CSS – Institut für Psychologie, Universität Graz
10.9.2024, Graz/Online, Austria

Measurement bias in intensive longitudinal data

Georg Krammer
Institut für Wirtschafts- und Berufspädagogik
Johannes Kepler Universität Linz, Austria

Institute of
Business and
Vocational Education

Georg Krammer - Googl ×

https://scholar.google.at/citations?user=GMM_XDsAAAAJ&hl=en&oi=ao 130%

Google Scholar

Georg Krammer

FOLLOW

GET MY OWN PROFILE

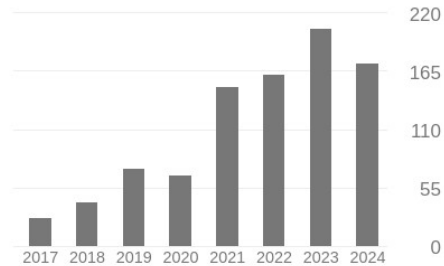
Johannes Kepler University Linz
Verified email at jku.at - Homepage

Educational Psychology Psychometrics Personality Open Science Teacher Education

TITLE	CITED BY	YEAR
Assessment of creativity evaluation skills: A psychometric investigation in prospective teachers M Benedek, N Nordtvedt, E Jauk, C Koschmieder, J Pretsch, G Krammer, ... Thinking Skills and Creativity 21, 75–84	167	2016
Creativity myths: Prevalence and correlates of misconceptions on creativity M Benedek, M Karstendiek, SM Ceh, RH Grabner, G Krammer, I Lebuda, ... Personality and Individual Differences 182, 111068	73	2021
Humor styles across 28 countries JA Schermer, R Rogoza, MM Kwiatkowska, CM Kowalski, S Aquino, ... Current Psychology, 1-16	53	2019
The life-skills program *Lions Quest* in Austrian schools: implementation and outcomes M Matischek-Jauk, G Krammer, H Reicher Health promotion international 33 (6), 1022-1032	45	2018
Believing in neuromyths makes neither a bad nor good student-teacher: The relationship between neuromyths and academic achievement in teacher education G Krammer, SE Vogel, RH Grabner Mind, Brain, and Education 15 (1), 54-60	44	2021
Aspects of online teaching and their relation to positive experience and motivation among teacher education students: mixed-method findings at the beginning of COVID-19 G Krammer, B Pflanzl, M Matischek-Jauk Zeitschrift für Bildungsforschung 10, 337-375	37 *	2020

Cited by VIEW ALL

	All	Since 2019
Citations	941	839
h-index	16	15
i10-index	28	25



Public access VIEW ALL

0 articles 4 articles

not available available

Based on funding mandates

TITLE	CITED BY	YEAR
Assessment of creativity evaluation skills: A psychometric investigation in prospective teachers M Benedek, N Nordtvedt, E Jauk, C Koschmieder, J Pretsch, G Krammer, ... Thinking Skills and Creativity 21, 75–84	167	2016
Creativity myths: Prevalence and correlates of misconceptions on creativity M Benedek, M Karstendiek, SM Ceh, RH Grabner, G Krammer, I Lebuda, ... Personality and Individual Differences 182, 111068	73	2021
Humor styles across 28 countries JA Schermer, R Rogoza, MM Kwiatkowska, CM Kowalski, S Aquino, ... Current Psychology, 1-16	53	2019
The life-skills program *Lions Quest* in Austrian schools: implementation and outcomes M Matischek-Jauk, G Krammer, H Reicher Health promotion international 33 (6), 1022-1032	45	2018
Believing in neuromyths makes neither a bad nor good student-teacher: The relationship between neuromyths and academic achievement in teacher education G Krammer, SE Vogel, RH Grabner Mind, Brain, and Education 15 (1), 54-60	44	2021
Aspects of online teaching and their relation to positive experience and motivation among teacher education students: mixed-method findings at the beginning of COVID-19 G Krammer, B Pflanzl, M Matischek-Jauk Zeitschrift für Bildungsforschung 10, 337-375	37 *	2020
Using students' feedback for teacher education: measurement invariance across pre-service teacher-rated and student-rated aspects of quality of teaching G Krammer, B Pflanzl, J Mayr Assessment & Evaluation in Higher Education 44 (4), 596-609	37	2019
Neuromythen sind zu Beginn des Lehramtsstudiums prävalent und unabhängig vom Wissen über das menschliche Gehirn G Krammer, SE Vogel, T Yardimci, RH Grabner Zeitschrift für Bildungsforschung 9 (2), 221-246	35	2019
TESAT–Ein neues Verfahren zur Eignungsfeststellung und Bewerberauswahl für das Lehramtsstudium A Neubauer, C Koschmieder, G Krammer, J Mayr, FH Müller, B Pflanzl, ... Zeitschrift für Bildungsforschung 7 (1), 5-21	30	2017
The psychometric costs of applicants' faking: Examining measurement invariance and retest correlations across response conditions G Krammer, M Sommer, ME Arendasy Journal of Personality Assessment 99 (5), 510-523	29	2017

Applicant faking of personality inventories in college admission: Applicants' shift from honest responses is unsystematic and related to the perceived relevance for the profession G Krammer Journal of Personality Assessment 102 (6), 758-769	13	2020
Assessing quality of teaching from different perspectives: Measurement invariance across teachers and classes G Krammer, B Pflanzl, G Lenske, J Mayr Educational Assessment 26 (2), 88-103	17	2021
Believing in neuromyths makes neither a bad nor good student-teacher: The relationship between neuromyths and academic achievement in teacher education G Krammer, SE Vogel, RH Grabner Mind, Brain, and Education 15 (1), 54-60	44	2021
Können wir jede Person lehren Lehrer*in zu werden? Sollen wir es? G Krammer, B Pflanzl journal für lehrerInnenbildung 19 (2), 28-39	5	2019
Open Science als Beitrag zur Qualität in der Bildungsforschung G Krammer, E Svecnik Zeitschrift für Bildungsforschung 10 (3), 263-278	12	2020
A cautionary note on aggregation in educational psychology and beyond G Krammer Theory & Psychology 33 (5), 681-700	2	2023

