So you want to use some time series data

Emorie D. Beck

Wednesday May 1, 2021

NORTHWESTERN UNIVERSITY
FEINBERG
SCHOOL OF MEDICINE

	EARLY BIRD	REGULAR
STUDENT	$30	$45
Requires an Undergraduate or Graduate SPSP Membership valid through Decemb		
EARLY CAREER	$30	$45
Requires an Early Career SPSP Membership valid through December 2021 - check		
FULL/RETIRED/ ASSOCIATE	$105	$130
Requires a Full, Retired or Associate SPSP Membership valid through December 20		
NON-MEMBER	$150	$175

SPF

Summer Psychology Forum

AN SPSP CONFERENCE

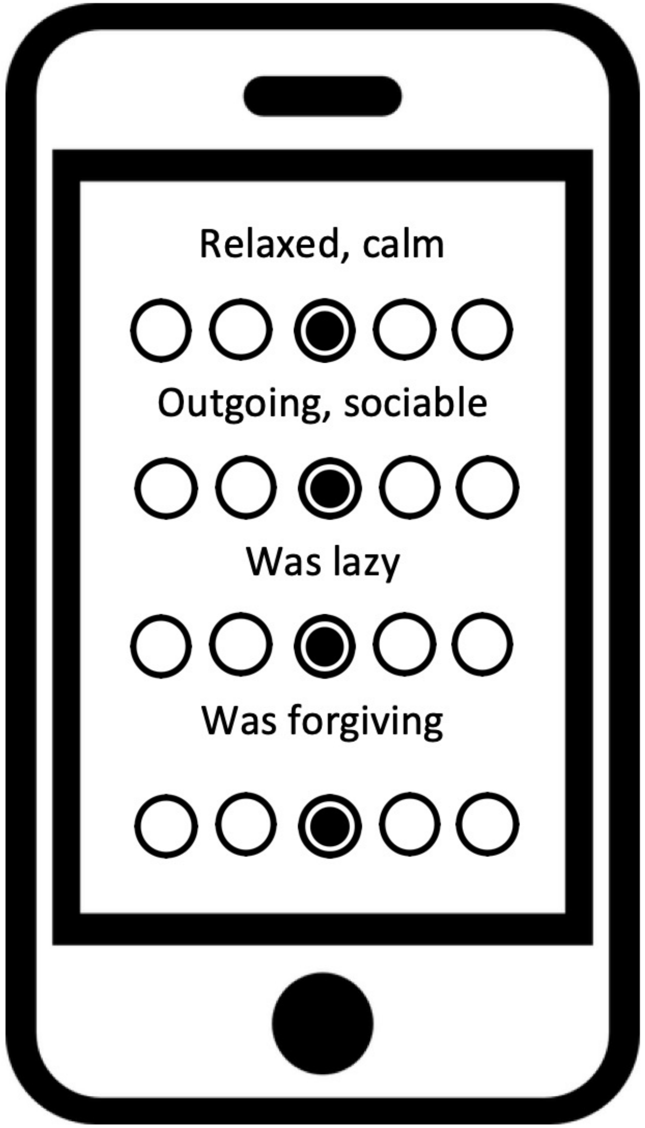
https://forum.spsp.org

Modeling and Analyzing the Dynamics of Motivation, Affect, and Social Interaction

Keynote	Robin Vallacher
Introduction to Complex Systems Theory	Brian Eiler
How to Set Up an EMA Study	Emorie Beck
Applied EMA Techniques	Ruben Arslan
Collecting Mobile Sensor Data	Alex Danvers
Translating Verbal Models to Theory	Eiko Fried and Don Robinaugh
Introduction to MLM for Longitudinal Data Analysis	Josh Jackson
VAR Models + GIMME	Aidan Wright
Dynamic Network Analysis	Julian Burger
Recurrence Quantification Analysis / Advanced Physiological Signal Processing	Aaron Likens
Neural Network Models for Building Theory	Stephen Read
Introduction to Agent-Based Modeling	Andrzej Nowak
Language as a Dynamic System: Where to Get it and What to Do Next	Rick Dale

What is a time series?

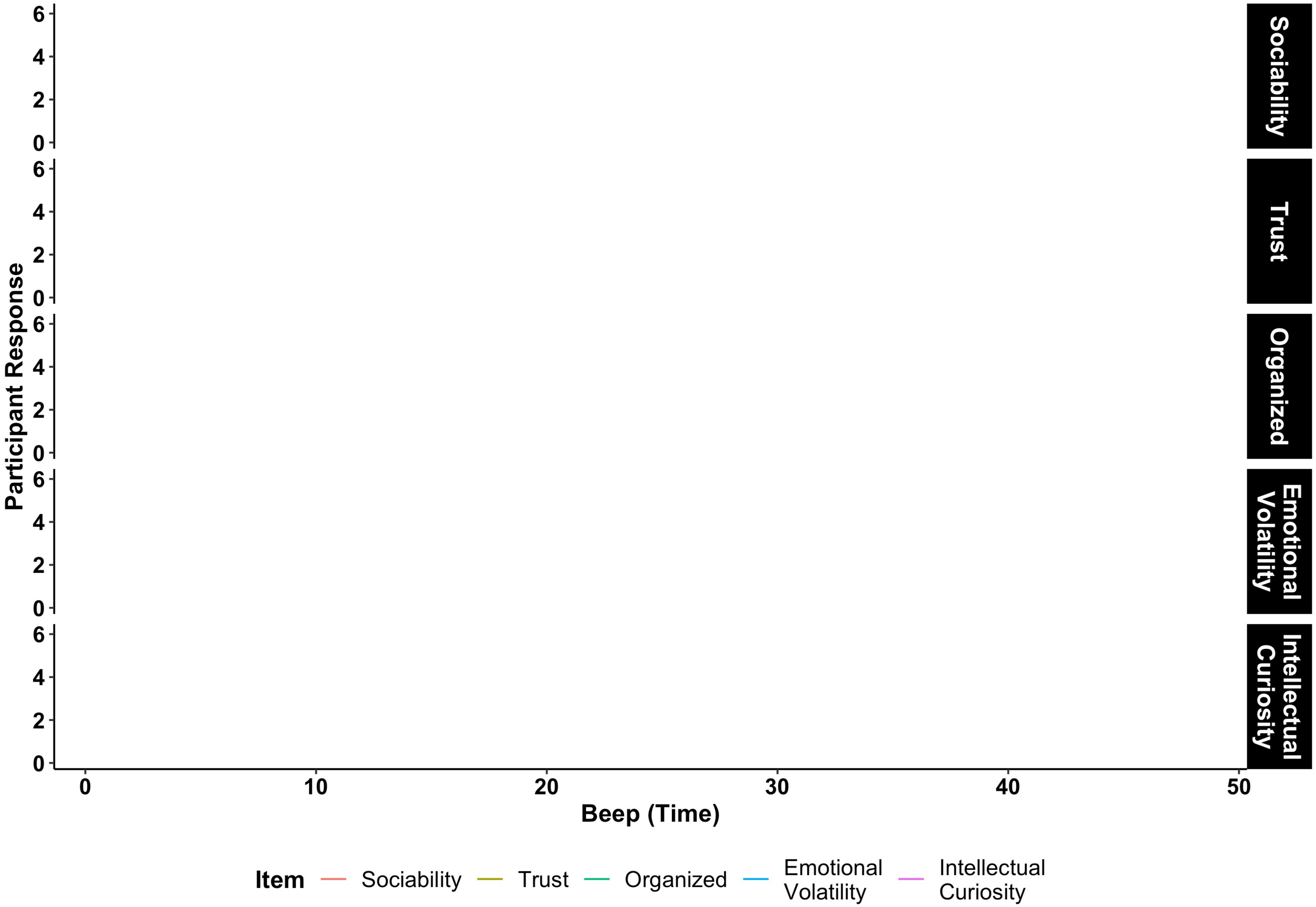
Brian Eiler’s “An introduction to complex systems”



