Item Validation in the WARN-D Study

Assessing EMA items to build an early warning system for depression

Björn Siepe, PhD Student
Department of Psychological Methods, Philipps University Marburg, Germany

@b_siepe bjoernsiepe@gmail.com

WARN-D

WARN-D

Goal: Develop a personalized early warning system for depression

How?

Follow 2,000 students over 2 years

Integrate multiple modern assessment types

This talk:

How can we learn more about all the EMA items used in the study?

Sample

Data Collection

Recruitment
- Four Cohorts
- 500 students each

Population
- Students of higher education in The Netherlands > 18 years

This Talk
- Cohort 1 + 2
- (C1: Nov 2021; C2: May 2022)
- *N* = 865

Measurement

Baseline

3 Months EMA-Phase
- 4x /day (18-21 items)
- Additional sunday survey (46 items)
- Many items: 1-7 Likert
- Smartwatch

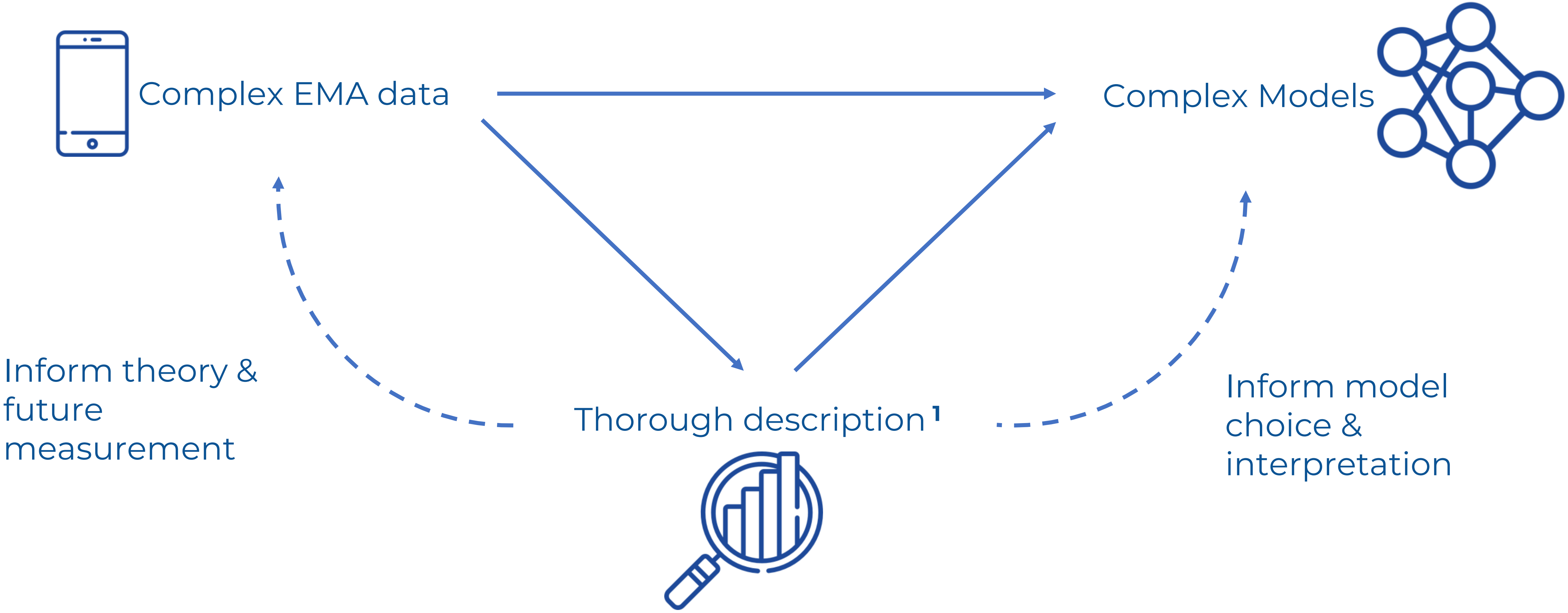
Multiple Follow-Ups

WARN-D

Theoretical Relevance

Complex EMA data → Complex Models

Theoretical Relevance



1 | e.g. Anscombe (1973), Tukey (1977), Haig (2013)

