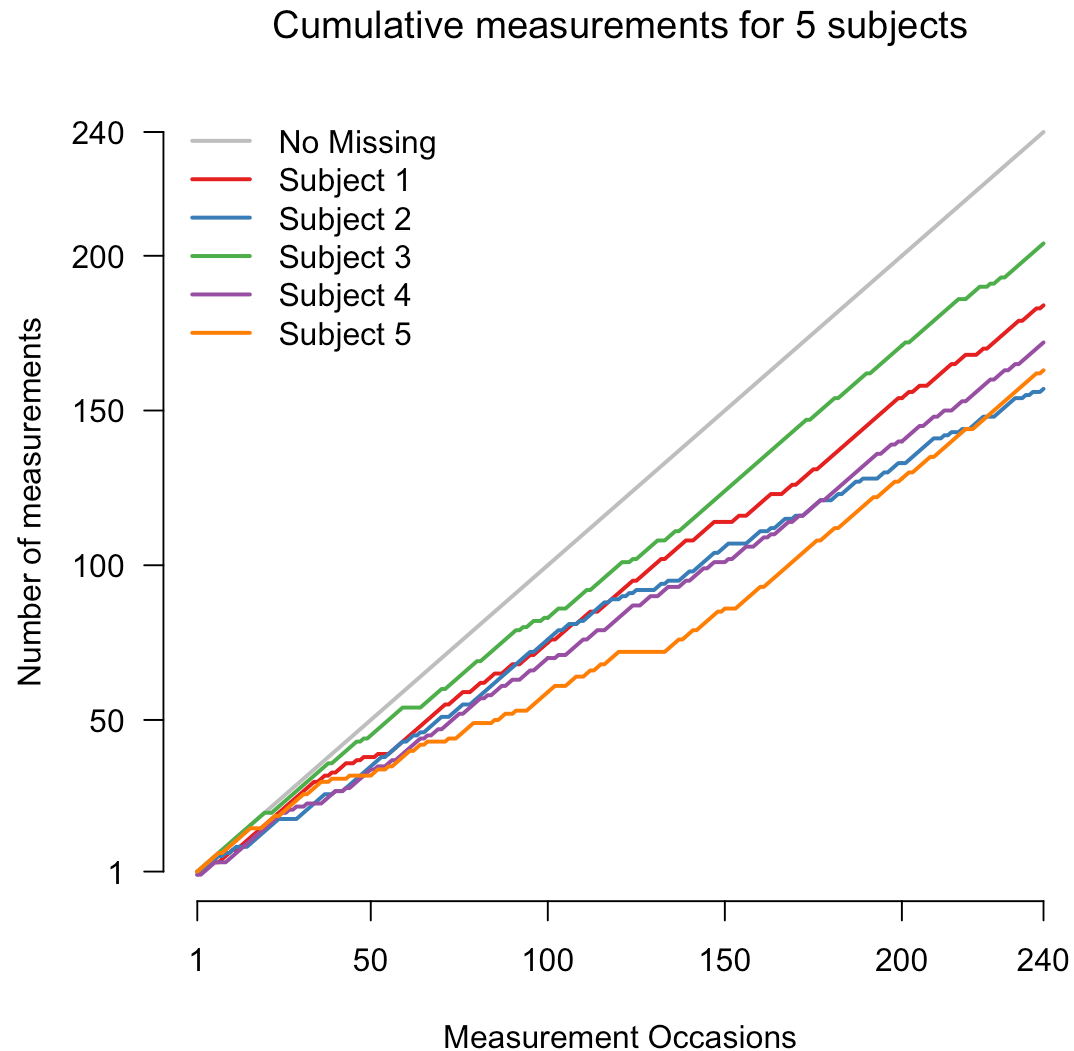
```
## [1] 125
```

Overall Missingness across Individuals

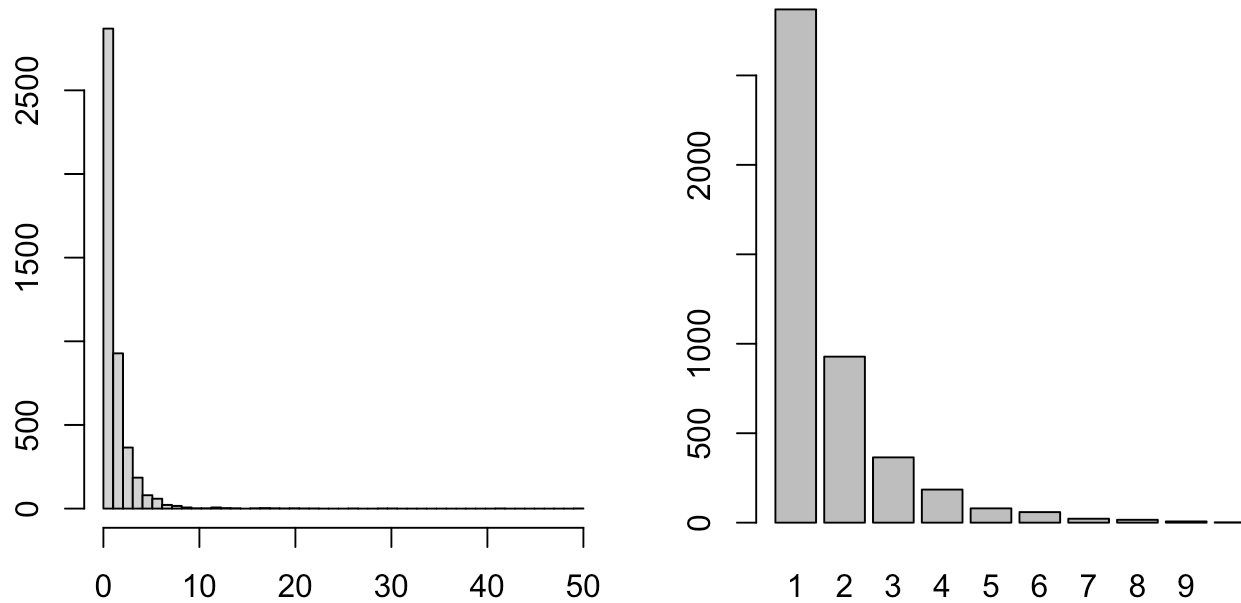
First check: Proportion of completed measurements across individuals:



Missingness across time

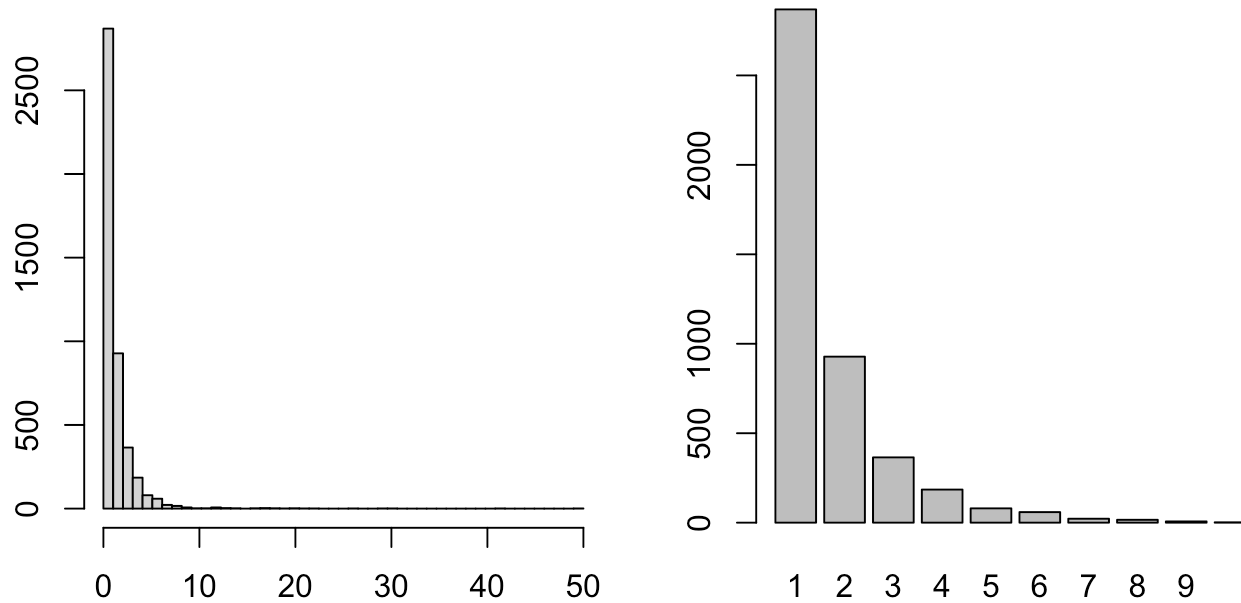


Length of missing sequences



Most missing sequences are only a few measurement occasions long

Length of missing sequences

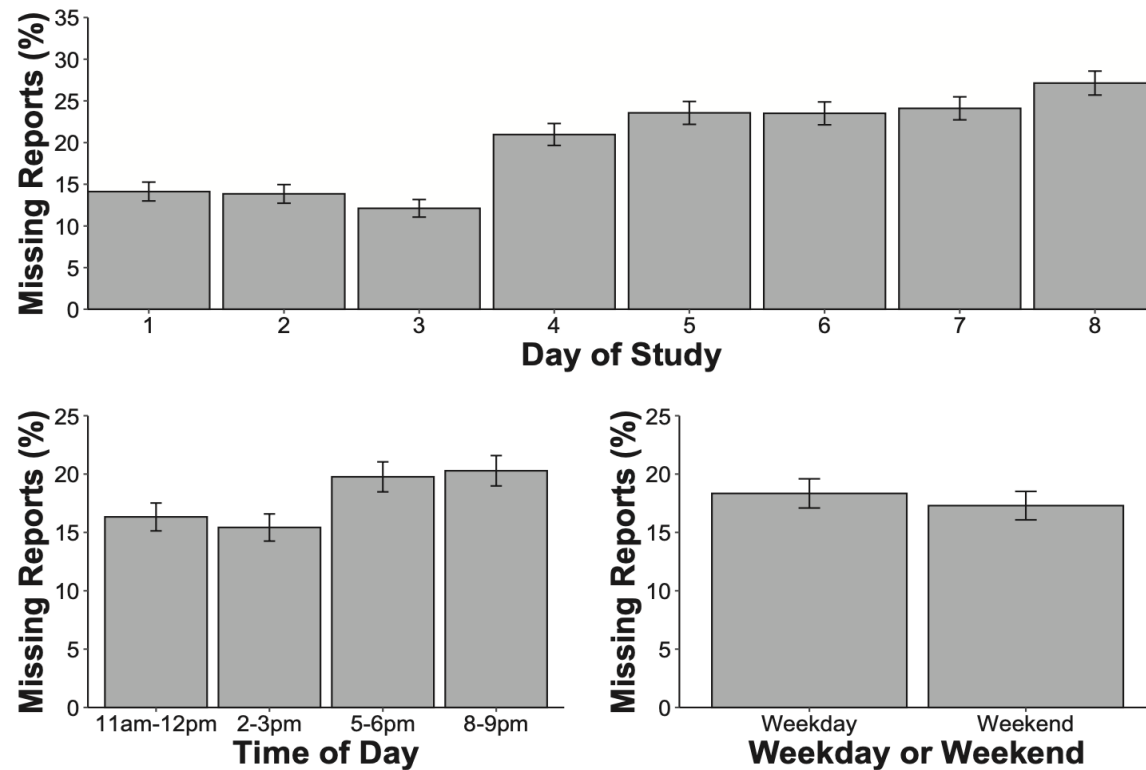


Most missing sequences are only a few measurement occasions long

- “Good” for trend-type analyses
- “Bad” for lagged-effects-type analyses

Is missingness random?

Time effects:



(from Sun, Rhemtulla, and Vazire (2021))

Is missingness random? II

Missingness is likely caused by the characteristics of people we are studying.

Table 1. Effects of Between-Person Factors on Nonresponse.

Model and Predictors	<i>b</i>	<i>SE</i>
1. Age	−.028	.021
1. Gender	−.199	.098
2. Positive schizotypy	.112	.044
2. Negative schizotypy	−.001	.044
3. Anxiety (BAI)	.050	.078
3. Depression (BDI)	.152	.074
4. Social anxiety	−.025	.052
5. Emotional expressivity	−.015	.041
6. ADHD: inattentiveness	.055	.096
6. ADHD: hyperactivity	.141	.089
7. Hypomania	.249	.080

(from Silvia et al. (2013))

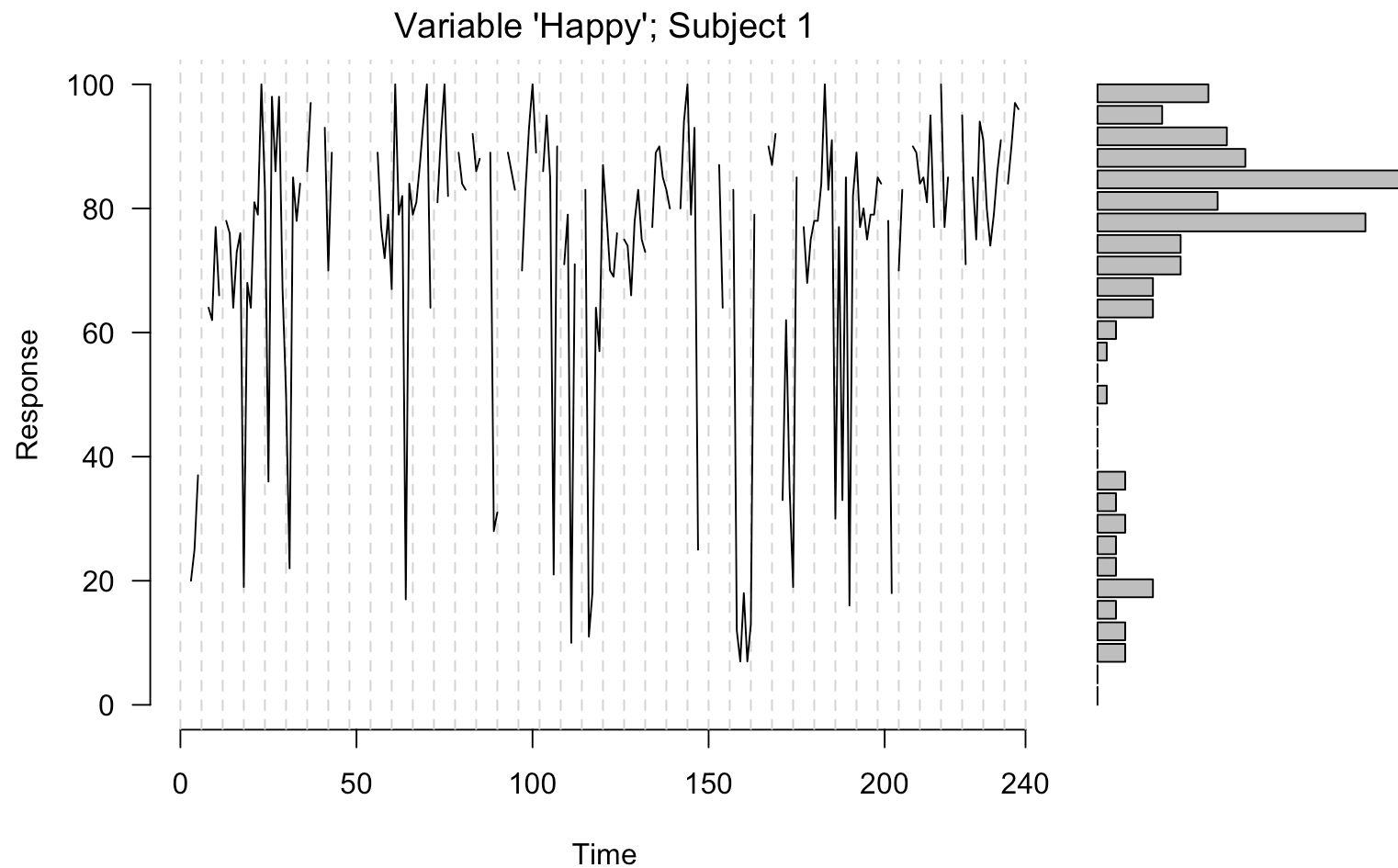
There are also results on dynamic predictors of missingness, see Rintala et al. (2020) and Sun, Rhemtulla, and Vazire (2021).

