Modeling Learners Beyond the Mean

https://bit.ly/PeterE_presentations

https://bit.ly/LikertVASPreprint





Experimenteller Vergleich von Likert Skala und visueller Analogskala im Ecological Momentary Assessment von Affekt bei Studierenden

Jonas M. B. Haslbeck*^{1,2}, Alberto Jover-Martínez*¹, Anne J. Roefs¹, Eiko I. Fried³, Lotte H. J. M. Lemmens¹, Esmee Groot¹, and Peter Adriaan Edelsbrunner^{4, 5}

¹Department of Clinical Psychological Science, Maastricht University
 ²Psychological Methods Group, University of Amsterdam
 ³Department of Psychology, Leiden University
 ⁴Department of Humanities, Social and Political Sciences, ETH Zurich
 ⁵Department of Psychology, LMU Munich

Within-person Perspektive Blume et al., 2022; Blume & Schmiedek, 2024

Ideographie 45. & Nomothetik

following	Please indicate to what extent you feel the following mood states and physical sensations at this moment:		
Sad			
Not at	all	Extremely	
Tap on the line to start!			

Kontrolle von Trait-Varianz Hamaker et al., 2007

Assessment von Dynamiken
Hasibeck & Ryan, 2022

Methodologie: Meistens unbegründet/intuitiv

Frühere Studien über Antwortformate: Hauptsächlich querschnittlich, höhere Antwortstufen (siehe simms et al., 2019; Reliabilität mit > 4 Fritz et al., 2024) Übertragbarkeit?

Typischer Ansatz: 7-stufige Likert oder visuelle lineares gemischtes Modell Analogskala, Haslbeck et al., 2023

Please indicate to what extent you feel the following mood states and physical sensations at this moment: Sad Not at all Extremely Tap on the line to start! Statistiken/Verteilungen

Wie wirkt sich die Antwortskala Forschungsfrage: aus? Auf... Statistiken/Verteilungen ...1) Univariate 3) Zusammenhänge mit externen

4) Subjektive Bewertung

2) Multivariate

Kovariaten

Häufig untersucht Häufig verwendet

Häufig untersucht

Design

Kontext: Emotionen, single Item, Uni-Studierende (eher wenige kl. Diagnosen)

Positive (happy, satisfed, ..., 5 Items) und negative (stressed, worried,... 9 Items) Valenz

7-stufige Antwortskala vs. visuelle Analogskala (101-stufig)

Möglichst parallel (z.B. nur Endlabels bei 7-stufiger Skala)

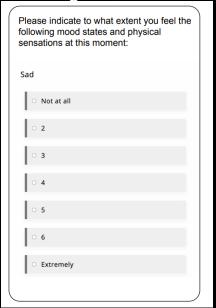
Einschränkung durch App:

7-stufig vertikal, VAS horizontal

7-stufig = Not at all, 2, 3, 4, 5, 6, Extremely

Between-person (Vermeidung von Übertragungseffekten/Manipulationsreaktanz):

Zwei Wochen, 3x täglich, entweder immer 7-stufig oder immer VAS



Please indicate to wh following mood states sensations at this mo	and physical
Sad	
Not at all	Extremely
Tap on the l	line to start!

N(Likert) = 63, N(VAS) = 56M(age) = 22 (SD = 4)

Modellierungen:

Bayesianische Mehrebenenmodelle in brms

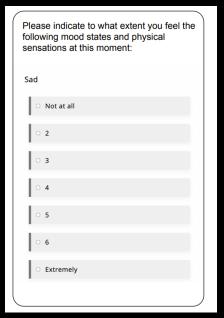
Verteilungen: Deskriptive Statistiken pro Person berechnet und deren Verteilungen Modelliert. z.B. Within-person means: zero-inflated beta

Fixed effects: Antwortformat (contrast coding, baseline = Likert), Valenz (effects coding) + Interaktion