Validation Workflow

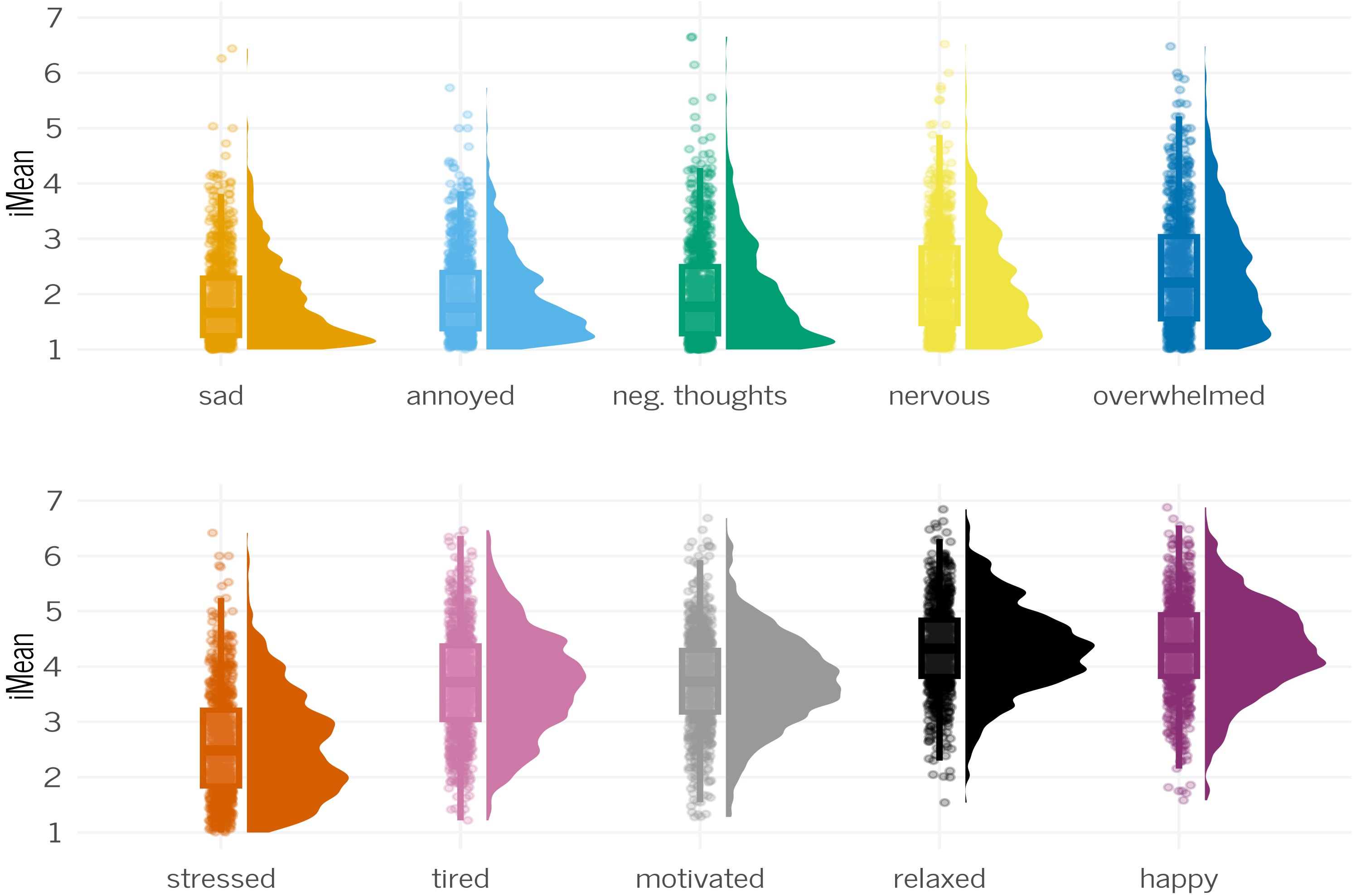
Visualization & Description

1. Distributions
2. Context
3. Time

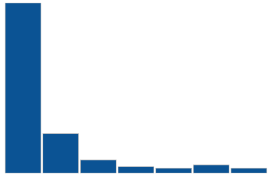
Other parts

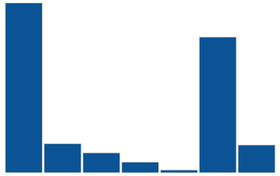
More on this later…

Distributions



Distributions





Floor Effects:

multiple items with ~30% of individuals answering 1 (“*not at all*”) >80% of the time

Multimodality:

multiple negative affect items with ~20% of individuals showing multimodality [1]

Theoretically: Conceptualization of responses [2]

Statistically: Important to account for [3]

1 | Haslbeck et al. (2021) **2** | Ryan et al. (2021) **3** | von Klipstein et al. (2021)