Looking at Data

```
head(data)
```

```
##      subj_id dayno beep group happy excited relaxed satisfied angry anxious
## 1          1     1    1     2    79      64      30       74     19      7
## 2          1     1    2     2    NA      NA      NA       NA     NA     NA
## 3          1     1    3     2    20      24      37       36     62     41
## 4          1     1    4     2    25      24      33       29     42     68
## 5          1     1    5     2    37      24      40       17     35     36
## 6          1     1    6     2    NA      NA      NA       NA     NA     NA
##      depressed sad
## 1           8  58
## 2          NA  NA
## 3          73  67
## 4          77  81
## 5          12  76
## 6          NA  NA
```

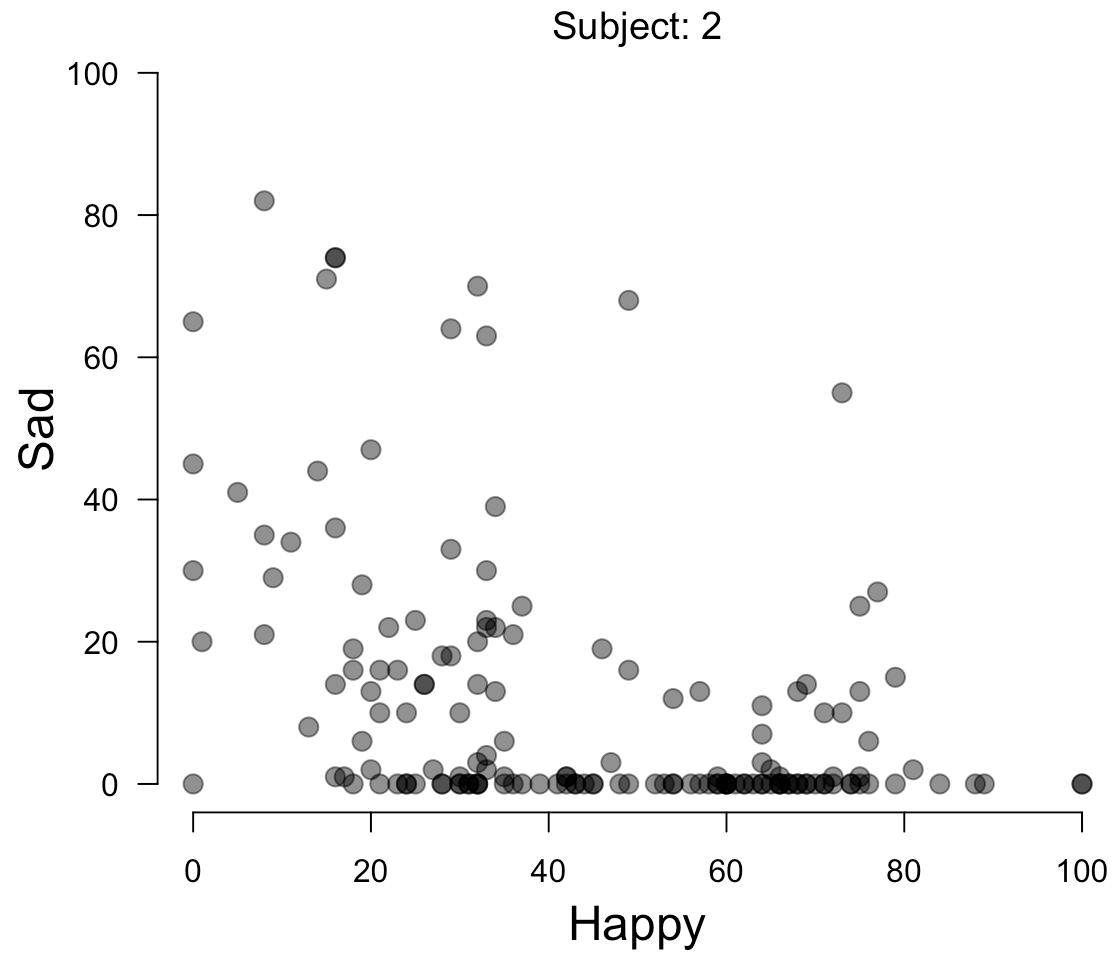
Looking at Data: Time Series + Marginals



Looking at Data: Time Series + Marginals

Heterogeneity: see [PDF](#)