Institute of
Business and
Vocational Education

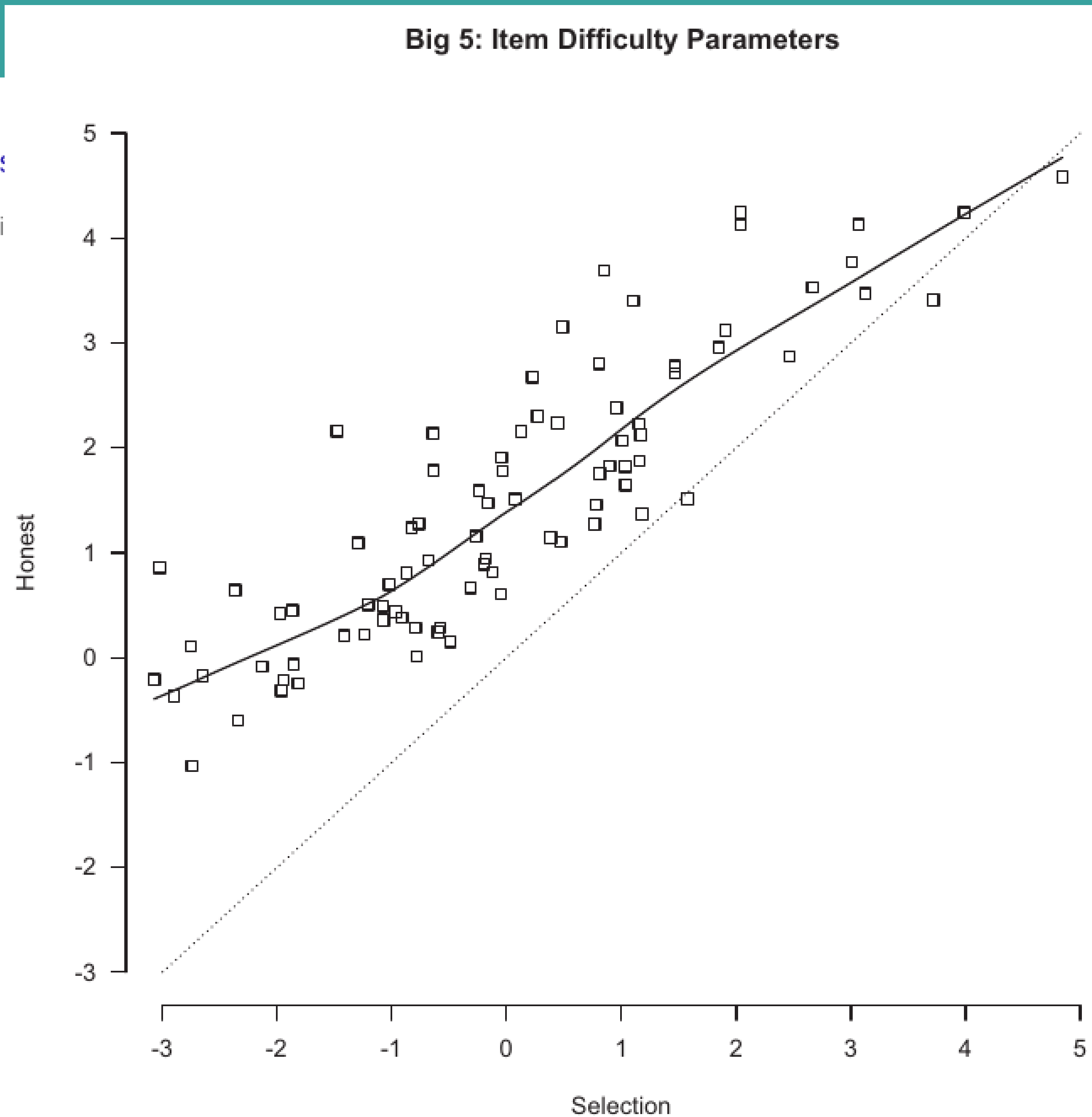
Applicant faking of personality inventories in college admission: Applicants' shift from honest responses is unsystematic and related to the perceived relevance for the profession | 13 | 2020

G Krammer

Journal of Personality Assessment 102 (6), 758-769

Applicant faking
responses is uns
G Krammer
Journal of Personali

13 2020

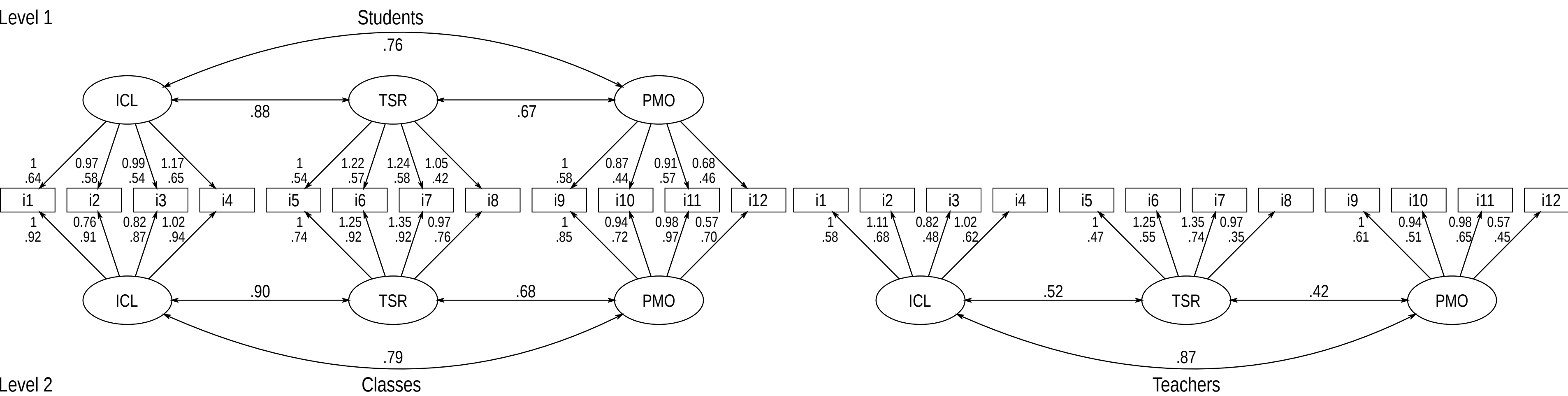


Institute of
Business and
Vocational Education

Assessing quality of teaching from different perspectives: Measurement invariance across teachers and classes 17 2021

G Krammer, B Pflanzl, G Lenske, J Mayr

Educational Assessment 26 (2), 88-103



Believing in neuromyths makes neither a bad nor good student-teacher: The relationship between neuromyths and academic achievement in teacher education G Krammer, SE Vogel, RH Grabner Mind, Brain, and Education 15 (1), 54-60	44	2021

Table 2
Comparison (BEST and Bayes Factor) of Overall grade point average (GPA) and GPAs of the Practical Courses of their First Academic Year (GPA 1st year) between ITE Students Believing (True) or Rejecting (False) Neuromyths (NM)

	Overall GPA			*GPA 1st year*		
	M *difference*	*Effect size*	*Bayes Factor*	M *difference*	*Effect size*	*Bayes Factor*
1. We only use 10% of our brain.	−0.021 [−0.092 0.053]	−0.094 [−0.425 0.226]	1/6.15	−0.004 [−0.017 0.003]	−0.136 [−0.467 0.191]	1/4.89
2. Individuals learn better when they receive information in their preferred learning style (e.g., auditory, visual, kinesthetic).	0.209 [−0.299 0.742]	0.560 [−0.622 1.789]	1/1.35	0.463 [0.121 0.71]	4.13 [−0.351 11.138]	1/1.86
3. Short bouts of co-ordination exercises can improve integration of left and right hemispheric brain function.	0.055 [−0.182 0.306]	0.213 [−0.656 1.106]	1/3.06	0.040 [−0.056 0.193]	0.401 [−0.796 1.658]	1/2.99
4. Differences in hemispheric dominance (left brain, right brain) can help explain individual differences amongst learners.	0.134 [−0.122 0.407]	0.532 [−0.429 1.488]	1/2.22	0.169 [−0.174 0.72]	0.618 [−0.945 2.436]	1/2.28
5. Children are less attentive after consuming sugary drinks and/or snacks.	0.013 [−0.076 0.103]	0.049 [−0.313 0.402]	1/5.26	0.000 [0.000 0.000]	0.001 [−0.303 0.302]	1/5.54
6. If pupils do not drink sufficient amounts of water (6–8 glasses a day) their brains shrink.	−0.096 [−0.273 0.076]	−0.407 [−1.141 0.316]	1/2.46	−0.128 [−0.28 0.007]	−0.929 [−1.988 0.053]	1/2.49
7. Learning problems associated with developmental differences in brain function cannot be remediated by education.	−0.124 [−0.24–0.008]	−0.463 [−0.907–0.027]	1.17	−0.027 [−0.091 0.014]	−0.320 [−0.887 0.211]	1/1.2
8. Children must acquire their native language before a second language is learned. If they do not do so neither language will be fully acquired.	0.076 [−0.014 0.169]	0.313 [−0.055 0.696]	1/1.82	0.000 [0.000 0.000]	0.002 [−0.321 0.317]	1/1.73
9. There are critical periods in childhood after which certain things can no longer be learned.	−0.034 [−0.112 0.043]	−0.144 [−0.465 0.187]	1/4.06	0.000 [0.000 0.000]	0.002 [−0.286 0.289]	1/6.17