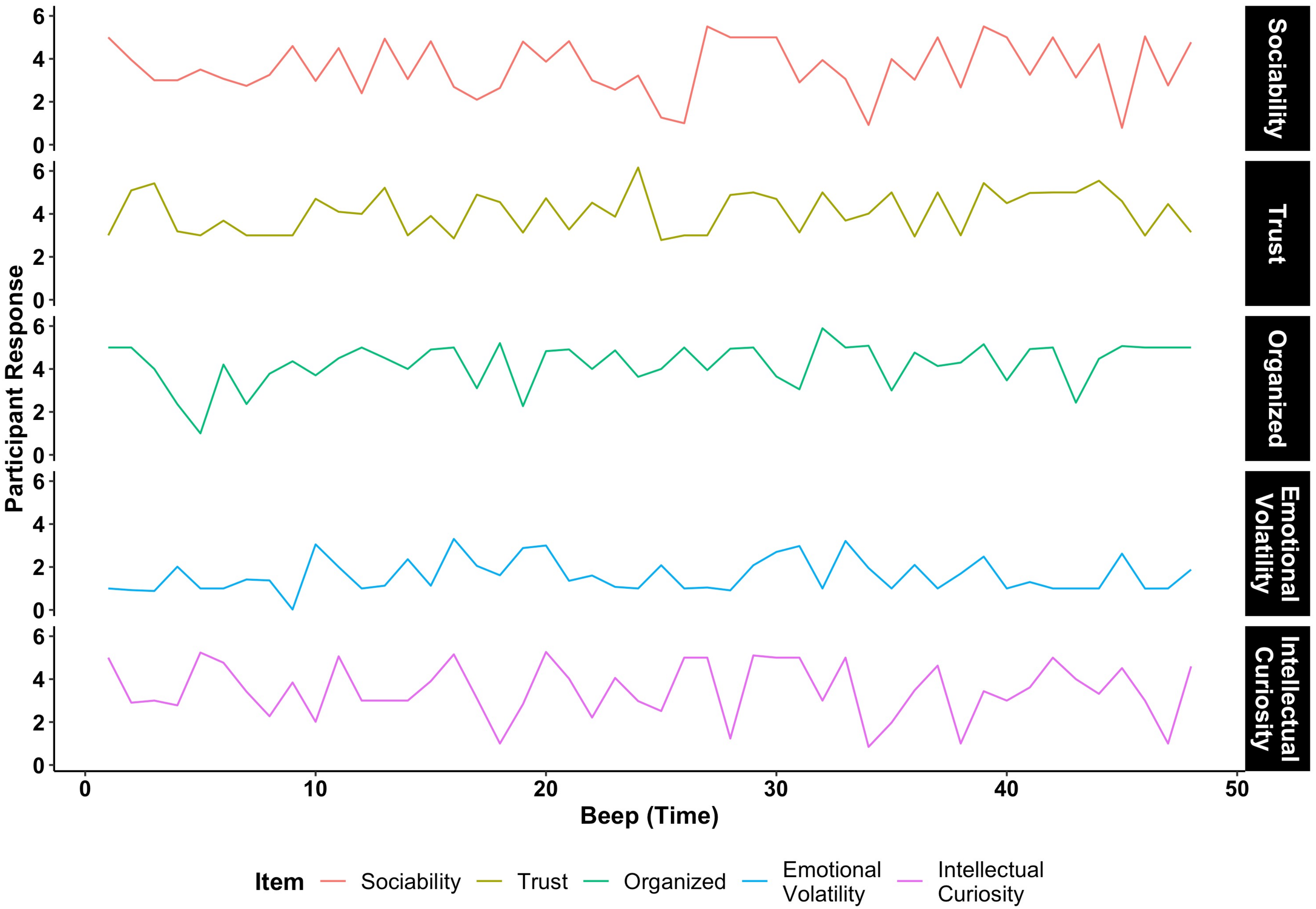
Emorie Beck's "How to set up an EMA study"; Ruben Arslan's "Applied EMA techniques"; Alex Danver's "Collecting Mobile Sensor Data"; Rick Dale's "Language as a Dynamic System"

ID	Date	Sociability	Trust	Organization	Emotional Volatility	Intellectual Curiosity
1	10/22/18 15:23	3	3	4	3	3
1	10/22/18 19:00					
1	10/22/18 23:23	2	2	1	4	2
1	10/22/18 23:25	3	1	1	3	4
1	10/23/18 3:00					
1	10/23/18 7:00					
1	10/23/18 11:00					
1	10/23/18 17:50	3	4	4	3	3
1	10/23/18 19:39	2	3	3	3	3
1	10/23/18 23:00					
1	10/24/18 0:00	5	4	1	4	1
1	10/24/18 3:00					
1	10/24/18 7:00					
1	10/24/18 11:00					
1	10/24/18 11:44	4	3	2	1	2
1	10/24/18 15:37	5	2	4	4	3

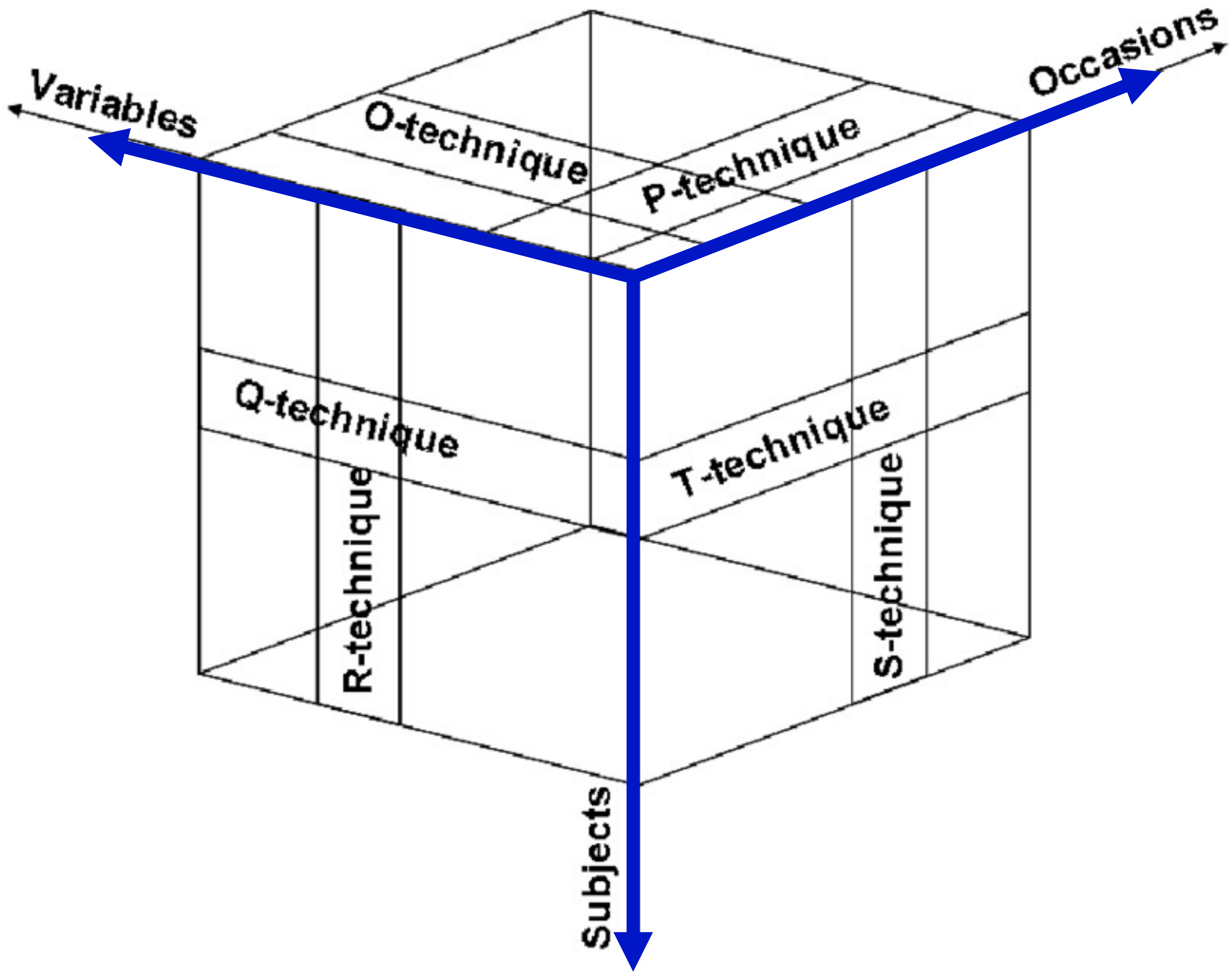
Emorie Beck's "How to set up an EMA study"; Ruben Arslan's "Applied EMA techniques"; Alex Danver's "Collecting Mobile Sensor Data"; Rick Dale's "Language as a Dynamic System"



Sociability
Trust
Organized
Emotional Volatility
Intellectual Curiosity
Participant Response
Beep (Time)
0
10
20
30
40
50
0
2
4
6
Item
Sociability
Trust
Organized
Emotional Volatility
Intellectual Curiosity



Participant Response
Beep (Time)
Sociability
Trust
Organized
Emotional Volatility
Intellectual Curiosity
0
2
4
6
0
10
20
30
40
50
Item
Sociability
Trust
Organized
Emotional Volatility
Intellectual Curiosity



See Boker & Martin (2018)

Questions	Aggregation
Description	Between-Person / Nomothetic
Prediction	Within-Person
Explanation	Idiographic

See Beck & Jackson (2021)

Part 1: Considerations

Basic / Data Driven

- Do you have sufficient N?
- Do you have a variance or frequency problem?
- The day-night problem
- What is your missingness structure?

More Challenging / Conceptual

- Is your question autoregressive?
- Are there likely individual differences in the phenomena?
- How prevalent are different response styles?

Part 1: Considerations

Do you have sufficient N?

Between **$N = 2$**

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1

$n_1 = 8$

$n_2 = 8$

Within

What Between and Within-Person N's do you need?

Depends on:

- Perspective (idiographic, within-person, etc.)
- Method (mean, SD, correlation)
- Number of indicators

But once you choose a question and look to a method, this should be the first question you ask yourself.

Part 1: Considerations

Do you have a variance or frequency problem?

S1:
SD_{y1} = 1.39
$Range_{y1}$ = 1 to 5

y_2: 0 = 4, 1 = 4

S2:
SD_{y1} = 1.39
Range = 1 to 2

y_2: 0 = 7, 1 = 1

ID	day	beep	y_1	y_2
S1	1	1	1	0
S1	1	2	3	1
S1	1	3	1	1
S1	1	4	2	0
S2	1	1	1	0
S2	1	2	1	0
S2	1	3	1	0
S2	1	4	1	1
S1	2	1	1	1
S1	2	2	3	0
S1	2	3	5	1
S1	2	4	2	0
S2	2	1	1	0
S2	2	2	1	0
S2	2	3	2	0
S2	2	4	1	0

"The method required it, but the life did not" (Allport, 1968).

What to do with low variances and frequencies:

Typical suggestions:
- standardize (within-person)
- throw out low frequency / variance folks

But the best thing you can do is design and test questions and provide instructions to participants that will provide variability in responses.

Part 1: Considerations

The Day-Night Problem

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1

Nighttime

Part 1: Considerations

The Day-Night Problem

What to do with overnight intervals

Typical suggestions:

- Add empty rows
- Cubic spline interpolation
- Linear normalization
- Ignore them (if not asking autoregressive questions)

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S1	1	5	NA
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S2	1	5	NA
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S1	2	5	NA
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1

Nighttime

Part 1: Considerations

What is your missingness structure?