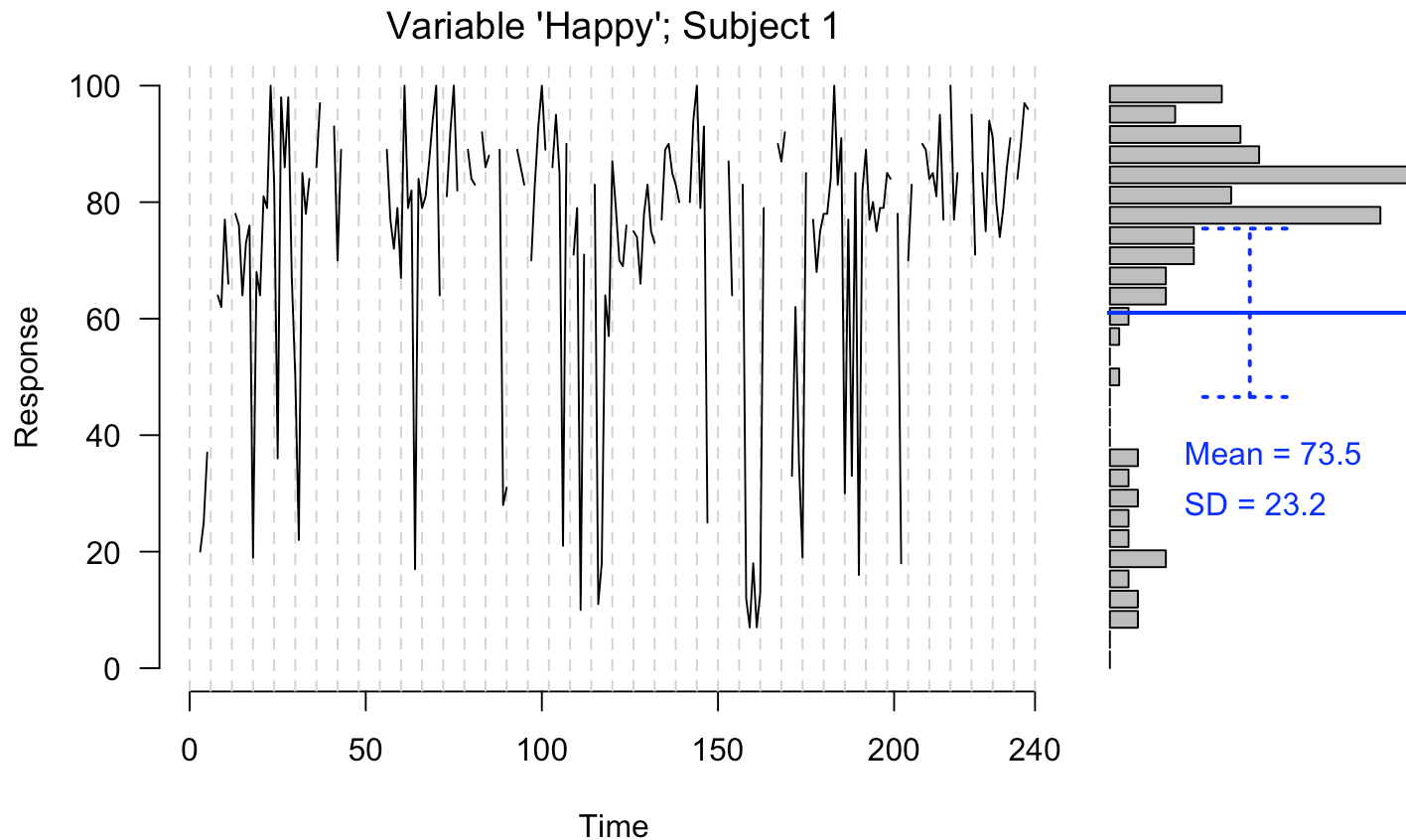
Looking at Data: Bivariate



Looking at Data: Bivariate

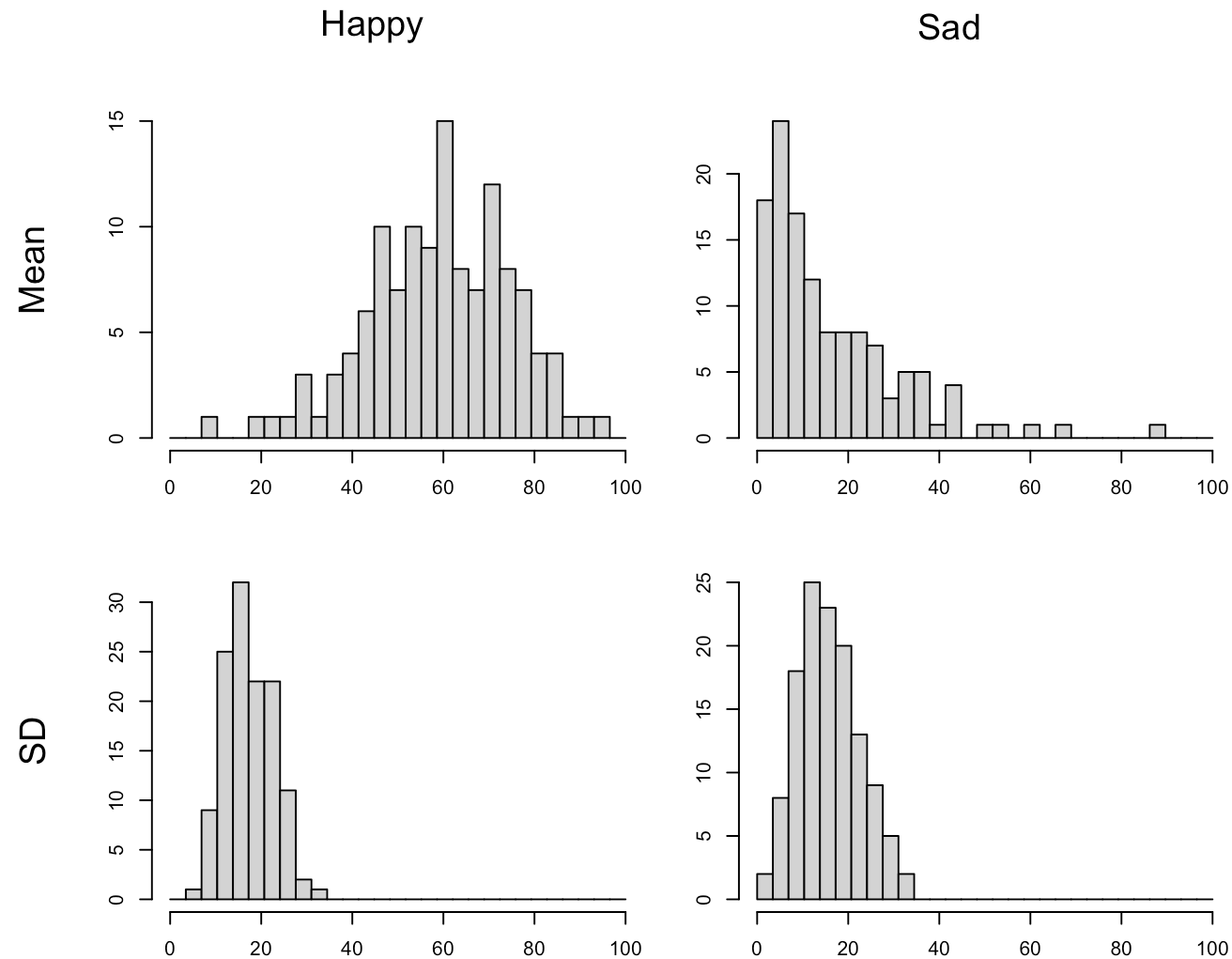
Heterogeneity: see [PDF](#)

Descriptive Statistics: Mean & SD



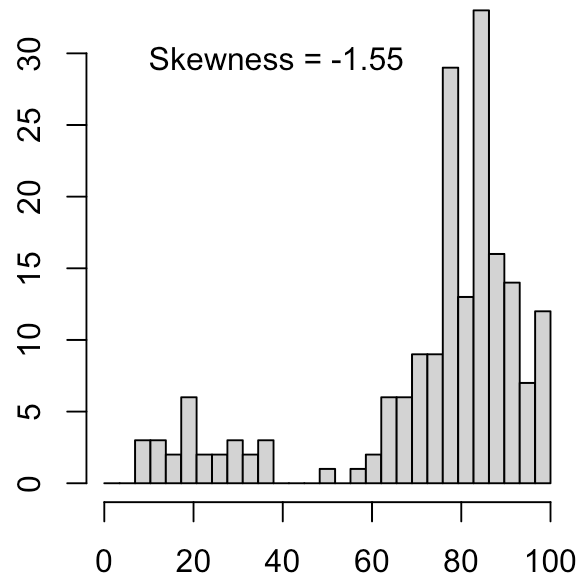
As always: means/SDs can be misleading if variables are not unimodal & symmetric.