Note. GPAs ranged from 1 to 5, with higher values representing higher academic achievement. The 95% high density interval of the estimated mean differences and effect sizes are given. Teacher education students indicated whether they believed in neuromyths (true), rejected neuromyths (false), or did not know.

Können wir jede Person lehren Lehrer*in zu werden? Sollen wir es? G Krammer, B Pflanzl journal für lehrerInnenbildung 19 (2), 28-39	5	2019
Open Science als Beitrag zur Qualität in der Bildungsforschung G Krammer, E Svecnik Zeitschrift für Bildungsforschung 10 (3), 263-278	12	2020

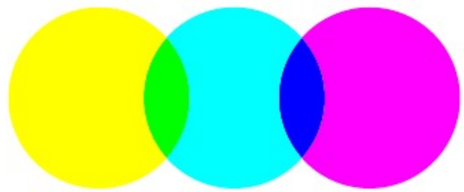
Institute of
Business and
Vocational Education

A cautionary note on aggregation in educational psychology and beyond G Krammer Theory & Psychology 33 (5), 681-700	2	2023

Institute of
Business and
Vocational Education

Habilitationsschrift

On aggregation in research on teacher education



Eingereicht an der Fakultät für Kulturwissenschaften
der Universität Klagenfurt

zur Erlangung der *venia docendi* in der Bildungsforschung

vorgelegt von
Georg Krammer, November 2021

A cautionary note on agg

G Krammer

Theory & Psychology 33 (5), 68

2 2023

Contents

Institute of Business and Vocational Education

681 1 of 20

Check for updates

Theory & Psychology

Article

Theory & Psychology
2023, Vol. 33(5) 681–700

DOI: 10.1177/09593543231172495
journals.sagepub.com/home/tap
Sage

A cautionary note on aggregation in educational psychology and beyond

Georg Krammer
University College of Teacher Education Styria

Abstract

This article addresses *aggregation* as a fundamental practice in educational psychology and ties it into the idiographic/nomothetic distinction, that is, distinguishing between studying *what once was* and studying *what always is*. I address the underlying assumptions of seminal educational research (OECD's large-scales assessment and Hattie's synthesizing meta-analyses). I argue that educational psychologists assume a priori general educational principles akin to nomothetic laws without sufficiently scrutinizing the limitations of aggregation. I then contextualize this assumption within the history of psychology, and address how these assumptions shape how educational psychologists view, collect, and examine data. Furthermore, I contextualize this assumption with an example showing a peculiarity of educational research: the existence of multiple perspectives on constructs. Finally, I argue that investing time and resources in the debate on aggregation and the epistemic nature of the insights that educational psychologists generate will ultimately advance the field and help bridge the theory–practice gap.

Keywords

aggregation, best practice, educational psychology, idiographic/nomothetic, quantitative psychology

Contents

Contents

Studying perspective-free and situation-unspecific overarching latent constructs may advance our field less than respecting differences in perspectives, across situations, and within aspects of constructs.

Measurement bias in intensive longitudinal data

The longer (preprinted) reads:

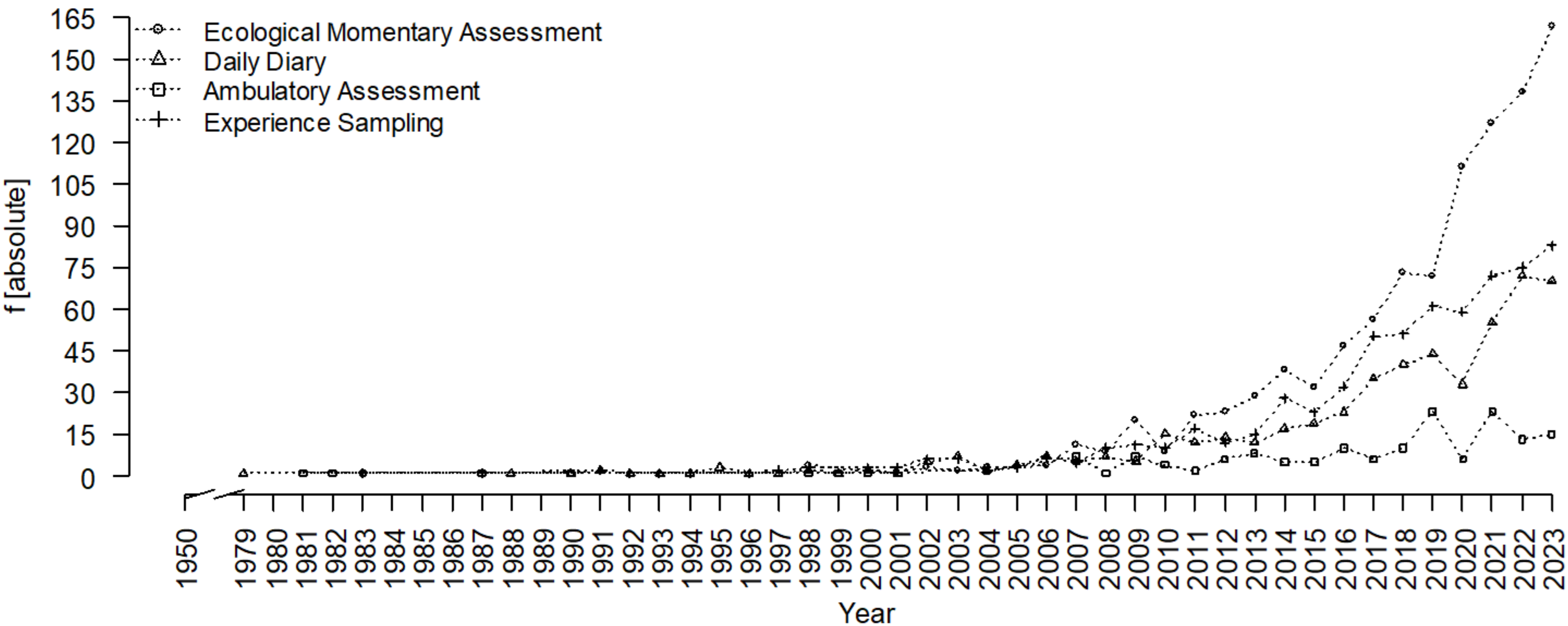
Krammer, G. (2024, July 7th). When we measure differently every day: a ML-SEM simulation study on within-person nonuniform measurement bias in intensive longitudinal data. https://doi.org/10.31219/osf.io/fm253