Random effects: Intercept über Personen, Random slopes über Items

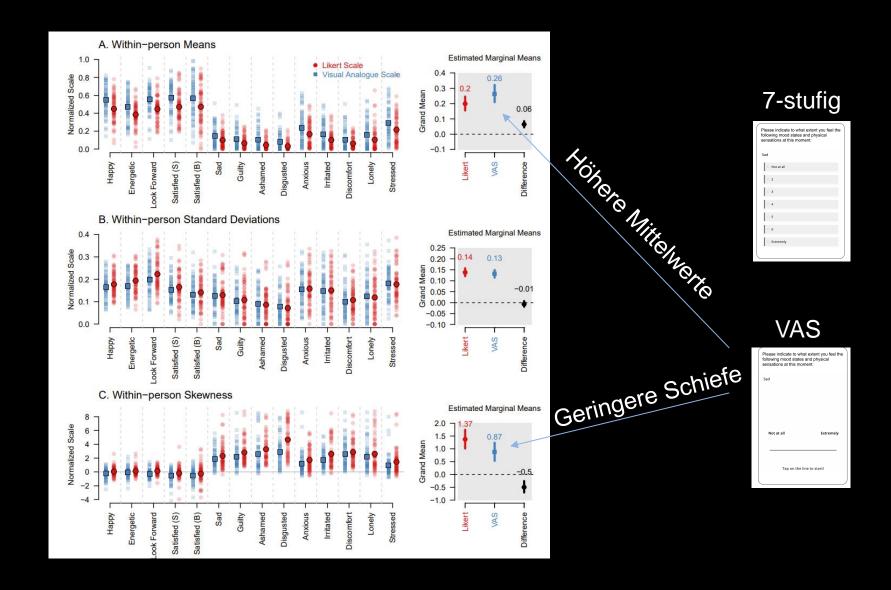
Distributional models: Fixed effects auf alle Verteilungsparameter (Zero-Anteil, beide beta-

Parameter)

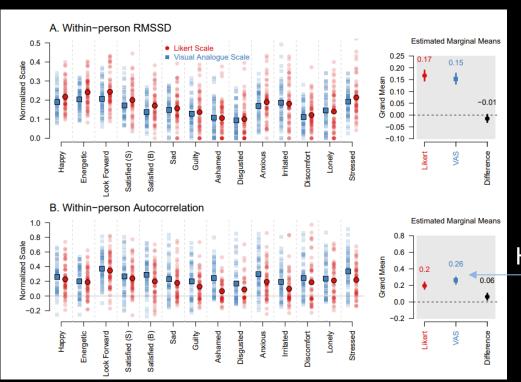


Please indicate to wh following mood states sensations at this mo	and physical
Sad	
Not at all	Extremely
Tap on the	ine to start!

Ergebnisse: Deskriptive Verteilungsparameter



Ergebnisse: Within-Person Dynamiken



7-stufig

Please indicate to what extent you feel the following mood states and physical sensations at this moment.

Sad

Net at all

2

3

4

5

Externely

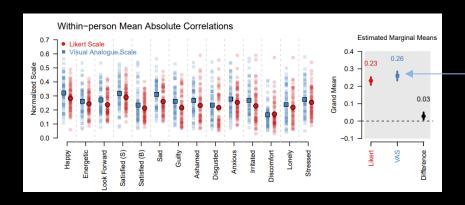
Höhere Autokorrelation



Ergebnisse: Multivariate Verteilungen

7-stufig

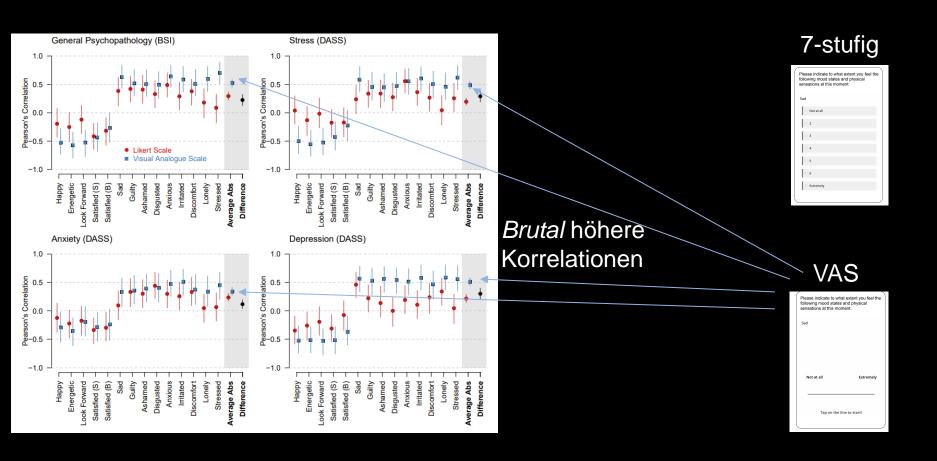




Etwas höhere within-Person Itemkorrelationen



Ergebnisse: Korrelationen mit externen Kriterien (Skalen klinischer Symptome)