Heterogeneity in Mean & SD

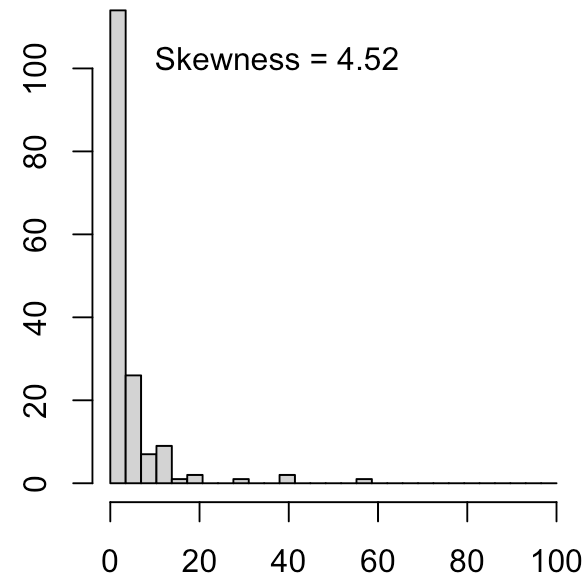


Descriptive Statistics: Skewness & Modality

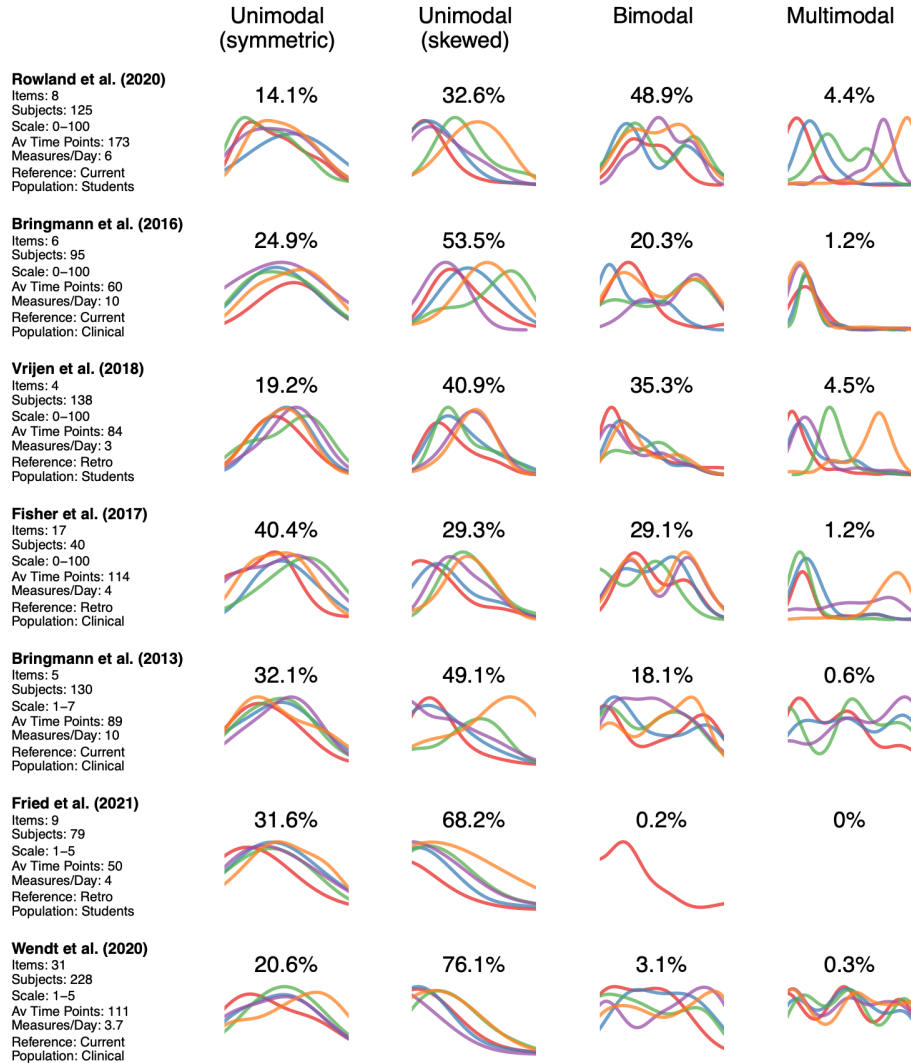
Happy (Subj 1)



Anxious (Subj 5)