Krammer, G. (2024, August 13th). The Between-Not-Within fallacy coined and exemplified: why studying a within-person uniform measurement bias is driven by between-person differences in intensive longitudinal data. https://doi.org/10.31219/osf.io/7x8sg

Web of Science Search (June 26th 2024)

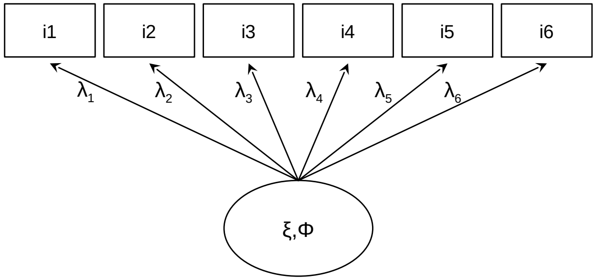
1950 → 2023

Title containing: “ecological momentary assessment”, “daily diary”, “ambulatory assessment” or “experience sampling”



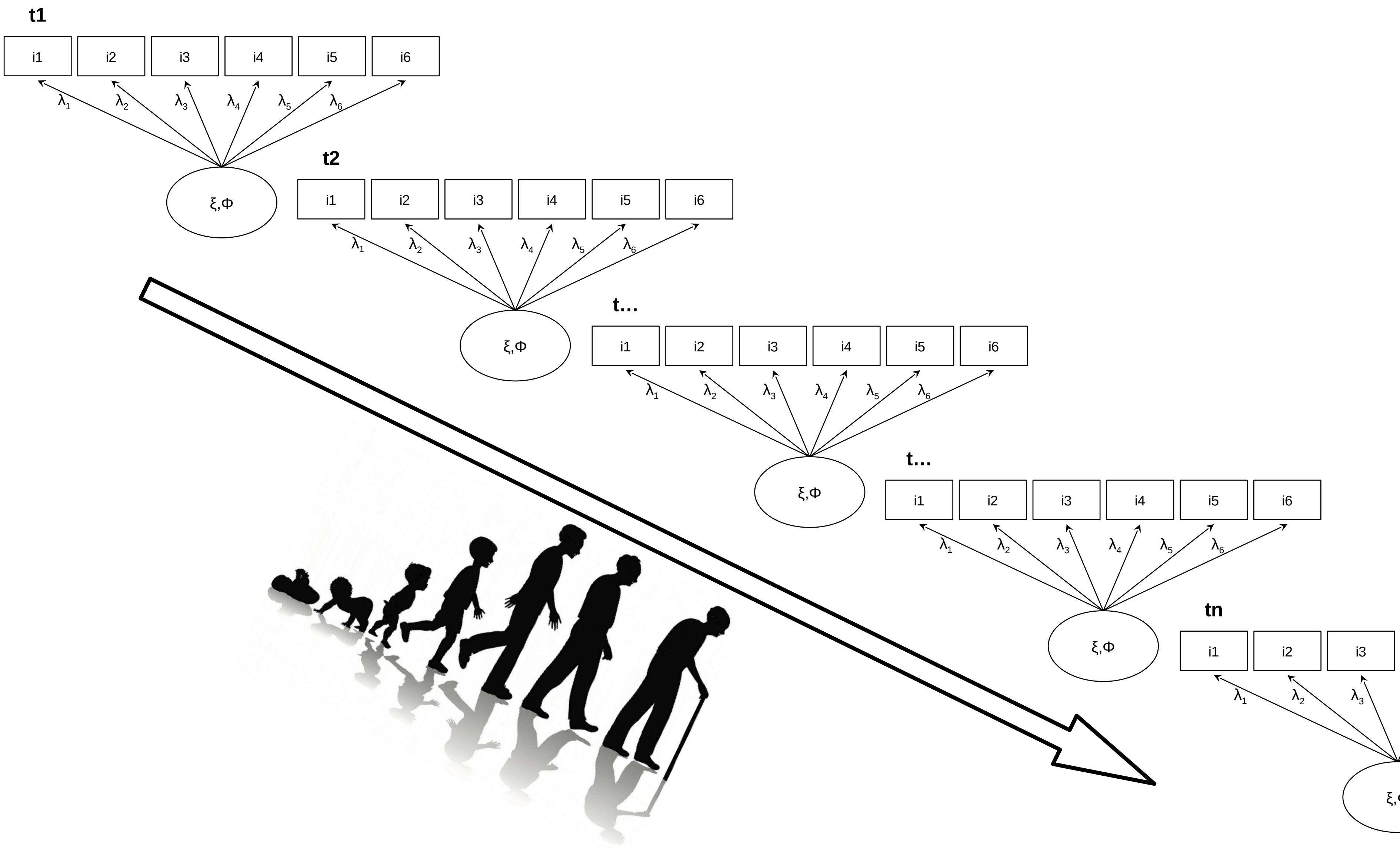
Intensive Longitudinal Data

Institute of
Business and
Vocational Education

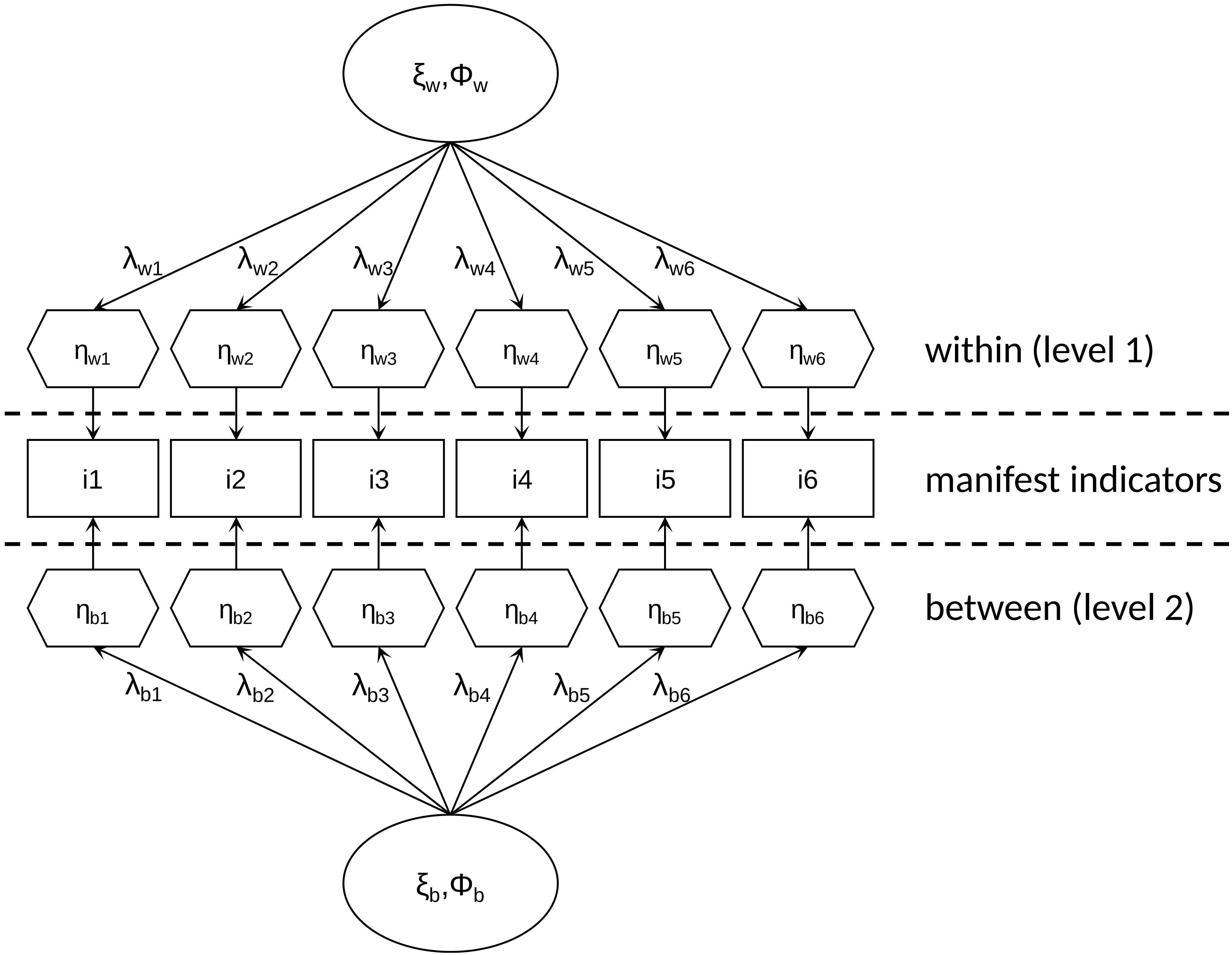