What to do with missingness

Options

- **Observation-wise deletion: delete missing observations**
- **Person-wise deletion: delete people with reporting issues**

Regardless of how you deal with it, sensitivity analyses with and without those folks (or controlling for those patterns) should be conducted

ID	Day	Beep	Sociability	Trust	Organization	Emotional Volatility	Intellectual Curiosity
		1	3	3	4	3	3
		2	2	2	1	4	2
		3	3	1	1	3	4
		4	3	4	4	3	3
		1	2	3	3	NA	3
		2	5	4	1	NA	1
		3	4	3	2	NA	2
		4	5	2	4	NA	3
		1	NA	NA	NA	NA	NA
		2	4	2	3	1	3
		3	4	3	2	4	1
		4	3	2	3	2	2
		1	4	4	1	4	4
		2	2	2	2	4	3
4	1	3	2	1	1	2	2
4	1	4	4	3	0	4	4

Part 1: Considerations

Is your question autoregressive? The unequal interval problem

ID	day	beep	y_1	$y_{1,\,t-1}$
S1	1	1	1	
S1	1	2	3	
S1	1	3	1	
S1	1	4	2	
S1	1	5	NA	
S1	2	1	1	
S1	2	2	3	
S1	2	3	1	
S1	2	4	2	
S1	2	5	NA	
S2	1	1	2	
S2	1	2	2	
S2	1	3	4	
S2	1	4	3	
S2	1	5	NA	
S2	2	1	1	
S2	2	2	2	
S2	2	3	2	
S2	2	4	1	

Lagged Indicator

Part 1: Considerations

Is your question autoregressive? The unequal interval problem

ID	day	beep	y_1	$y_{1,\,t\text{-}1}$
S1	1	1	1	NA
S1	1	2	3	1
S1	1	3	1	3
S1	1	4	2	1
S1	1	5	NA	2
S1	2	1	1	NA
S1	2	2	3	1
S1	2	3	1	3
S1	2	4	2	1
S1	2	5	NA	2
S2	1	1	2	NA
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
S2	1	5	NA	3
S2	2	1	1	NA
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2

Lagged Indicator

Part 1: Considerations

Are there likely individual differences in the phenomena?

ID	day	beep	y_1	$y_{1, t-1}$
S1	1	1	1	
S1	1	2	3	
S1	NA	NA	NA	
S1	1	4	2	
S1	1	5	NA	
S1	2	1	1	
S1	NA	NA	NA	
S1	2	3	1	
S1	2	4	2	
S1	2	5	NA	
S2	1	1	2	
S2	1	2	2	
S2	1	3	4	
S2	1	4	3	
S2	1	5	NA	
S2	2	1	1	
S2	2	2	2	
S2	2	3	2	
S2	2	4	1	

Part 1: Considerations

Are there likely individual differences in the phenomena?

ID	day	beep	y_1	$y_{1,\ t-1}$
S1	1	1	1	NA
S1	1	2	3	1
S1	1	4	2	3
S1	1	5	NA	2
S1	2	1	1	NA
S1	2	3	1	1
S1	2	4	2	1
S1	2	5	NA	2
S2	1	1	2	NA
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
S2	1	5	NA	3
S2	2	1	1	NA
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2
S2	2	5	NA	1

Part 1: Considerations

Are there likely individual differences in the phenomena?

ID	day	beep	y_1	$y_{1,\,t-1}$
S1	1	1	1	NA
S1	1	2	3	1
S1	1	4	2	3
S1	1	5	NA	2
S1	2	1	1	NA
S1	2	3	1	1
S1	2	4	2	1
S1	2	5	NA	2
S2	1	1	2	NA
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
S2	1	5	NA	3
S2	2	1	1	NA
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2
S2	2	5	NA	1

Part 1: Considerations

Are there likely individual differences in the phenomena?

ID	day	beep	y_1	$y_{1,t-1}$
S1	1	1	1	NA
S1	1	2	3	1
S1	NA	NA	NA	3
S1	1	4	2	NA
S1	1	5	NA	2
S1	2	1	1	NA
S1	NA	NA	NA	1
S1	2	3	1	NA
S1	2	4	2	1
S1	2	5	NA	2
S2	1	1	2	NA
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
S2	1	5	NA	3
S2	2	1	1	NA
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2

Part 1: Considerations

Are there likely individual differences in the phenomena?

What to do with unequal intervals:

Options:
- Don't lag (best) and use alternative model
- Cubic spline interpolation
- Linear normalization

Once data are collected, they are what they are. Better to design a study that minimizes unequal lags if basic lagged (often AR(1)) questions are your goal

ID	day	beep	y_1	$y_{1,\,t\text{-}1}$
~~S1~~	~~1~~	~~1~~	~~1~~	~~NA~~
S1	1	2	3	1
~~S1~~	~~NA~~	~~NA~~	~~NA~~	~~3~~
~~S1~~	~~1~~	~~4~~	~~2~~	~~NA~~
~~S1~~	~~1~~	~~5~~	~~NA~~	~~2~~
~~S1~~	~~2~~	~~1~~	~~1~~	~~NA~~
~~S1~~	~~NA~~	~~NA~~	~~NA~~	~~1~~
~~S1~~	~~2~~	~~3~~	~~1~~	~~NA~~
S1	2	4	2	1
~~S1~~	~~2~~	~~5~~	~~NA~~	~~2~~
~~S2~~	~~1~~	~~1~~	~~2~~	~~NA~~
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
~~S2~~	~~1~~	~~5~~	~~NA~~	~~3~~
~~S2~~	~~2~~	~~1~~	~~1~~	~~NA~~
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2

Part 1: Considerations

How prevalent are different response styles?

Are restrictions in variance reflective of participants' actual experiences or of response styles?

What to do with response styles:

Options:

- Best to tackle this at the collection stage

Once data are collected, they are what they are. Good instructions on how you *want* participants to interpret and use scales can alleviate this issue. Different response styles may be interesting, but they will hamper inference.

ID	day	beep	y_1	$y_{1,\,t-1}$
S1	1	1	4	
S1	1	2	3	
S1	1	3	4	
S1	1	4	2	
S1	1	5	NA	
S1	2	1	4	
S1	2	2	3	
S1	2	3	4	
S1	2	4	2	
S1	2	5	NA	
S2	1	1	1	
S2	1	2	2	
S2	1	3	4	
S2	1	4	3	
S2	1	5	NA	
S2	2	1	1	
S2	2	2	5	
S2	2	3	2	
S2	2	4	4	
S2	2	5	NA	

Part 2: Statistical Methods

Basic / Univariate

Univariate: Variability, Instability, Inertia

Bi/Multivariate: Correlations and (Basic) Structural Models

Variance Decomposition: Multilevel Models

Advanced / Multivariate

Cross-Lagged VAR, GIMME, and EGA

Machine Learning

Differential Equation Models

Part 2: Statistical Methods

Univariate: Variability, Instability, Inertia

ID	day	beep	y_1	$y_{1,\,t-1}$
S1	1	1	1	NA
S1	1	2	3	1
S1	1	3	1	3
S1	1	4	2	1
S1	1	5	NA	2
S1	2	1	1	NA
S1	2	2	3	1
S1	2	3	1	3
S1	2	4	2	1
S1	2	5	NA	2
S2	1	1	2	NA
S2	1	2	2	2
S2	1	3	4	2
S2	1	4	3	4
S2	1	5	NA	3
S2	2	1	1	NA
S2	2	2	2	1
S2	2	3	2	2
S2	2	4	1	2
S2	2	5	NA	1

Intraindividual Variability: the standard deviation of an indicator

$$SD_{y,i} = \sqrt{\frac{\sum(y_t - \bar{y})^2}{N-1}}$$

R: `sd()`, `psych::SD()`

Instability: mean squared successive difference (MSSD)

$$MSSD_{y,i} = \frac{\sum(y_t - y_{t-1})^2}{N}$$

R: `psych::mssd()`

Inertia: the univariate, lagged autocorrelation

$$r_{yy_{t-1},i} = \frac{\sum(y_t - \bar{y})(y_{t-1} - \overline{y_{t-1}})}{\sqrt{\sum(y_t - \bar{y})^2(y_{t-1} - \overline{y_{t-1}})^2}}$$

R: `cor()`, `psych::cor.ci()`

See Ong & Ram (2017)