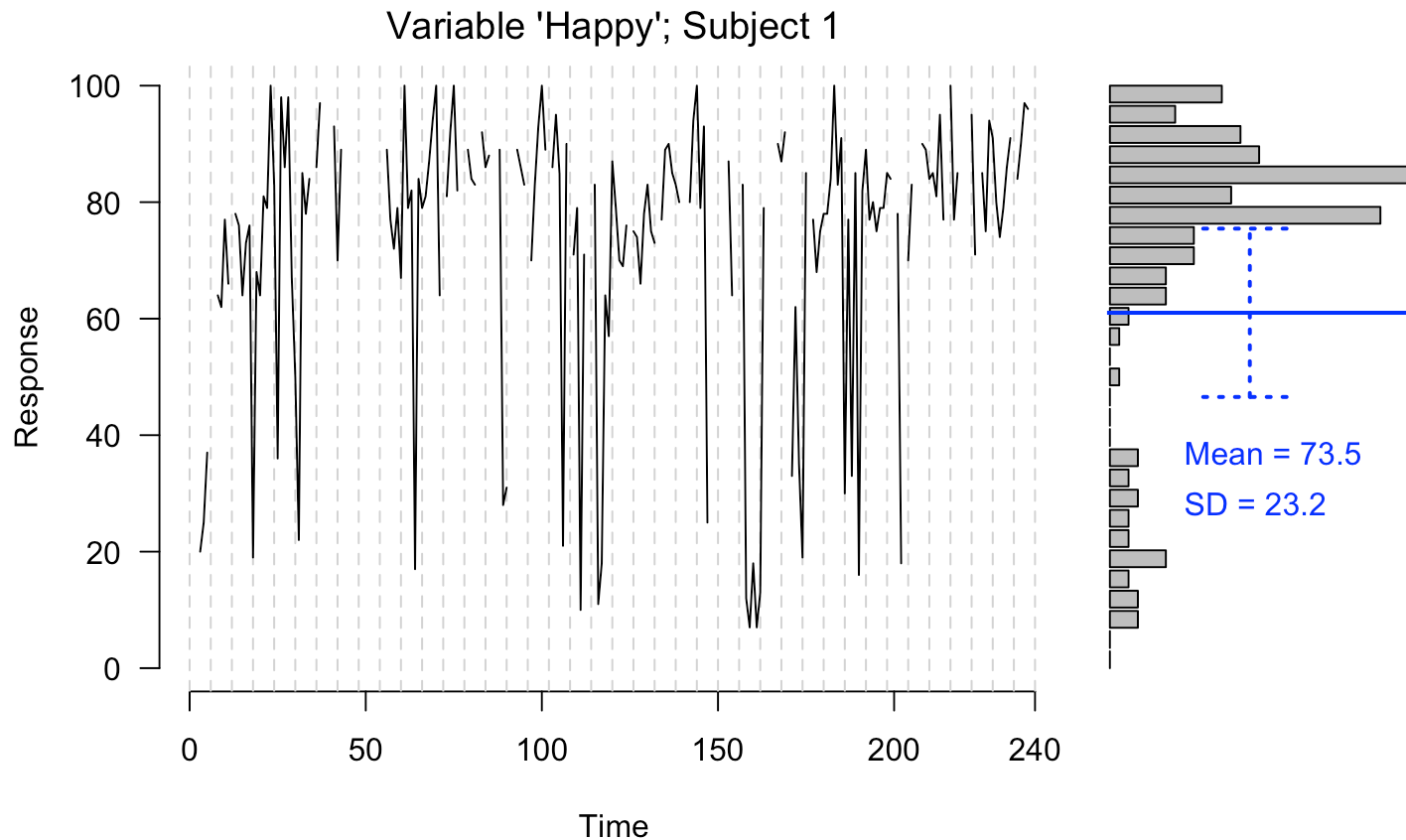


Paper on Skewness & Modality



Preprint: <https://psyarxiv.com/qudr6>

Descriptive Statistics: RMSSD & AR

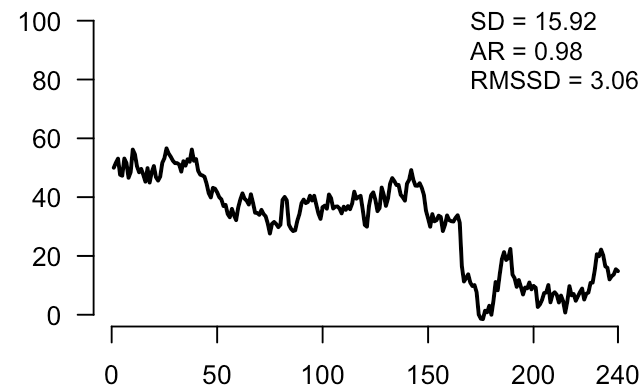
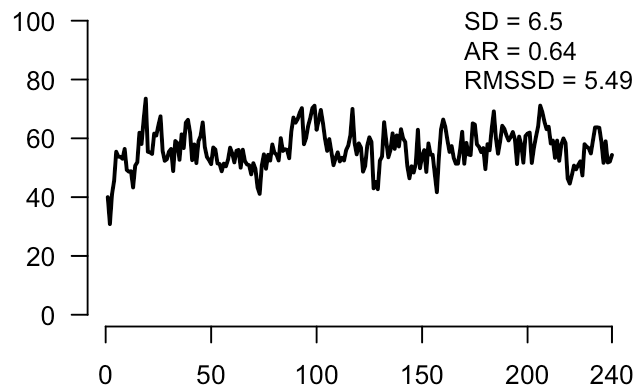
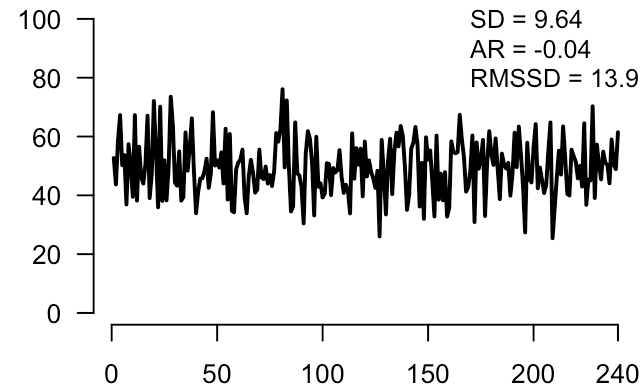
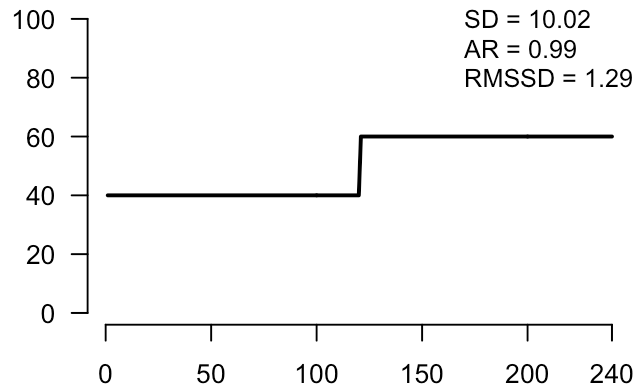


Variance captures only a small aspect of time evolution.