Examples with six items per scale:
Agentic/Neurotic narcissism via NGS & NVS (Crowe et al., 2016, 2018)
Grandiose/Vulnerable narcissism via SB-PNI (Pincus et al., 2009; Schoenleber et al., 2015)

Intensive Longitudinal Data



ML-SEM



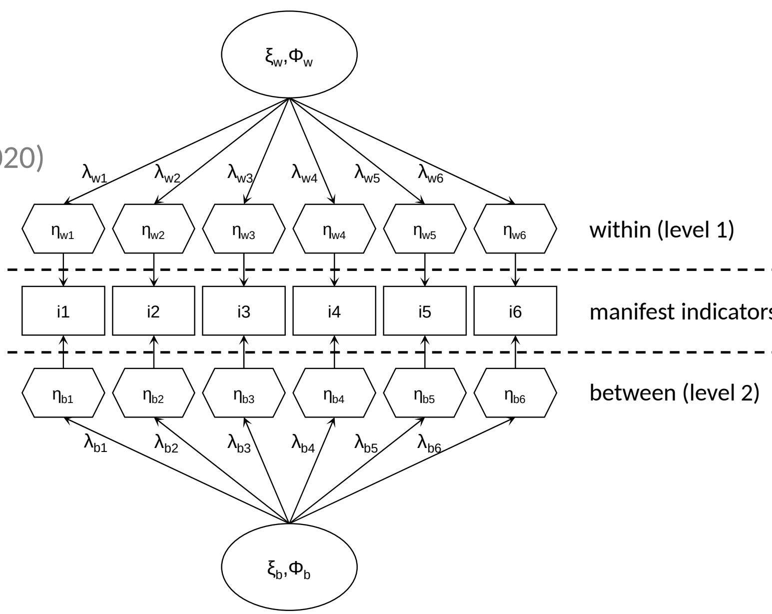
ξw,Φw
λw1
λw2
λw3
λw4
λw5
λw6
ηw1
ηw2
ηw3
ηw4
ηw5
ηw6
within (level 1)
i1
i2
i3
i4
i5
i6
manifest indicators
ηb1
ηb2
ηb3
ηb4
ηb5
ηb6
between (level 2)
λb1
λb2
λb3
λb4
λb5
λb6
ξb,Φb

ML-SEM

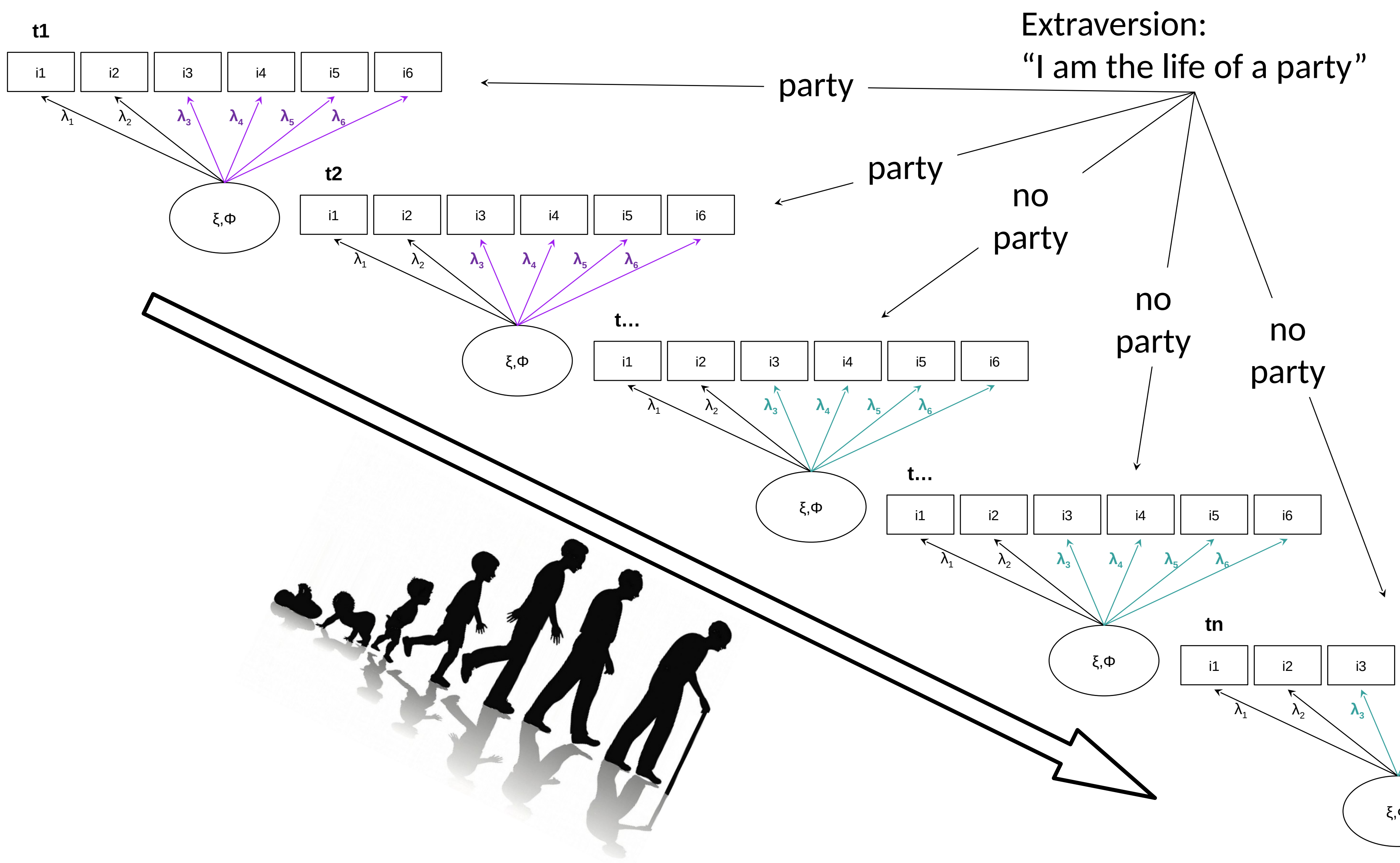
Multilevel structural equation models

(Lüdtke et al., 2007; Mehta & Neale, 2005; Muthén & Satorra, 1995; Stapleton, 2013)