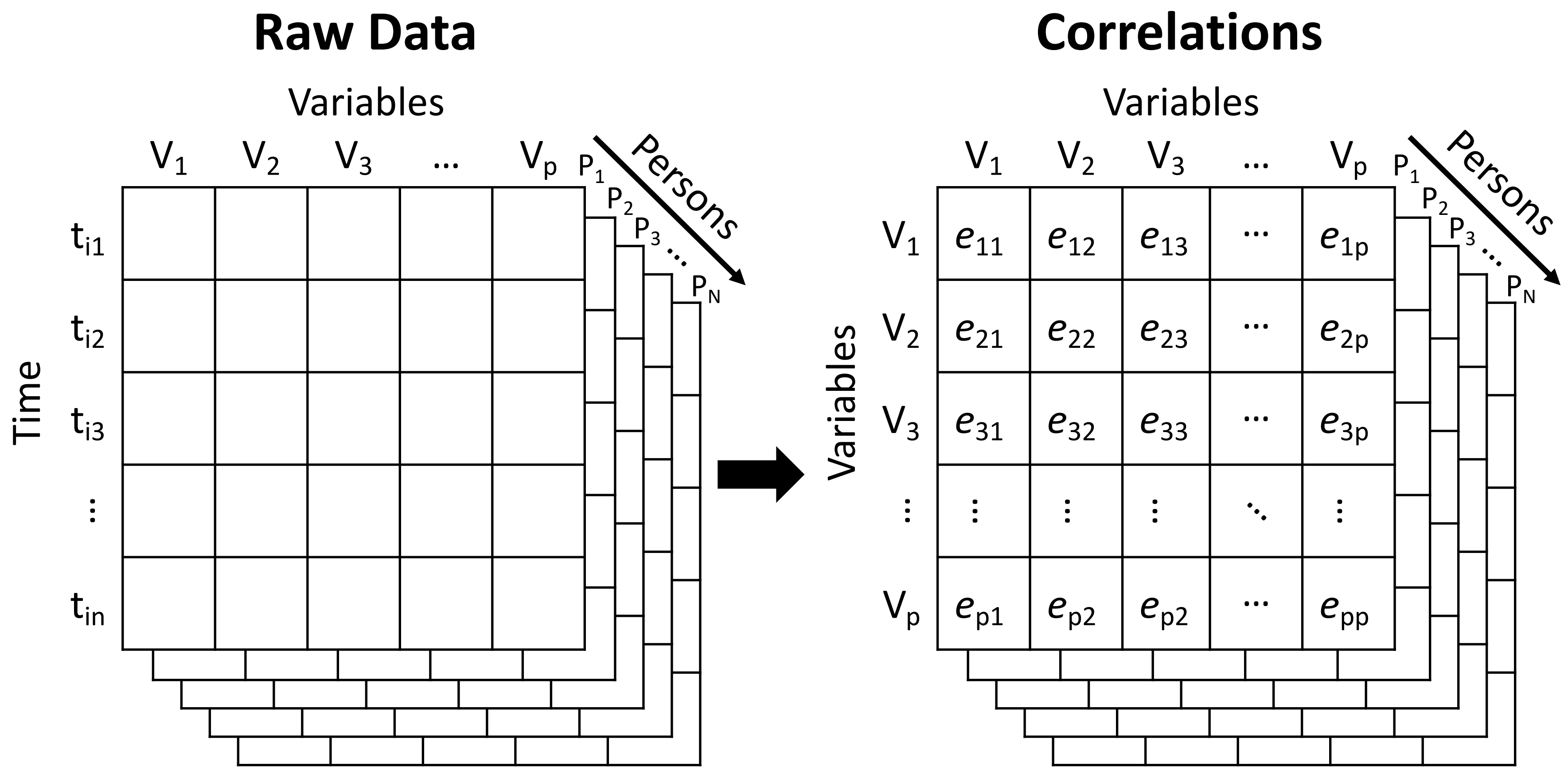
Part 2: Statistical Methods

Univariate: Variability, Instability, Inertia

ID	SD_y	$MSSD_y$	$r_{y,yt-1}$
S1	1.07	0.94	0.12
S2	0.88	1.28	-0.35
S3	0.91	1.39	-0.25
S4	1.55	2.98	0.12
S5	0.93	1.02	-0.08
S6	0.99	0.52	0.03
S7	0.81	1.05	-0.11
S8	1.05	1.17	0.18
S9	1.03	0.91	0.38
S10	1.06	1.01	0.29

Part 2: Statistical Methods

Bi/Multivariate: Correlations and (Basic) Structural Models



Part 2: Statistical Methods

Bi/Multivariate: Correlations and (Basic) Structural Models

	V1	V2	V3	V4	V5
V1					
V2	0.02				
V3	-0.05	0.08			
V4	0.64	-0.12	-0.10		
V5	0.26	-0.27	0.43	0.07	

Factor Analysis
Dynamic Factor Analysis
`psych::fa.parallel(),psych::fa()`

Principle Components Analysis
`psych::fa.parallel(),`
`psych::pca()`

Other clustering techniques

See Beck & Jackson, 2020

Part 2: Statistical Methods

Variance Decomposition: Multilevel Models

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S1	1	5	NA
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S2	1	5	NA
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S1	2	5	NA
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1
S2	2	5	NA

Observations

Level 1: $Y_{bi} = \beta_{0i} + \varepsilon_{bi}$

, where *b* = beep, and *i* = person

β_{0di} = average level for person *i* ε_{bdi}= residual for beep *b*, person *i*

Person

Level 2: $\beta_{0i} = \mu_{00} + r_{0i}$

μ_{000} = average level across persons

r_{00i} = deviation for person *i* across beeps

var(r_{00i}) = between-person variance in within-person levels

```
lme4::lmer(y ~ 1 + (1 | ID), data = d)
```

See Haan-Rietdijk et al., 2016

Josh Jackson's "Multilevel Model (for EMA data)"

Part 2: Statistical Methods

Variance Decomposition: Multilevel Models

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S1	1	5	NA
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S2	1	5	NA
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S1	2	5	NA
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1
S2	2	5	NA

Observations

Level 1: $\boldsymbol{Y_{bdi} = \beta_{0di} + \varepsilon_{bdi}}$

, where *b* = beep, *d* = day, and *i* = person

β_{0di} = average level for person *i* on day *d*

ε_{bdi}= residual for beep *b*, day *d*, person *i*

Day

Level 2: $\boldsymbol{\beta_{0di} = \gamma_{00i} + u_{0di}}$

γ_{00i} = average level for person *i* across days

$\boldsymbol{u_{0di}}$ = deviation for person *i* on day *d* across beeps *b*

Person

Level 3: $\boldsymbol{\gamma_{00i} = \mu_{000} + r_{00i}}$

μ_{000} = average level across person & days

r_{00i} = deviation for person *i* across day & beep

var($\boldsymbol{r_{00i}}$) = between-person variance in within-person levels

var($\boldsymbol{u_{0di}}$) = variance across days

```
lme4::lmer(y ~ 1 + (1 | ID/day), data = d)
```

See Haan-Rietdijk et al., 2016

Josh Jackson's "Multilevel Model (for EMA data)"

Part 2: Statistical Methods

Variance Decomposition: Multilevel Models

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S1	1	5	NA
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S2	1	5	NA
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S1	2	5	NA
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1
S2	2	5	NA

Observations

Level 1:

$$Y_{bi} = \beta_{0i} + \beta_{1i} * (X_{bi} - \bar{X}_i) + \varepsilon_{bi}$$

Person

Level 2:

$$\beta_{0i} = \mu_{00} + \mu_{01} * (\bar{X}_i - \bar{X}) + r_{0i}$$
$$\beta_{1i} = \mu_{10} + \mu_{11} * (\bar{X}_i - \bar{X}) + r_{1i}$$

μ_{10} = average change in Y as a function of deviations from within-person averages of X across people *i* and beeps *b*.

μ_{01}= average change in Y as a function of between person differences in average levels of X across people *i* and beeps *b*.

μ_{11}= average change in Y as a function of both within-person deviations from person-level means and between-person differences average levels of X across people *i* and beeps *b*.

See Gordon (2018) for a tutorial PDF

Josh Jackson's "Multilevel Model (for EMA data)"

Part 2: Statistical Methods

Variance Decomposition: Multilevel Models

ID	day	beep	y_1
S1	1	1	1
S1	1	2	3
S1	1	3	1
S1	1	4	2
S1	1	5	NA
S2	1	1	2
S2	1	2	2
S2	1	3	4
S2	1	4	3
S2	1	5	NA
S1	2	1	1
S1	2	2	3
S1	2	3	1
S1	2	4	2
S1	2	5	NA
S2	2	1	1
S2	2	2	2
S2	2	3	2
S2	2	4	1
S2	2	5	NA

Observations

Level 1:

$$Y_{bi} = \beta_{0i} + \beta_{1i} * (X_{bi} - \bar{X}_i) + \varepsilon_{bi}$$

Person

Level 2:

$$\beta_{0i} = \mu_{00} + \mu_{01} * (\bar{X}_i - \bar{X}) + r_{0i}$$
$$\beta_{1i} = \mu_{10} + \mu_{11} * (\bar{X}_i - \bar{X}) + r_{0i}$$

β_{0i} = average levels of Y for person *i* across beeps *b*.
β_{1i} = average change in Y as a function of within-person deviations in X for person *i* across beeps *b*.

r_{0i} = deviations from average levels of *Y* for person *i* across beeps *b*
r_{0i} =deviations in change in Y as a function of deviations in within-person levels of X for person *i* across beeps *b*.

See Gordon (2018) for a tutorial PDF

Josh Jackson's "Multilevel Model (for EMA data)"

Part 2: Statistical Methods

Cross-Lagged VAR, GIMME, and EGA

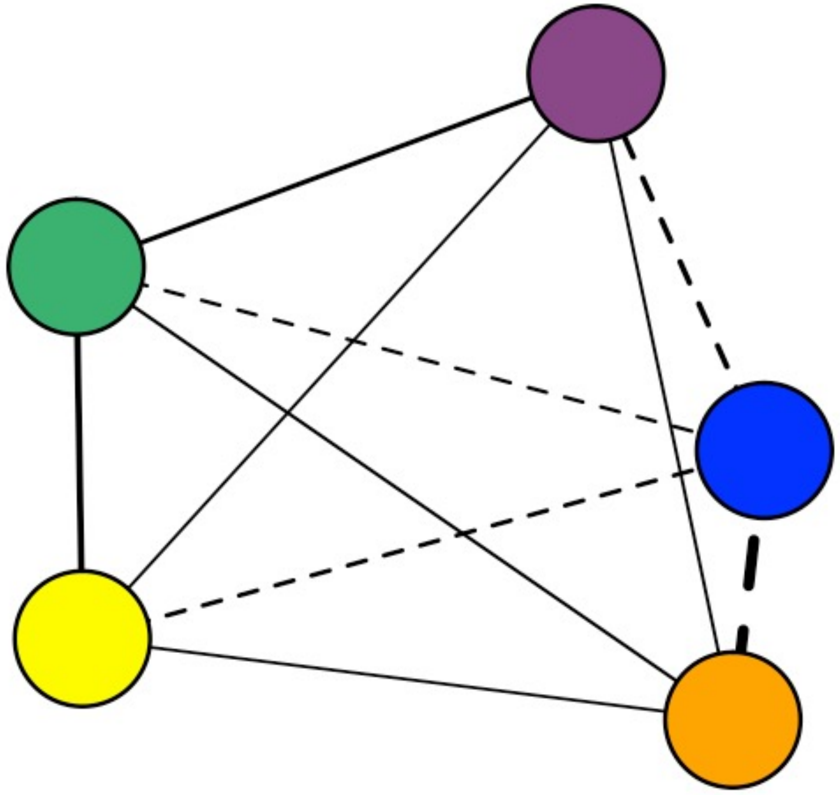
ID	Date	V1	V2	V3	V4	V5
1	10/22/18 15:23	3	3	4	3	3
1	10/22/18 23:23	2	2	1	4	2
1	10/22/18 23:25	3	1	1	3	4
1	10/23/18 17:50	3	4	4	3	3
1	10/23/18 19:39	2	3	3	3	3
1	10/24/18 0:00	5	4	1	4	1
1	10/24/18 11:44	4	3	2	1	2
1	10/24/18 15:37	5	2	4	4	3
1	10/24/18 20:46	2	3	3	2	3
1	10/25/18 21:07	4	2	3	1	3
1	10/25/18 22:54	4	3	2	4	0
1	10/25/18 23:47	3	2	3	2	2
1	10/26/18 8:44	4	4	1	4	4
1	10/27/18 16:46	2	2	2	4	3
1	10/27/18 22:55	2	1	1	2	2
1	10/28/18 3:40	4	3	0	4	4