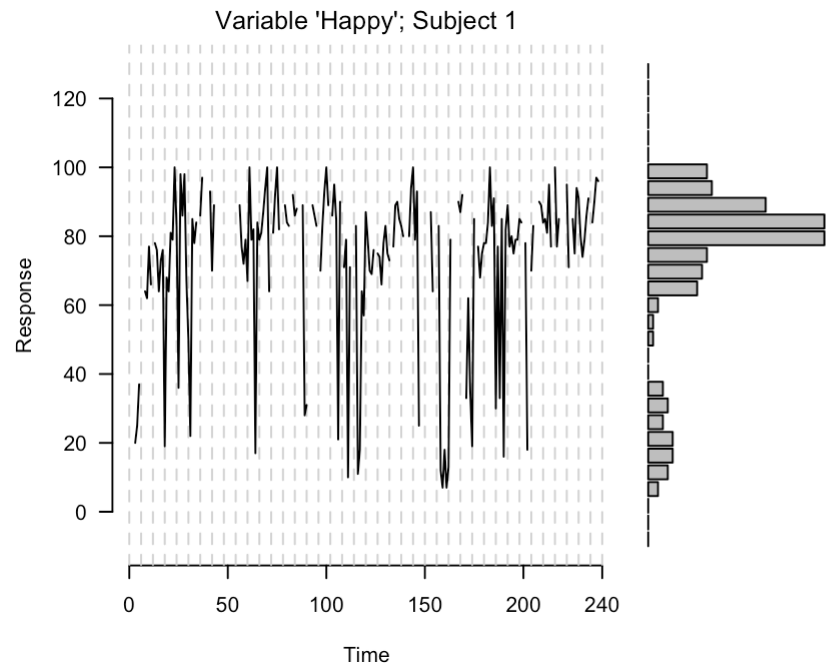
Descriptive Statistics: RMSSD & AR

$$\text{RMSSD} = \sqrt{\frac{1}{N} \sum_t^N (X_t - X_{t-1})^2}$$

$$\text{AR} = \text{cor}(X_t, X_{t-1})$$



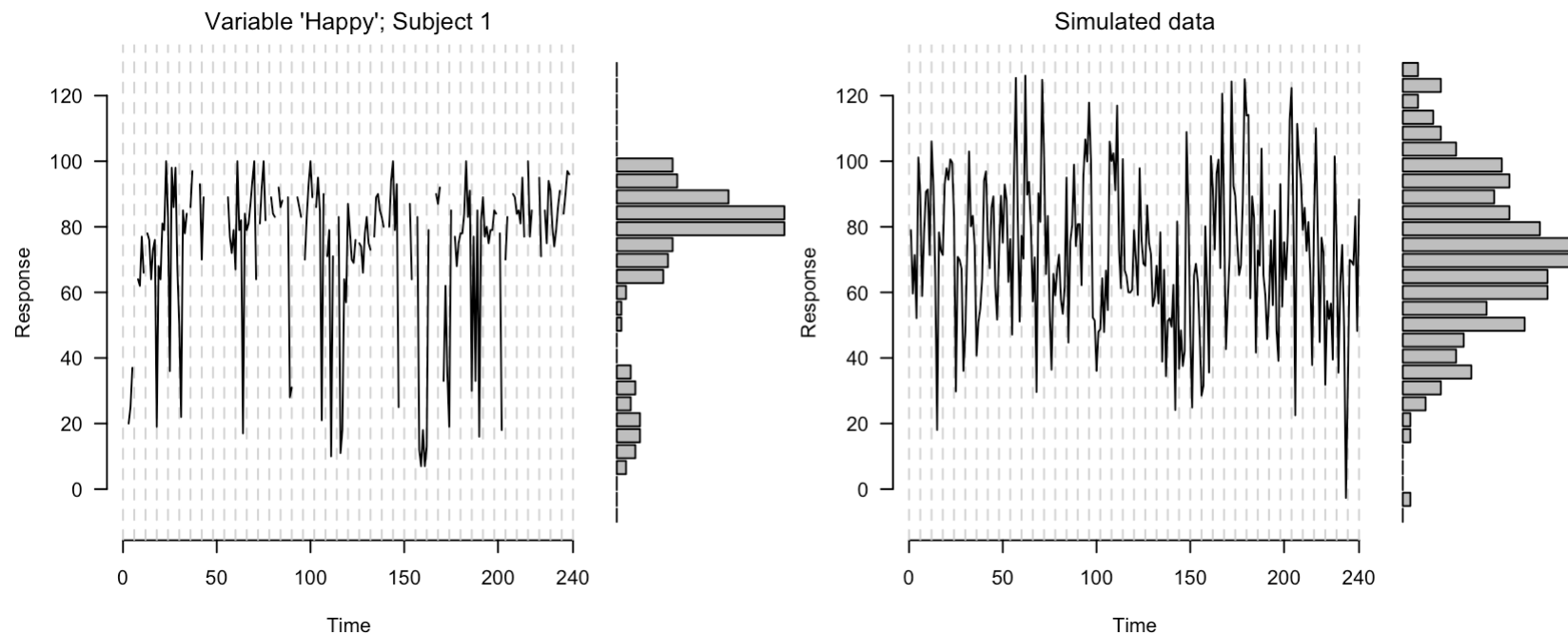
Absolute model fit of AR model



```
out <- lm(X[-1] ~ X[-length(X)])  
out$coefficients
```

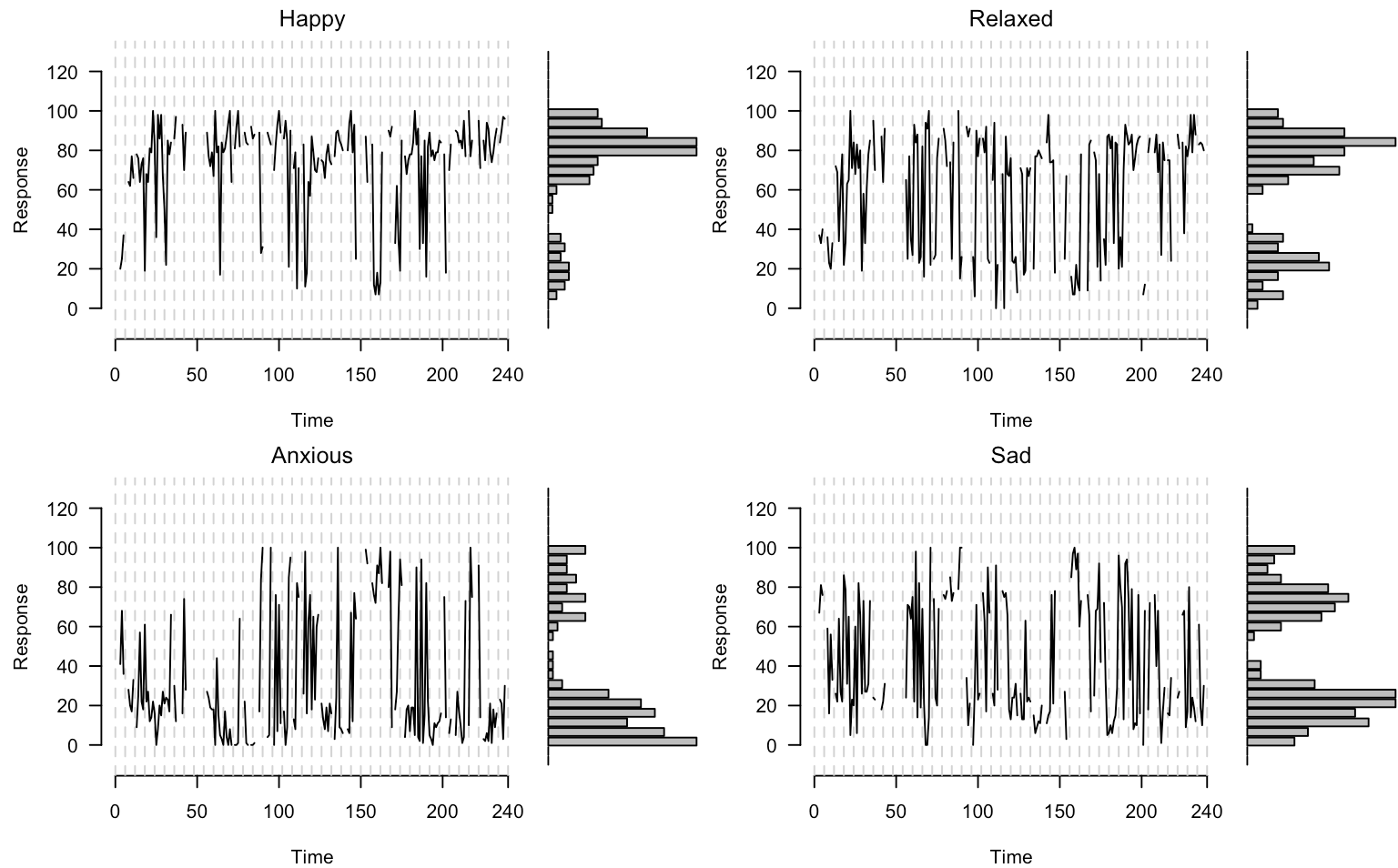
```
## (Intercept) X[-length(X)]  
## 46.0438002 0.3542108
```

Absolute model fit of AR model

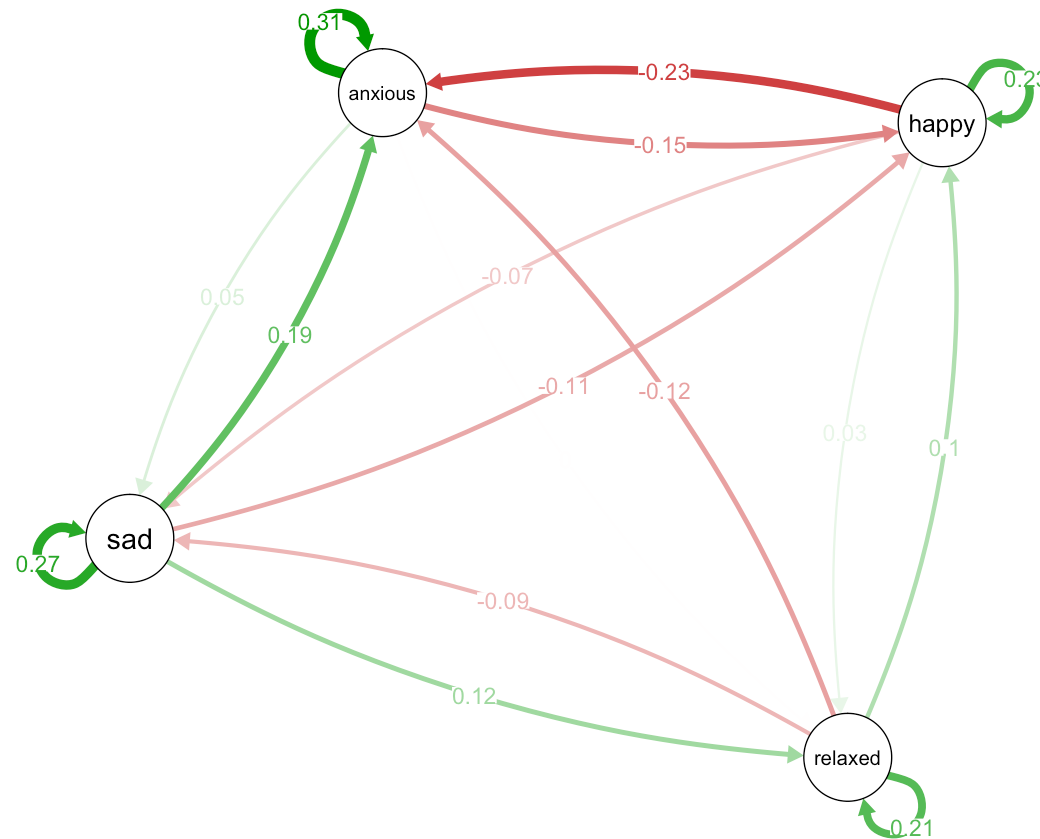


```
set.seed(1)
X_sim <- rep(NA, 240)
X_sim[1] <- X[1]
coefs <- out$coefficients
res_sd <- sd(residuals(out))
for(i in 2:240) X_sim[i] <- coefs[1] + coefs[2]*X_sim[i-1] + rnorm(1, 0, res_sd)
```