- Intensive longitudinal data → two-level data structure:
  the respondents are the nesting factor
- SEM have a long tradition of testing psychometrical soundness multiple-item questionnaires.
- Emerging reviews show a lack of studies reporting psychometric properties in intensive longitudinal data.
  - h For example, ambulatory assessment: only 30% of the surveyed studies report psychometric properties and origin of items/scales. (Trull & Ebner-Priemer, 2020)

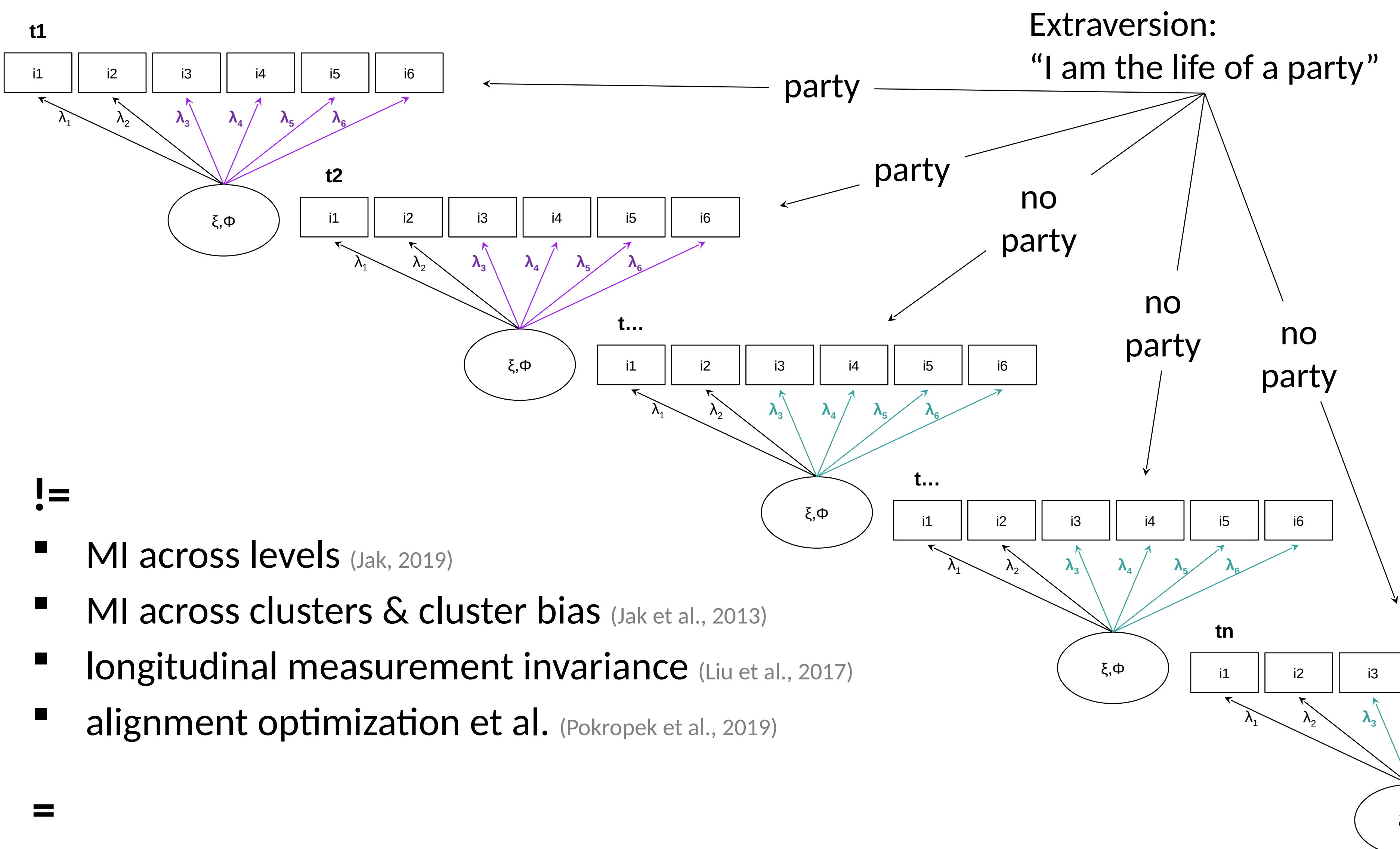


Measurement Bias



Measurement Bias

Institute of Business and Vocational Education



!=

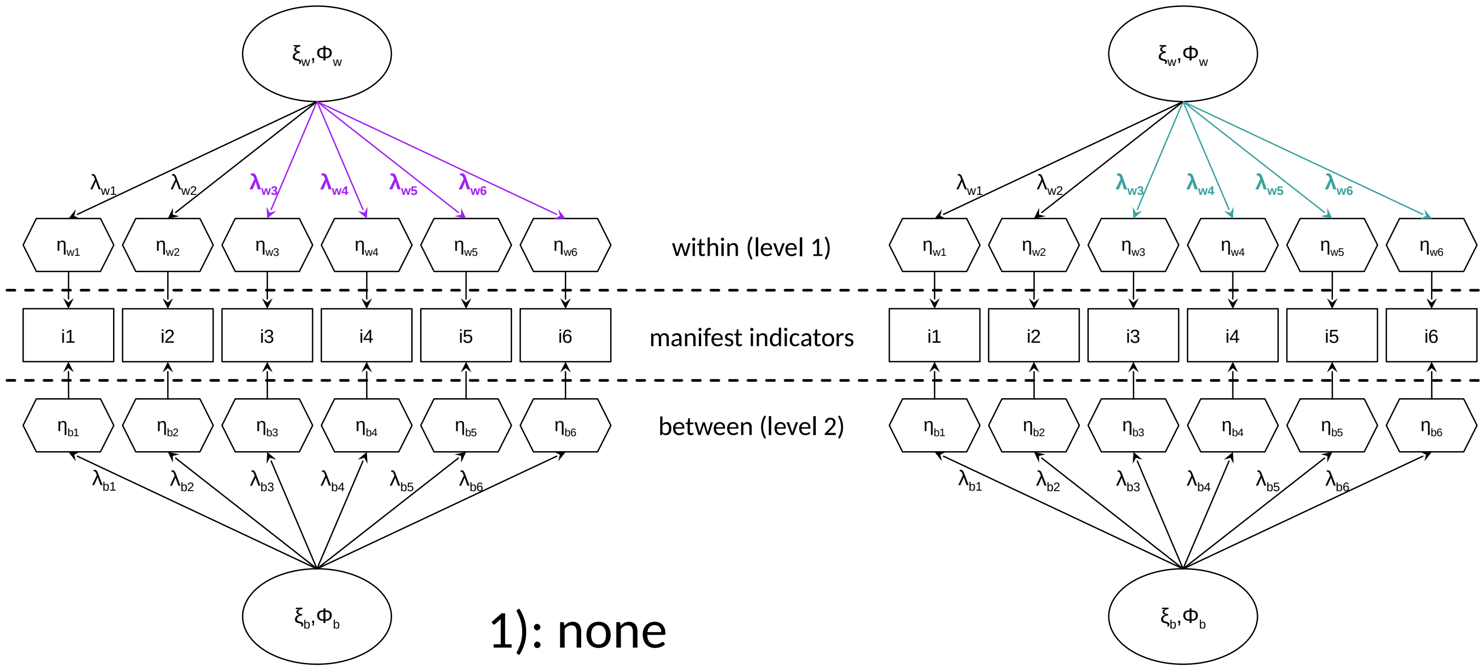
- MI across levels (Jak, 2019)
- MI across clusters & cluster bias (Jak et al., 2013)
- longitudinal measurement invariance (Liu et al., 2017)
- alignment optimization et al. (Pokropek et al., 2019)