Julian Burger’s “Dynamic Network Analysis”; Aidan Wright’s “GIMME”

Part 2: Statistical Methods

Cross-Lagged VAR, GIMME, and EGA

	V1	V2	V3	V4	V5
V1					
V2	0.02				
V3	-0.05	0.08			
V4	0.64	-0.12	-0.10		
V5	0.26	-0.27	0.43	0.07	

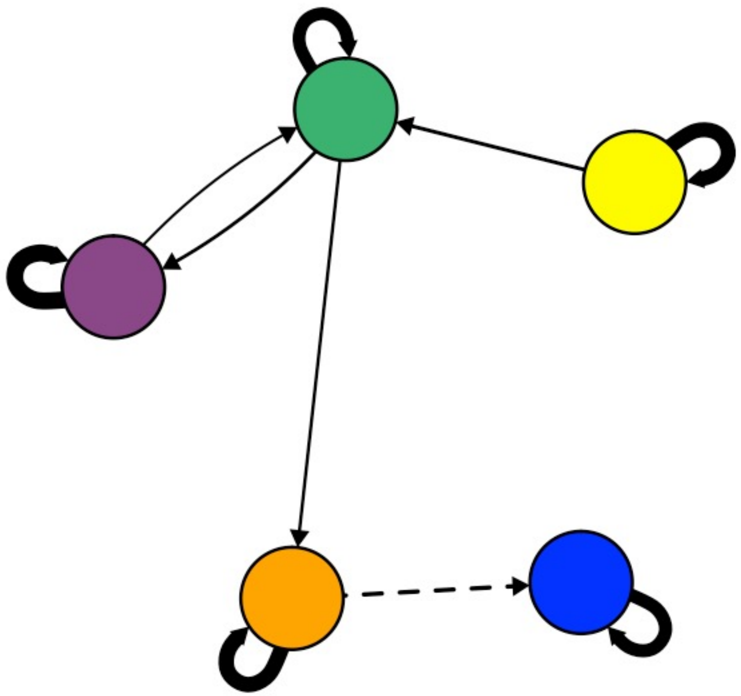


Julian Burger's "Dynamic Network Analysis"; Aidan Wright's "GIMME"

Part 2: Statistical Methods

Cross-Lagged VAR, GIMME, and EGA

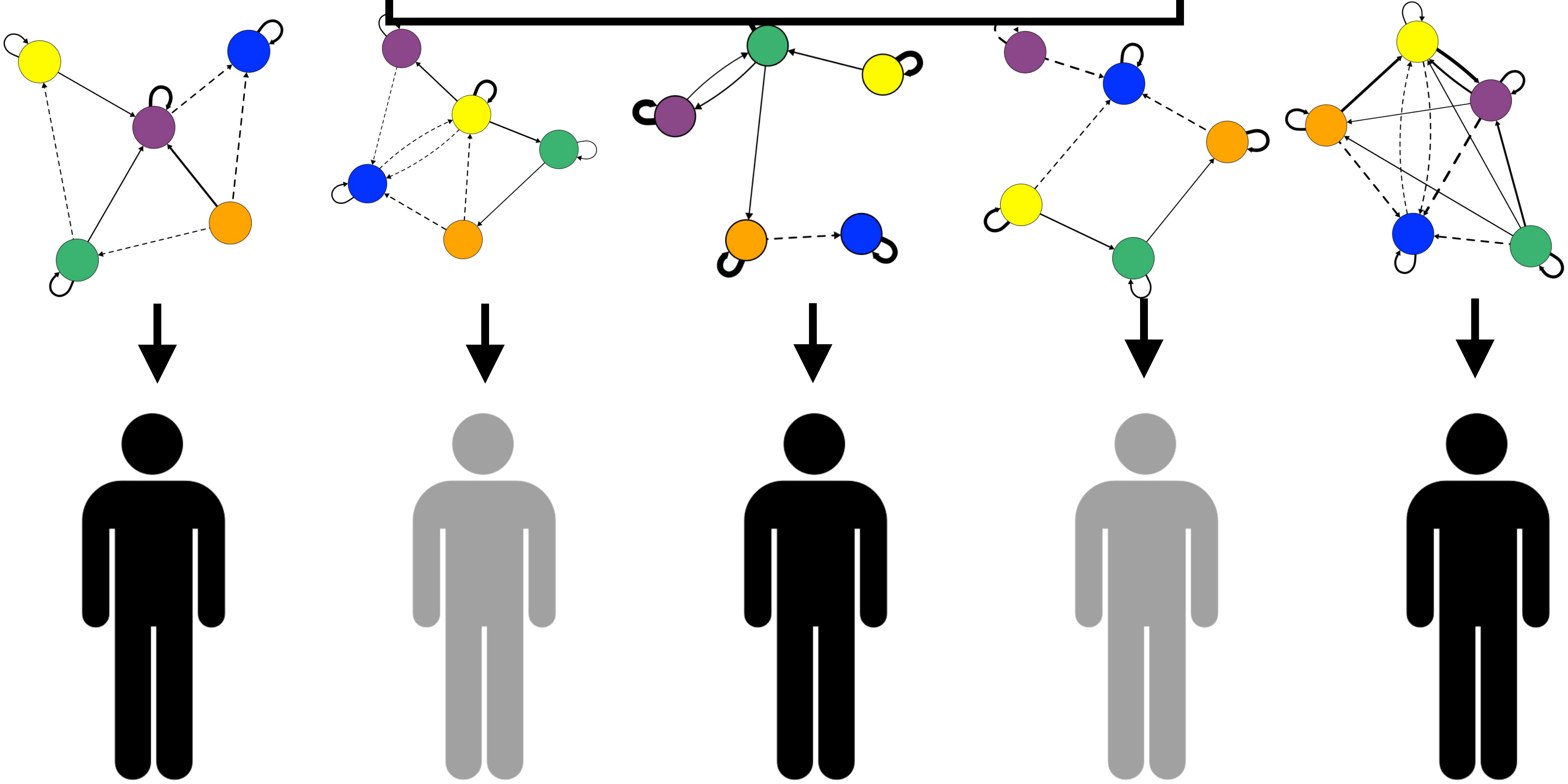
	V1	V2	V3	V4	V5
V1	.11	.04	-.12	.22	.15
V2	0.02	.43	.16	-.20	-.36
V3	-0.05	0.08	-.25	-.12	.37
V4	0.64	-0.12	-0.10	.60	.10
V5	0.26	-0.27	0.43	0.07	.33



Julian Burger's "Dynamic Network Analysis"; Aidan Wright's "GIMME"

Part 2: Statistical Methods

Cross-Lagged VAR, GIMME, and EGA



Part 2: Statistical Methods

Cross-Lagged VAR, GIMME, and EGA

Multilevel Vector Autoregressive Models	Graphical Vector Autoregressive Models	GIMME (Unified Structural Equation Models)	Dynamic Exploratory Graph Analysis
Between- and within-person (shrinkage)	Idiographic	Idiographic with constrained group-level pathways (by consensus)	Idiographic (with some multilevel options under development)
Frequentist or Bayesian Significance	Graphical LASSO (i.e. regularization)	Step-forward using Lagrange Multiplier Tests	Graphical LASSO & triangulated maximally filtered graph
Between-, within-, and person-specific effects	Results in correlation terms	Sub-group analysis, moderator pathways	Uses GLLA and embedding dimensions to capture change in continuous time metrics
`mlVAR::mlVAR()`	`graphicalVAR::graphicalVAR()`	`gimme::gimme()`	`EGAnet::dynEGA()`
Bringmann et al., 2016	Wild et al., 2010; Epskamp et al., 2018	Beltz & Gates, 2017; Lane et al., 2019	Golino & Epskamp, 2017; Golino et al., 2020