Context

- One of the main pillars of EMA

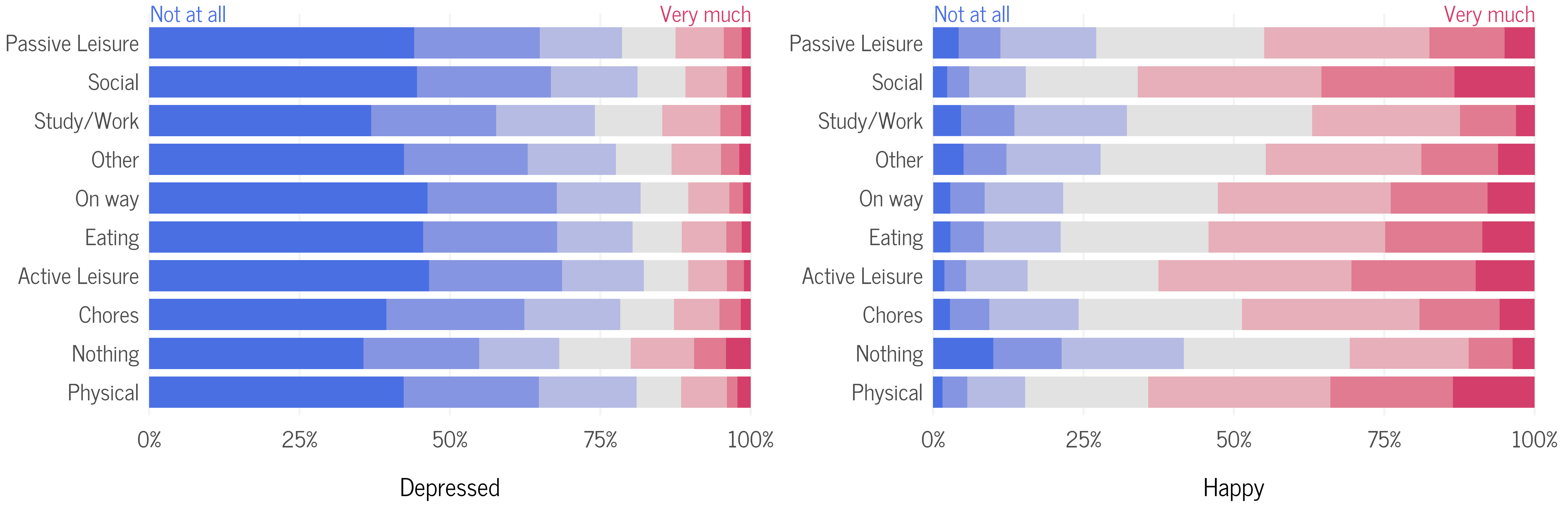
- In WARN-D:
  - Current activity
  - Online & offline social contacts
  - Location

enjoyment thereof

How similar are responses across different contexts?