=

- Within-person nonuniform measurement bias

Simulation study

5 qualitatively different types of nonuniform measurement bias



1): none

2): 1 item biased

3) and 4): 2 items biased

5): 4 items biased

Simulation study

Institute of Business and Vocational Education

- 3 x 5 x 3 x 5 x 2 = 450 conditions
  - 3 sample sizes with n ∈ {50, 100, 200}
  - 5 numbers of re-testing per subject with t ∈ {10, 20, 30, 50, 80}
  - 3 ICCs: ξ_b with $M = 0$ and $SD_b \in \{1, 2, 3\}$
  - 5 qualitatively different types of nonuniform measurement bias
  - 2 strengths of nonuniform measurement bias (low, high): $\Delta\lambda_{wi} \in \{.3, .5\}$
- 1000 data sets each in R (R Core Team, 2023b)
  - packages: *lavaan* (Rosseel, 2012) *psych* (Revelle, 2019) *multilevel* (Bliese, 2022) *parallel* (R Core Team, 2023a) *doParallel* (Corporation & Weston, 2022)
  - Response format: visual analogue scale
    (Jauk, Blum, et al., 2023; Jauk, Olaru, et al., 2023; Maliske et al., 2023)
- $\lambda_{wi} = .7$
  - cf. ML-SEM in the literature
    (Kim et al. (2016): found level 1 factor loadings with an average range of 0.41 - 0.83)
  - cf. prior simulation studies
    (Hsu et al., 2015; Kim & Cao, 2015)
  - realistic value for multiple-item questionnaires in intensive longitudinal data
    (cf. Study 1 and Study 3b in Rogoza et al., 2024)
  - leaves ample room for varying it across time points of measurement to introduce within-person nonuniform measurement bias

Simulation study

For each data set…

- … items' ICC
- … fit ML-SEM
- … χ^2-statistic: $p < .05$
- … CFI: .99, .95, .90
- … RMSEA: .10, .08, .06
- … SRMR-b: .08, .11
- … SRMR-w: .08, .11

Guiding principals:

- varying suggestion for cut-offs
  (Byrne, 2013; Hu & Bentler, 1999; Marsh et al., 2004; Schermelleh-Engel et al., 2003)
- single-level SEM as guideline
  (guidelines for evaluating ML-SEM fit are predominantly based on the single-level SEM: Kim et al., 2016)
- computing ML-SEM fit indices can be ambiguous: what is the sample size
  (Mehta & Neale, 2005).

Results: Summary

Institute of Business and Vocational Education

Type I error: χ^2-statistic <.5% & fit indices even with strictest cut-offs < .6%

Power: (adequate power: ≥.80, at least medium ICCs, at least 2 biased items)

ICC	medium		high	
to detect a…	…low strengths bias	…high strengths bias	…low strengths bias	…high strengths bias
$p < .05$	n = 100, each 50 times n = 200, each 30 times	n = 50, each 20 times	n = 50, each 30 times n = 100, each 20 times n = 200, each 10 times	n = 50, each 10 times
CFI ≥ .99	underpowered	n = 50, each 20 times	n = 50, each 80 times n = 100, each 50 times n = 200, each 30 times	n = 50, each 10 times
CFI ≥ .95	underpowered	underpowered	underpowered	n = 50, each 30 times n = 100, each 20 times
CFI ≥ .90	underpowered	underpowered	underpowered	n = 200, each 80 times
RMSEA < .06	underpowered	underpowered	underpowered	n = 50, each 20 times
RMSEA < .08	underpowered	underpowered	underpowered	n = 200, each 30 times
RMSEA < .10	underpowered	underpowered	underpowered	underpowered
SRMR-between < .08	underpowered	underpowered	underpowered	underpowered
SRMR-between < .11	underpowered	underpowered	underpowered	underpowered
SRMR-within < .08	underpowered	underpowered	underpowered	underpowered
SRMR-within < .11	underpowered	underpowered	underpowered	underpowered

Side note on SRMR-w: in certain conditions power decreased with larger data sets.

(similar, Marsh et al. (2004): in certain conditions single-level SRMR less power with higher sample sizes)

Take-Home-Message

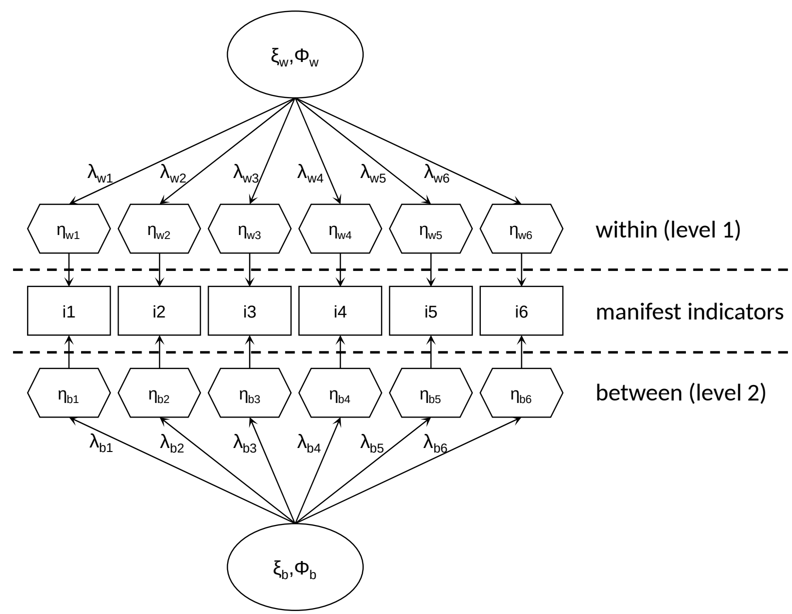
Institute of Business and Vocational Education

- ML-SEM fares very well in assessing psychometric properties of multiple-item questionnaires in intensive longitudinal data.

- Type I error: very good – too good?
  - still, don‘t ignore χ^2-statistic (Greiff & Heene, 2017)

- Power for detecting within-person nonuniform measurement bias:
  - χ^2-statistic +
  - CFI +
  - RMSEA + -
  - SRMRs -

- When using short scales in intensive longitudinal data:
  Please check psychometric properties!