Part 2: Statistical Methods

Machine Learning

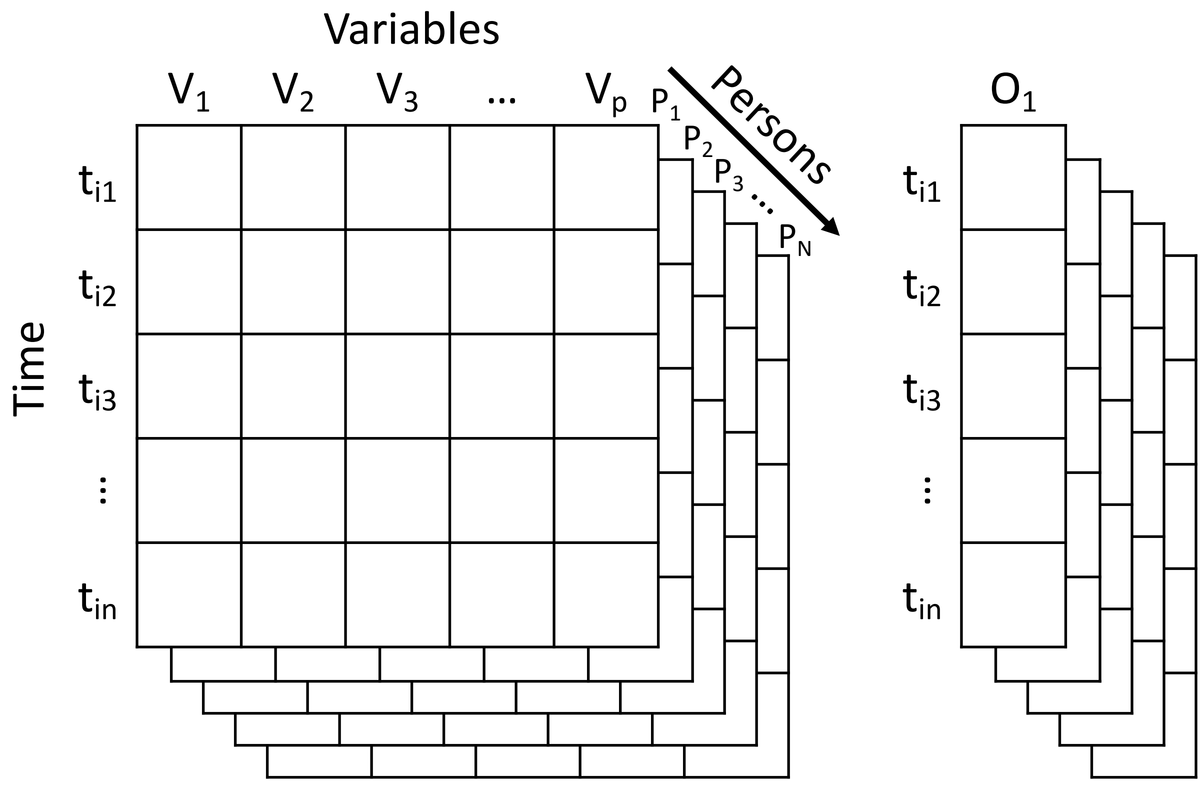
~~Prediction: Does X predict Y?~~

~~Prediction: Does X predict Y across people?~~

Are there individual differences in person-specific prediction?

Part 2: Statistical Methods

Machine Learning



Regularized (Logistic) Regression

Random Forest

Naïve Bayes

BISCU(W)IT

Neural Networks

Support Vector Machines

… And more

All supported by cross-validation, regularization, ensemble methods (e.g., bagging)

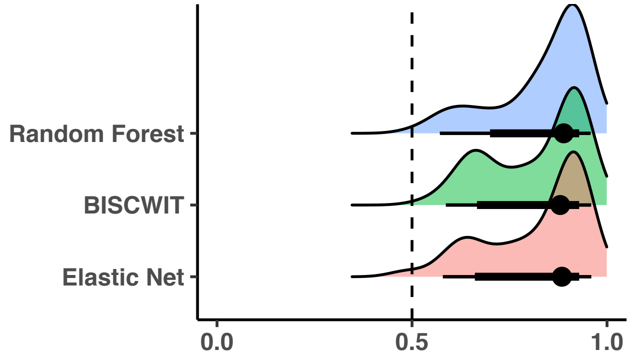
Part 2: Statistical Methods

Machine Learning

Are there individual differences in person-specific prediction?

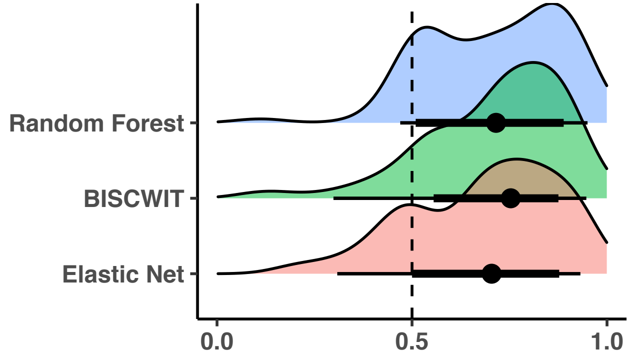
Procrastination

Accuracy



M (SD)	Median	N	Range
0.82 (0.13)	0.88	82	0.48–0.96
0.82 (0.12)	0.88	89	0.52–0.96
0.83 (0.12)	0.89	87	0.50–0.96

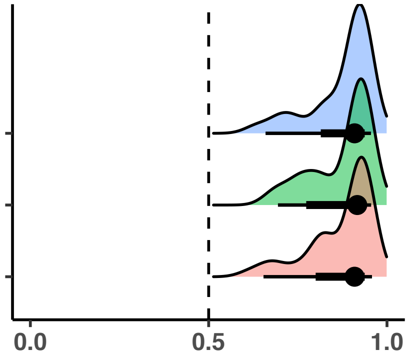
AUC



M (SD)	Median	N	Range
0.69 (0.19)	0.70	82	0.20–0.98
0.71 (0.18)	0.75	89	0.10–0.98
0.71 (0.17)	0.72	86	0.11–0.96

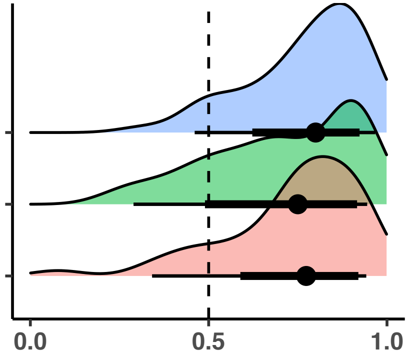
Loneliness

Accuracy



M (SD)	Median	N	Range
0.87 (0.09)	0.91	50	0.61–0.96
0.87 (0.08)	0.92	51	0.67–0.96
0.87 (0.09)	0.91	50	0.62–0.96

AUC

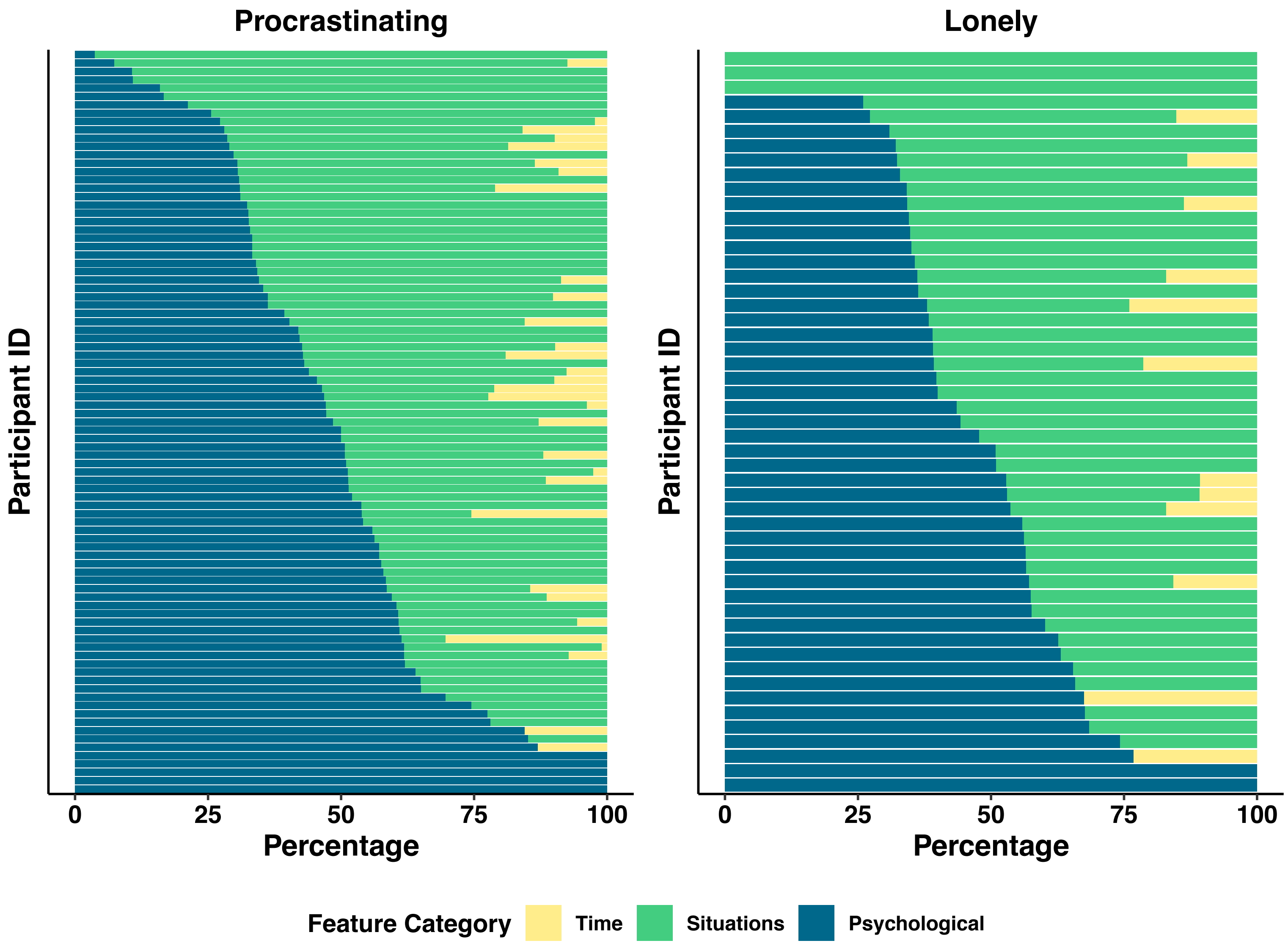


M (SD)	Median	N	Range
0.74 (0.19)	0.77	50	0.08–0.95
0.71 (0.21)	0.75	51	0.22–0.97
0.77 (0.16)	0.80	50	0.32–0.97