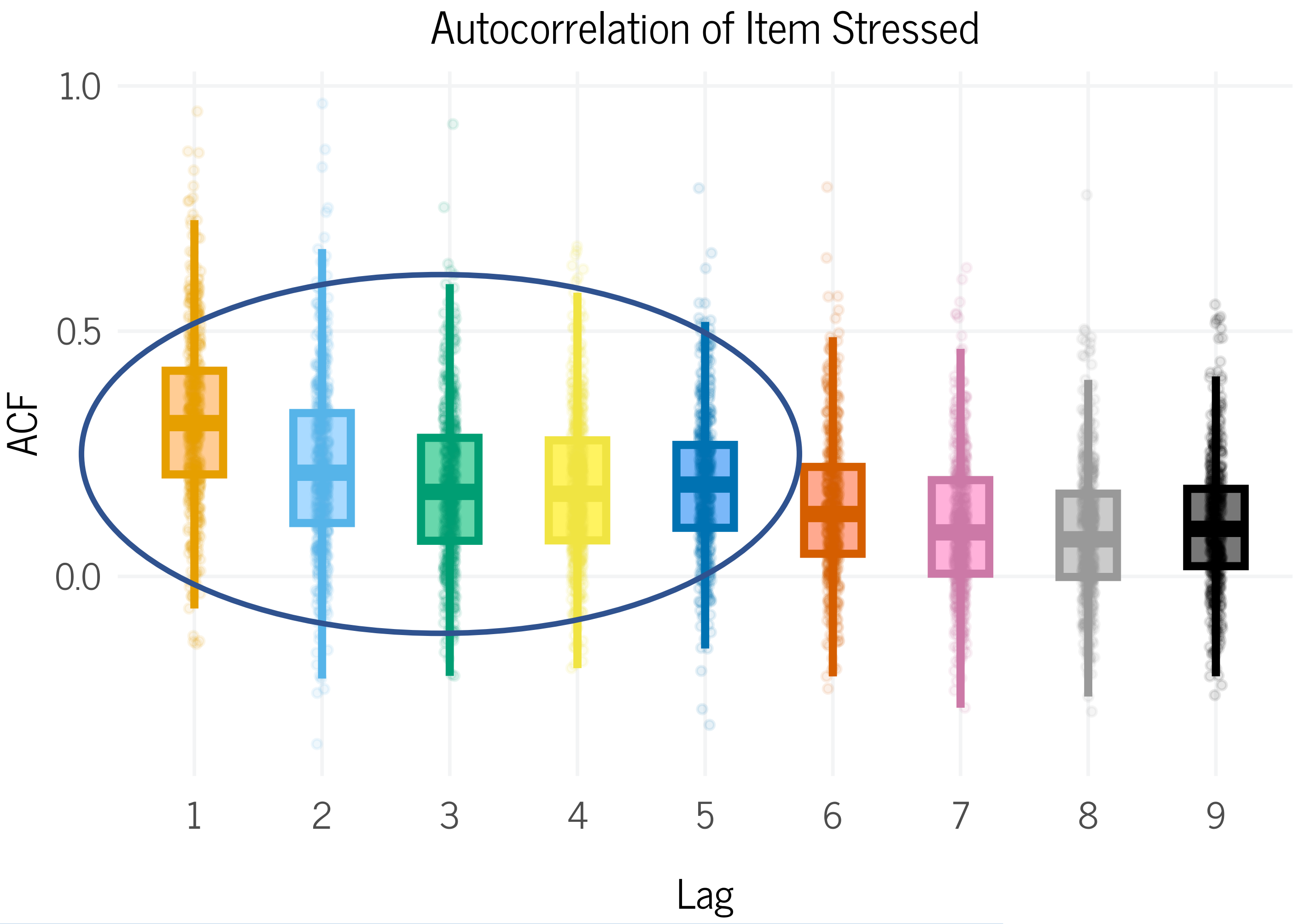
WARN-D

Context



WARN-D

Time


Autocorrelation of Item Stressed
1.0
0.5
0.0
ACF
1
2
3
4
5
6
7
8
9
Lag

WARN-D

Validation Workflow

Visualization & Description

Other parts

1. Distributions
2. Context
3. Time

WARN-D

Validation Workflow

Other parts

1. Validating EMA with sensor data
2. Self-reported change vs. inferred change from EMA
3. Prognostic validity
4. Concurrent validity
5. Measurement models

1. Overlap between EMA stress and sensor stress
2. Asking people if they changed vs. calculating change gives different answers
3. Predicting follow-up with EMA
4. Overlap daily vs. weekly depression items
5. Work-in-Progress

WARN-D

Summary

Implications for WARN-D	General
• Improved understanding of items and their characteristics • Structured workflow • Internal package, external extensive codebook	• Importance of descriptive work • Flexible modeling • More quantitative and qualitative work on item responses, response processes • Broad conception of validity and item quality

WARN-D

Thank you

The WARN-D team

ERC Starting Grant 2020, agreement No. 949059

Leiden University, the Netherlands

@b_siepe

bjoernsiepe@gmail.com

WARN-D

Keeping up

Slides &
My Homepage

WARN-D

Other WARN-D talks today:

Today, 1pm:
Introducing Fred, Open Software to Create Personalized Reports Utilizing Time-Series Data and Network Science
(Eiko I. Fried)

Today, 4 pm:
Augmenting Self-Reports with Passive Sensor Data to Understand Changes in Mental Health Presenter
(Carlotta L. Rieble)

@b_siepe

bjoernsiepe@gmail.com

WARN-D

Image Sources

https://www.flaticon.com/free-icons/data
Data icons created by Freepik - Flaticon

https://www.flaticon.com/free-icons/machine-learning"
Machine learning icons created by Becris - Flaticon

https://www.flaticon.com/free-icons/smartphone
Smartphone icons created by Freepik – Flaticon

https://www.flaticon.com/free-icons/idea"
Idea icons created by Freepik - Flaticon

References

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician, 27* (1), 17–21.
Haig, B. D. (2013). Detecting psychological phenomena: Taking bottom-up research seriously. *The American Journal of Psychology, 126* (2), 135–153. https://doi.org/10.5406/amerjpsyc.126.2.0135
Haslbeck, J., Ryan, O., & Dablander, F. (2023). Multimodality and skewness in emotion time series. *Emotion*. Advance online publication. https://doi.org/10.1037/emo0001218
Ryan, O., Dablander, F., & Haslbeck, J. M. B. (2023, March 17). Towards a Generative Model for Emotion Dynamics. https://doi.org/10.31234/osf.io/x52ns
Tukey, J. W. (1977). *Exploratory data analysis (Vol. 2).* Reading, MA
von Klipstein, L., Servaas, M., Lamers, F., Schoevers, R. A., Wardenaar, K. J., & Riese, H. (2022, May 20). Can Floor Effects Explain Increased Affective Reactivity Among Depressed Individuals Found in Experience Sampling Research?. https://doi.org/10.31234/osf.io/6vtew