Measurement bias in intensive longitudinal data

The longer (preprinted) reads:

Krammer, G. (2024, July 7th). When we measure differently every day: a ML-SEM simulation study on within-person nonuniform measurement bias in intensive longitudinal data. https://doi.org/10.31219/osf.io/fm253

Krammer, G. (2024, August 13th). The Between-Not-Within fallacy coined and exemplified: why studying a within-person uniform measurement bias is driven by between-person differences in intensive longitudinal data. https://doi.org/10.31219/osf.io/7x8sg

PD Dr. Georg Krammer
Senior Researcher @ Johannes Kepler University Linz, Institute of Business and Vocational Education

Leave of absence @ UCTE Styria as University College Professor for Educational Measurement and Applied Psychometrics

venia docendi: Educational Science

Co-editor of the *Zeitschrift für Bildungsforschung [Journal of Educational Research]*

Mail: georg.krammer@jku.at or find me on: RG | OSF | GoogleScholar | Mastodon

References 1/2

Bliese, P. (2022). *multilevel: Multilevel Functions*. https://cran.r-project.org/package=multilevel

Byrne, B. M. (2013). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. routledge.

Corporation, M., & Weston, S. (2022). *doParallel: Foreach Parallel Adaptor for the "parallel" Package*. https://cran.r-project.org/package=doParallel

Crowe, M. L., Carter, N. T., Campbell, W. K., & Miller, J. D. (2016). Validation of the Narcissistic Grandiosity Scale and creation of reduced item variants. *Psychological Assessment, 28*, 1550–1560. https://doi.org/10.1037/pas0000281

Crowe, M. L., Edershile, E. A., Wright, A. G. C., Campbell, W. K., Lynam, D. R., & Miller, J. D. (2018). Development and validation of the Narcissistic Vulnerability Scale: An adjective rating scale. *Psychological Assessment*, *30*, 978–983. https://doi.org/10.1037/pas0000578

Greiff, S., & Heene, M. (2017). Why psychological assessment needs to start worrying about model fit. *European Journal of Psychological Assessment*, *33*(5), 313–317. https://doi.org/10.1027/1015-5759/a000450

Hsu, H. Y., Kwok, O. man, Lin, J. H., & Acosta, S. (2015). Detecting Misspecified Multilevel Structural Equation Models with Common Fit Indices: A Monte Carlo Study. *Multivariate Behavioral Research*, *50*(2), 197–215. https://doi.org/10.1080/00273171.2014.977429

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jak, S. (2019). Cross-Level Invariance in Multilevel Factor Models. *Structural Equation Modeling*, *26*(4), 607–622. https://doi.org/10.1080/10705511.2018.1534205

Jak, S., Oort, F. J., & Dolan, C. V. (2013). A Test for Cluster Bias: Detecting Violations of Measurement Invariance Across Clusters in Multilevel Data. *Structural Equation Modeling*, *20*(2), 265–282. https://doi.org/10.1080/10705511.2013.769392

Jauk, E., Blum, C., Hildebrandt, M., Lehmann, K., Maliske, L., & Kanske, P. (2023). Psychological and neural correlates of social affect and cognition in narcissism: A multimethod study of self-reported traits, experiential states, and behavioral and brain indicators. *Personality Disorders: Theory, Research, and Treatment*.

Jauk, E., Olaru, G., Schürch, E., Back, M. D., & Morf, C. C. (2023). Validation of the German Five-Factor Narcissism Inventory and construction of a brief form using ant colony optimization. *Assessment*, *30*(4), 969–997.

Kim, E. S., & Cao, C. (2015). Testing Group Mean Differences of Latent Variables in Multilevel Data Using Multiple-Group Multilevel CFA and Multilevel MIMIC Modeling. *Multivariate Behavioral Research*, *50*(4), 436–456. https://doi.org/10.1080/00273171.2015.1021447

Kim, E. S., Dedrick, R. F., Cao, C., & Ferron, J. M. (2016). Multilevel Factor Analysis: Reporting Guidelines and a Review of Reporting Practices. *Multivariate Behavioral Research*, *51*(6), 881–898. https://doi.org/10.1080/00273171.2016.1228042

Liu, Y., Millsap, R. E., West, S. G., Tein, J.-Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological Methods*, *22*(3), 486–506.

Lüdtke, O., Trautwein, U., Schnyder, I., & Niggli, A. (2007). Simultane Analysen auf Schüler- und Klassenebene. *Zeitschrift Fur Entwicklungspsychologie Und Padagogische Psychologie*, *39*(1), 1–11. https://doi.org/10.1026/0049-8637.39.1.1

References 2/2

Maliske, L., Lehmann, K., Schurz, M., Hildebrandt, M., Jauk, E., & Kanske, P. (2023). *To feel and think what others feel and think: Functional network reorganization underlies context-changes in naturalistic social cognition*. https://doi.org/10.31234/osf.io/c6gsz

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling*, *11*(3), 452–483. https://doi.org/10.1207/s15328007sem1103

Mehta, P. D., & Neale, M. C. (2005). People are variables too: multilevel structural equations modeling. *Psychological Methods*, *10*(3), 259–284. https://doi.org/10.1037/1082-989X.10.3.259

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*(1995), 267–316. http://www.jstor.org/stable/271070%5Cnhttp://statmodel2.com/download/SMMuthenSatorra1995.pdf

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N. (2009). Initial construction and validation of the Pathological Narcissism Inventory. *Psychological Assessment*, *21*, 365–379. https://doi.org/10.1037/a0016530

Pokropek, A., Davidov, E., & Schmidt, P. (2019). A Monte Carlo Simulation Study to Assess The Appropriateness of Traditional and Newer Approaches to Test for Measurement Invariance. *Structural Equation Modeling*, 1–21. https://doi.org/10.1080/10705511.2018.1561293

R Core Team. (2023a). *parallel*. https://www.r-project.org/

R Core Team. (2023b). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research*. https://cran.r-project.org/package=psych

Rogoza, R., Krammer, G., Jauk, E., Flakus, M., Baran, L., Di Sarno, M., Di Pierro, R., Zajenkowski, M., Dufner, M., & Fatfouta, R. (2024). The Peaks and Valleys of Narcissism: The Factor Structure of Narcissistic States and Their Relations to Trait Measures. *Psychological Assessment*, *36*(2), 147–161. https://doi.org/10.1037/pas0001295

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Schermelleh-Engel, K., Moosbrugger, H., Müller, H., & others. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schoenleber, M., Roche, M. J., Wetzel, E., Pincus, A. L., & Roberts, B. W. (2015). Development of a brief version of the Pathological Narcissism Inventory. *Psychological Assessment*, *27*, 1520–1526. https://doi.org/10.1037/pas0000158

Stapleton, L. M. (2013). Multilevel structural equation modeling with complex sample data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural Equation Modeling: A Second Course* (pp. 521–562). Information Age Publishing Inc.

Trull, T. J., & Ebner-Priemer, U. W. (2020). Ambulatory Assessment in Psychopathology Research: A Review of Recommended Reporting Guidelines and Current Practices. *Journal of Abnormal Psychology*, *129*, 56–63. https://doi.org/10.1037/abn0000473