Ergebnisse: Weitere Ergebnisse

Subjektive Einschätzung recht ähnlich; deskriptiv VAS etwas anstrengender und unklarer

Wenig Multimodalität auf beiden Skalen (vgl. Haslbeck et al., 2023)

Missing data & Bearbeitungszeit ähnlich

Kein Initial Elevation Bias (Shrout et al., 2017)

7-stufig





Ergebnisse: Schlussfolgerung

Unsere VAS greift mehr systematische Variation im niedrigen Bereich ab.

(Höhere Mittelwerte, weniger Schiefe, höhere Korrelationen mit externen Kriterien) – vorläufige Empfehlung in *ähnlichen Studienkontexten*.

Einschränkung:

- Konstrukte, bei denen sich wichtige Varianz nah an den Skalenlimits (Min./Max.) abspielt (z.B. klinische Symptome bei nicht-klinischen Populationen)
- 7-stufige Skala: Zweite Kategorie «2», was wäre bei «1» oder (echtem Likert-)Label «gering»/»sehr gering»?
- Wir brauchen (Replikationen &) faktorielle Studien, die viele Faktoren (Anzahl Stufen, Position & Inhalte der Labels, horizontal vs. vertikal, Startpunkt auf VAS) gleichzeitig randomisieren, (Baribault et al., 2018), potentiell within-Person
 - Sowie kognitive Surveys und formale Modelle (Antwortprozess ergründen & modellieren)
 - Multi-Item Skalen

7-stufig





https://bit.ly/PeterE_presentations

https://bit.ly/LikertVASPreprint





Baribault, B., Donkin, C., Little, D. R., Trueblood, J. S., Oravecz, Z., van Ravenzwaaij, D., White, C. N., De Boeck, P., & Vandekerckhove, J. (2018). Metastudies for robust tests of theory. Proceedings of the National Academy of Sciences, 115(11), 2607–2612. https://doi.org/10.1073/pnas.1708285114

Blume, F., Irmer, A., Dirk, J., & Schmiedek, F. (2022). Day-to-day variation in students' academic success: The role of self-regulation, working memory, and achievement goals. Developmental Science, 25(6), e13301. https://doi.org/10.1111/desc.13301 Blume, F., & Schmiedek, F. (2024). It counts in every single lesson: Between-and within-person associations of teaching quality and student self-regulation. Learning and Instruction, 92, 101908.

Fritz, J., Piccirillo, M. L., Cohen, Z. D., Frumkin, M., Kirtley, O., Moeller, J., Neubauer, A. B., Norris, L. A., Schuurman, N. K., Snippe, E., & Bringmann, L. F. (2024). So You Want to Do ESM? 10 Essential Topics for Implementing the Experience-Sampling Method. Advances in Methods and Practices in Psychological Science, 7(3), 25152459241267912.

https://doi.org/10.1177/25152459241267912

Hamaker, E. L., Nesselroade, J. R., & Molenaar, P. C. (2007). The integrated trait–state model. Journal of research in personality, 41(2), 295–315. Haslbeck, J. M. B., & Ryan, O. (2022). Recovering Within-Person Dynamics from Psychological Time Series. Multivariate Behavioral Research, 57(5), 735–766. https://doi.org/10.1080/00273171.2021.1896353

Haslbeck, J., Ryan, O., & Dablander, F. (2023). Multimodality and skewness in emotion time series. Emotion.

https://psycnet.apa.org/record/2023-72233-001

Molenaar, P. C., & Valsiner, J. (2009). How generalization works through the single case: A simple idiographic process analysis of an individual psychotherapy. YIS: Yearbook of idiographic science, 1, 23–38.

Shrout, P. E., Stadler, G., Lane, S. P., McClure, M. J., Jackson, G. L., Clavél, F. D., Iida, M., Gleason, M. E. J., Xu, J. H., & Bolger, N. (2018). Initial elevation bias in subjective reports. Proceedings of the National Academy of Sciences, 115(1), E15–E23.

Simms, L. J., Zelazny, K., Williams, T. F., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. Psychological Assessment, 31(4), 557–566. https://doi.org/10.1037/pas0000648