Beck & Jackson, under review; See also [Fisher & Soyster, 2019]; [Soyster et al., 2020], [Butter et al., 2020]

Part 2: Statistical Methods

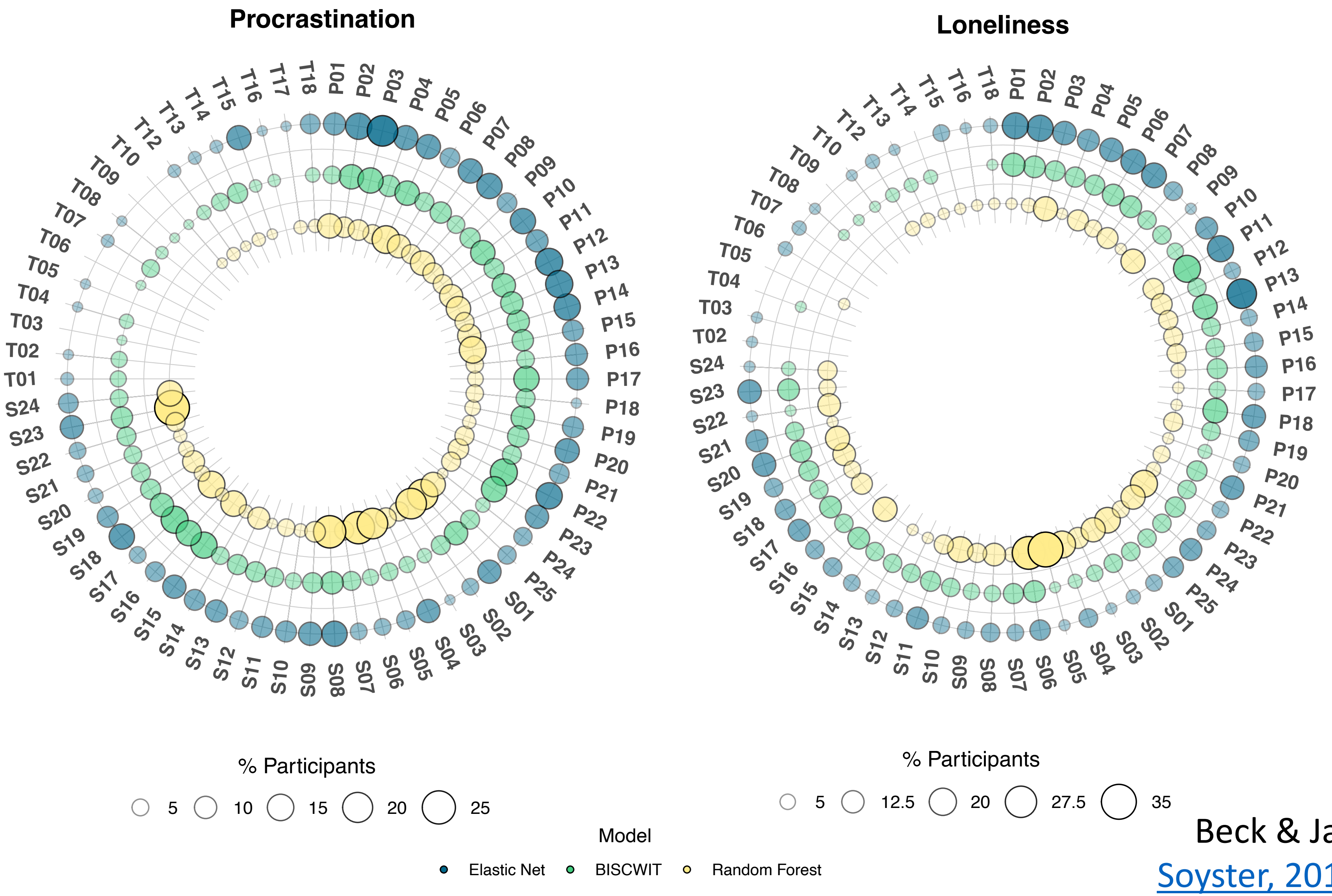
Machine Learning



Beck & Jackson, under review; See also Fisher & Soyster, 2019; Soyster et al., 2020, Butter et al., 2020

Part 2: Statistical Methods

Machine Learning



P01: Extraversion: Sociability
P02: Extraversion: Assertiveness
P03: Extraversion: Energy Level
P04: Agreeableness: Compassion
P05: Agreeableness: Respectfulness
P06: Agreeableness: Trust
P07: Conscientiousness: Organization
P08: Conscientiousness: Productivene
P09: Conscientiousness: Responsibilit
P10: Neuroticism: Anxiety
P11: Neuroticism: Depression
P12: Neuroticism: Emotional Volatility
P13: Openness: Intellectual Curiosity
P14: Openness: Aesthetic Sensitivity
P15: Openness: Creative Imagination
P16: Negative: Angry
P17: Negative: Afraid
P18: Positive: Happy
P19: Positive: Excited
P20: Positive: Proud
P21: Negative: Guilty
P22: Positive: Attentive
P23: Positive: Content
P24: Neutral: Purposeful
P25: Neutral: Goal–directed
S01: Duty
S02: Intellect
S03: Adversity
S04: Mating
S05: pOsitivity
S06: Negativity
S07: Deception
S08: Sociability
S09: Studying
S10: Argument
S11: Interacted
S12: Lost something
S13: Late
S14: Forgot something
S15: Bored with schoolwork
S16: Excited about schoolwork
S17: Anxious about schoolwork
S18: Tired
S19: Sick
S20: Sleeping
S21: In Class
S22: Listening to music
S23: On the internet
S24: Watching TV
T01: Monday
T02: Tuesday
T03: Wednesday
T04: Thursday
T05: Friday
T06: Saturday
T07: Sunday
T08: Morning
T09: Midday
T10: Evening
T12: Linear Trend
T13: Quadratic Trend
T14: Cubic Trend
T15: 24 hour Sinusoidal Cycle
T16: 12 hour Sinusoidal Cycle
T17: 24 hour Cosinusoidal Cycle
T18: 12 hour Cosinusoidal Cycle

Beck & Jackson, under review; See also Fisher & Soyster, 2019; Soyster et al., 2020, Butter et al., 2020

Part 2: Statistical Methods

Differential Equation Models

Dynamical Systems Theory

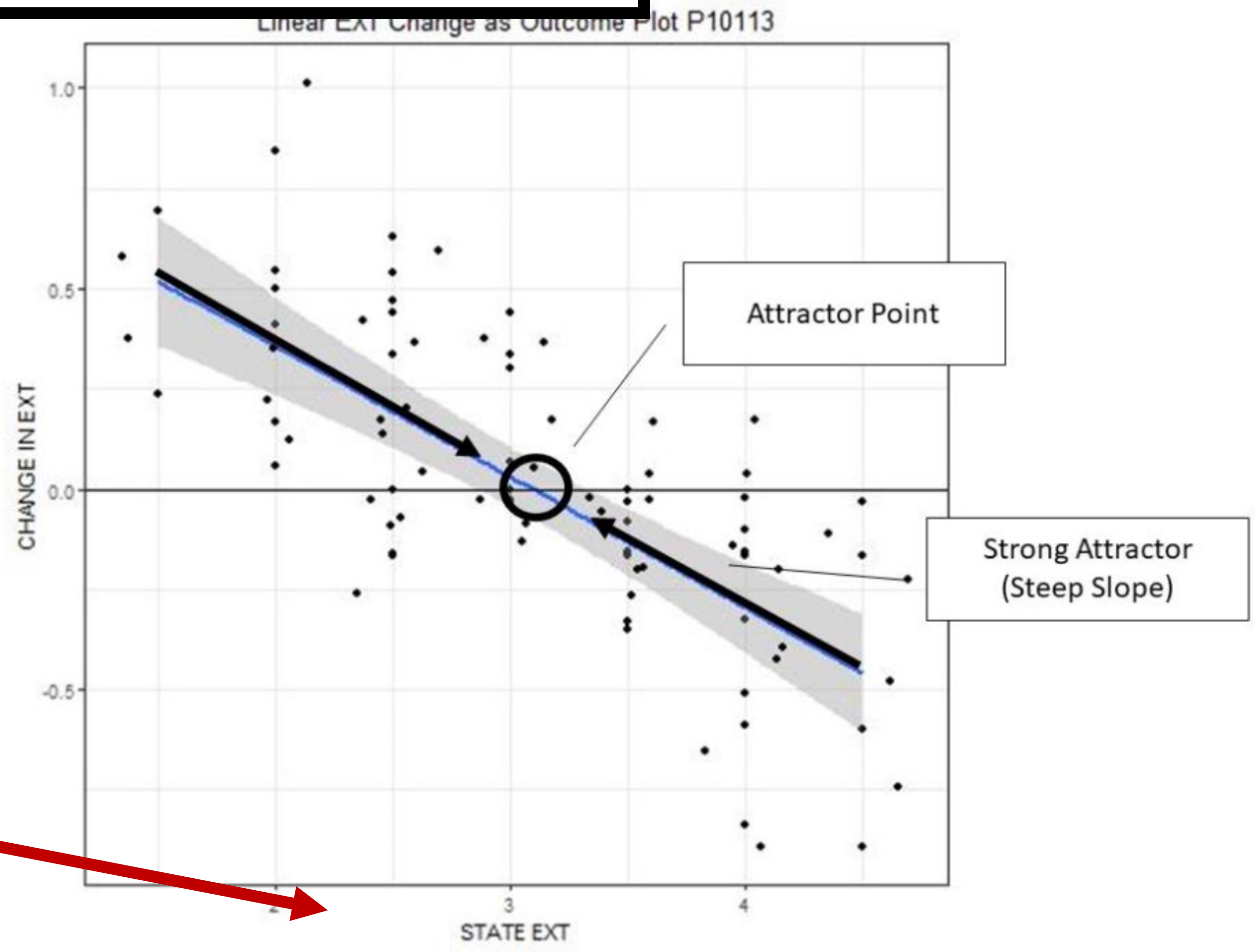
Equilibria

Attractors

Repellers

$$\Delta X = b_0 + b_1 X$$

Location: $0 = b_0 + b_1 X$



Danvers, Wundrack, & Mehl, 2020; Butner, 2014, Boker et al., 2009
Butler & Cuelz, 2020, Revelle & Condon, 2015, Revelle & Wilt, 2020