Absolute model fit of VAR model



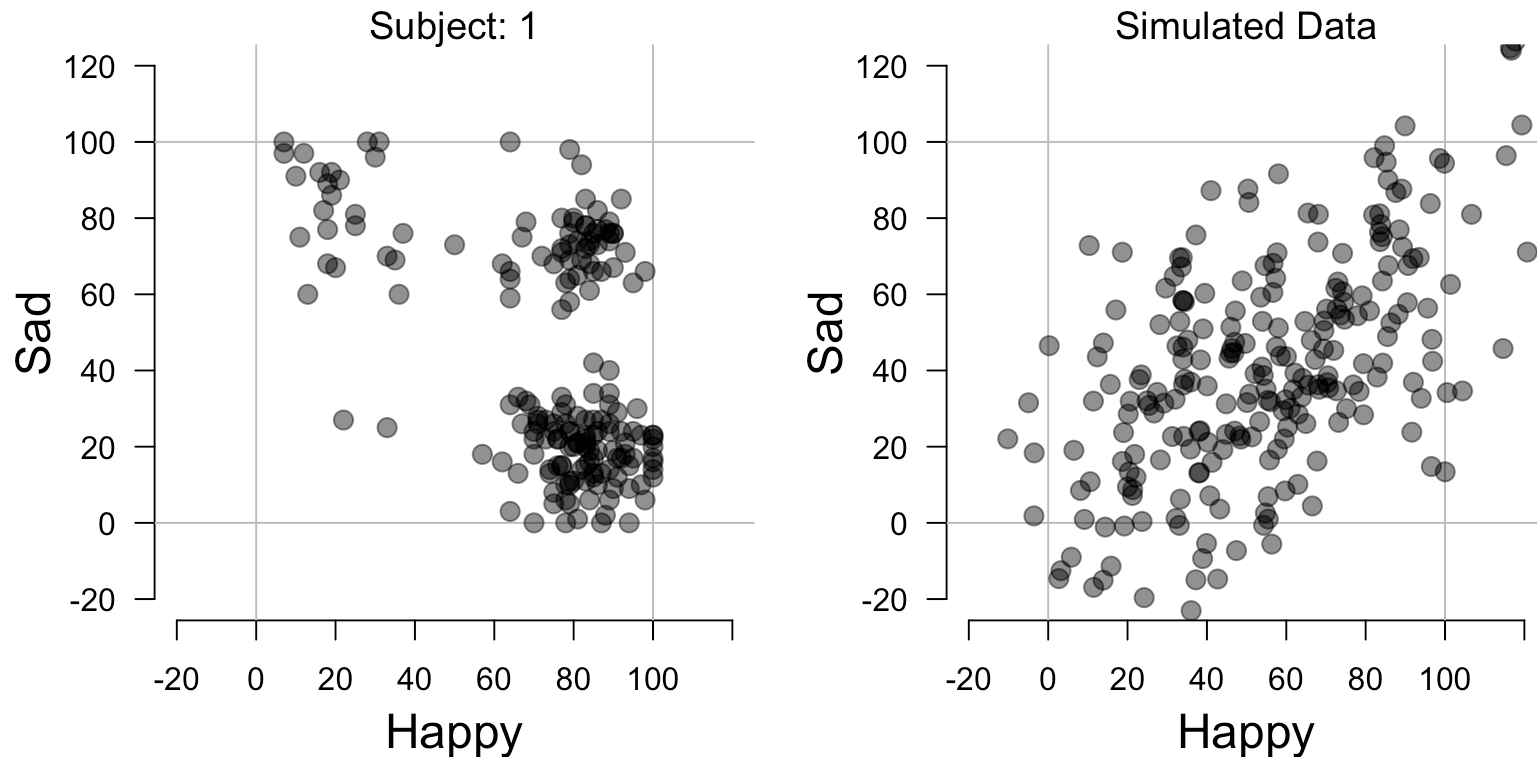
Absolute model fit of VAR model



```
round(ints, 2) # Intercepts
```

```
## [1] 61.81 46.45 33.22 27.58
```

Absolute model fit of VAR model



- Useful summary of the data (conditional autocorrelations and cross-correlations)
- But: Unlikely a (good approximation of a) generating mechanism

Descriptive vs. Generative Models

- How can we use statistical models if they are no plausible generating/mechanistic models?

An illustration of the issue:

Haslbeck, J. M., & Ryan, O. (2021). Recovering within-person dynamics from psychological time series. *Multivariate Behavioral Research*, 1-32. [LINK](#)

Some ideas on how to move forward:

Haslbeck, J., Ryan, O., Robinaugh, D. J., Waldorp, L. J., & Borsboom, D. (2021). Modeling psychopathology: From data models to formal theories. *Psychological Methods*. [LINK](#)

Summary

Missingness

- Many ways to analyze it
- Can be seen as data/behavior

Looking at Data

- Critical to do before modeling
- Many discoveries: Modality, skewness, low variance, huge heterogeneity, etc.

Descriptives

- Mean, SD, RMSSD, Skew, Modality
- Can provide powerful characterization of data

Basic Modeling

- AR, VAR
- Distinction: Descriptive & generative models

References

Rintala, Aki, Martien Wampers, Inez Myin-Germeys, and Wolfgang Viechtbauer. 2020. "Momentary Predictors of Compliance in Studies Using the Experience Sampling Method." *Psychiatry Research* 286: 112896.

Rowland, Zarah, Mario Wenzel, and Thomas Kubiak. 2020. "A Mind Full of Happiness: How Mindfulness Shapes Affect Dynamics in Daily Life." *Emotion* 20 (3): 436.

Silvia, Paul J, Thomas R Kwapil, Kari M Eddington, and Leslie H Brown. 2013. "Missed Beeps and Missing Data: Dispositional and Situational Predictors of Nonresponse in Experience Sampling Research." *Social Science Computer Review* 31 (4): 471–81.

Sun, Jessie, Mijke Rhemtulla, and Simine Vazire. 2021. "Eavesdropping on Missing Data: What Are University Students Doing When They Miss Experience Sampling Reports?" *Personality and Social Psychology Bulletin* 47 (11): 1535–49.