Eiko Fried & Don Robinaugh's "Formalizing Verbal Theories"

Part 2: Statistical Methods

Differential Equation Models

Dynamical Systems Theory

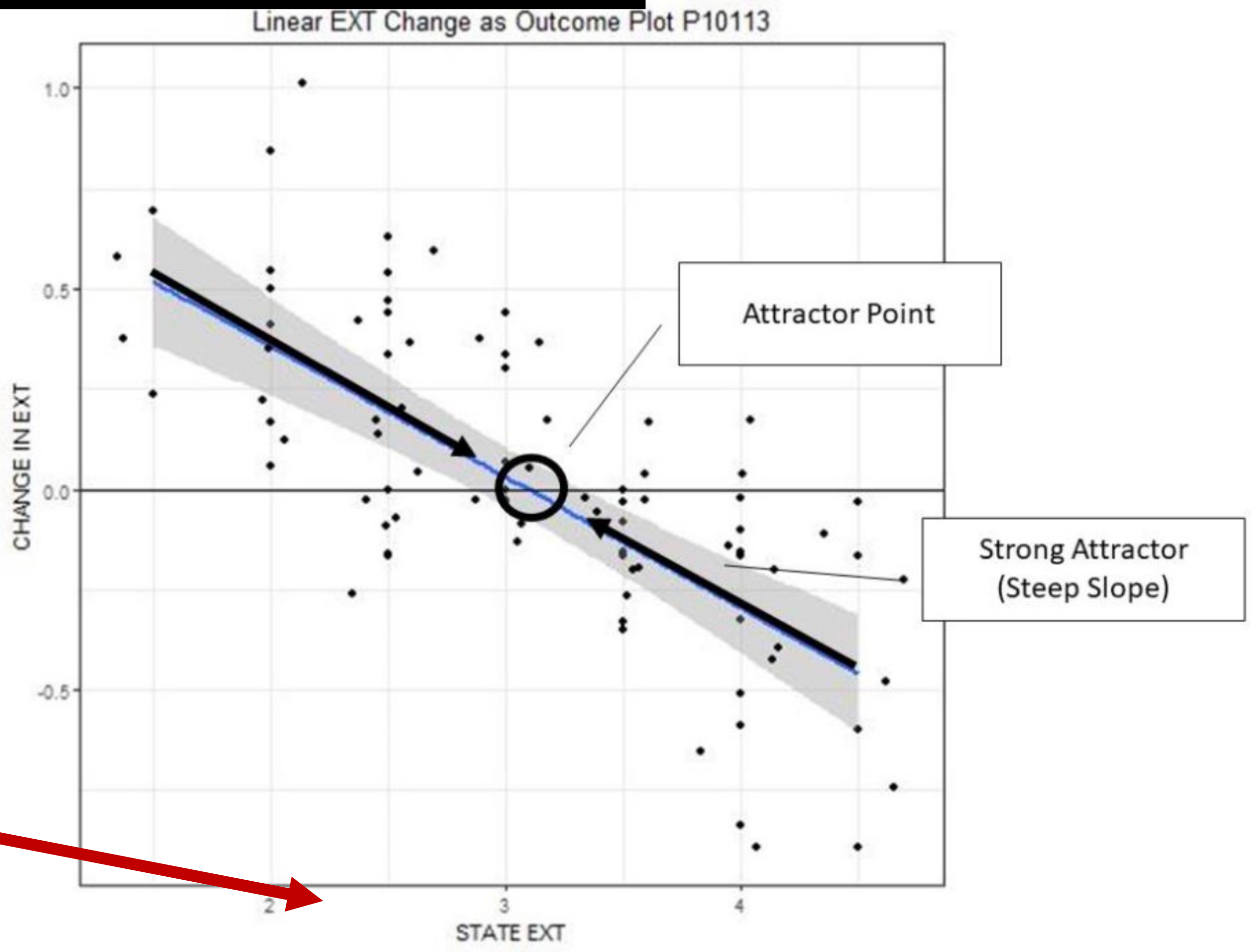
Equilibria

Attractors

Repellers

$$\Delta X = b_0 + b_1 X$$

Location: $X = -\dfrac{b_0}{b_1}$



Danvers, Wundrack, & Mehl, 2020; Butner, 2014, Boker et al., 2009
Butler & Cuelz, 2020, Revelle & Condon, 2015, Revelle & Wilt, 2020

Eiko Fried & Don Robinaugh's "Formalizing Verbal Theories"

Part 2: Statistical Methods

Differential Equation Models

Dynamical Systems Theory

Equilibria

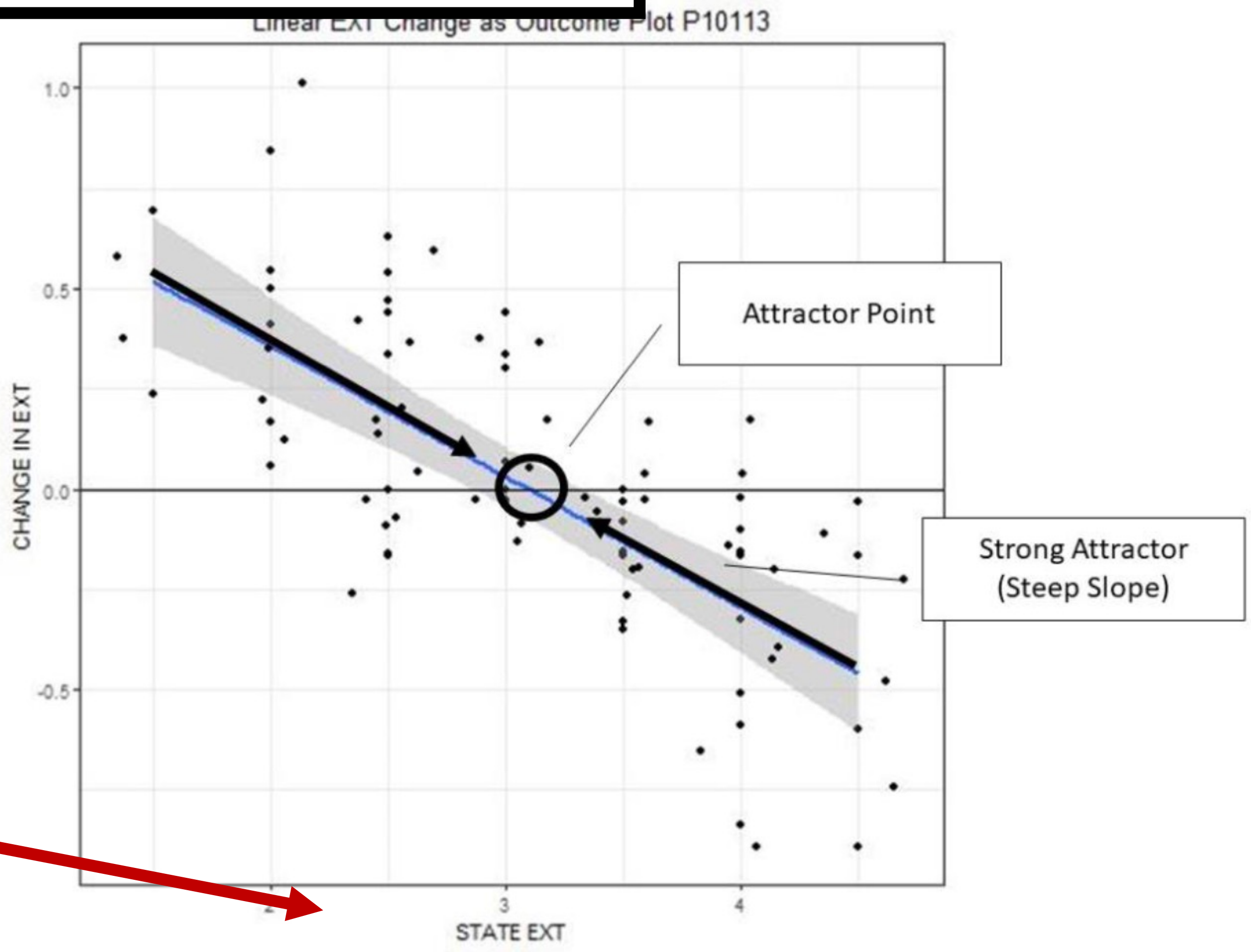
Attractors

Repellers

$$\Delta X = b_0 + b_1 X$$

Location: $X = -\dfrac{b_0}{b_1}$

Strength: b_1




```
lm()
EGAnet::glla()
psych::cta()
rties::estDerivs(); rties::indivClo()
```

Danvers, Wundrack, & Mehl, 2020; Butner, 2014, Boker et al., 2009
Butler & Cuelz, 2020, Revelle & Condon, 2015, Revelle & Wilt, 2020

Thank you!

https://github.com/emoriebeck/R-tutorials

@EmorieBeck

emorie_beck@northwestern.edu