

PSYCHOMETRIC PROPERTIES OF THE DUAL-RANGE SLIDER RESPONSE FORMAT

Matthias Kloft

International Meeting of the Psychometric Society 2023

MOTIVATION

When **ONE** Response Value is **NOT** Enough

“What percentage of your daily work time did you spend on preparing for IMPS 2023 in the last week?”

INTERVAL RESPONSES

Single-range slider / visual analog scale (**VAS**)



Dual-range slider (**DRS**)

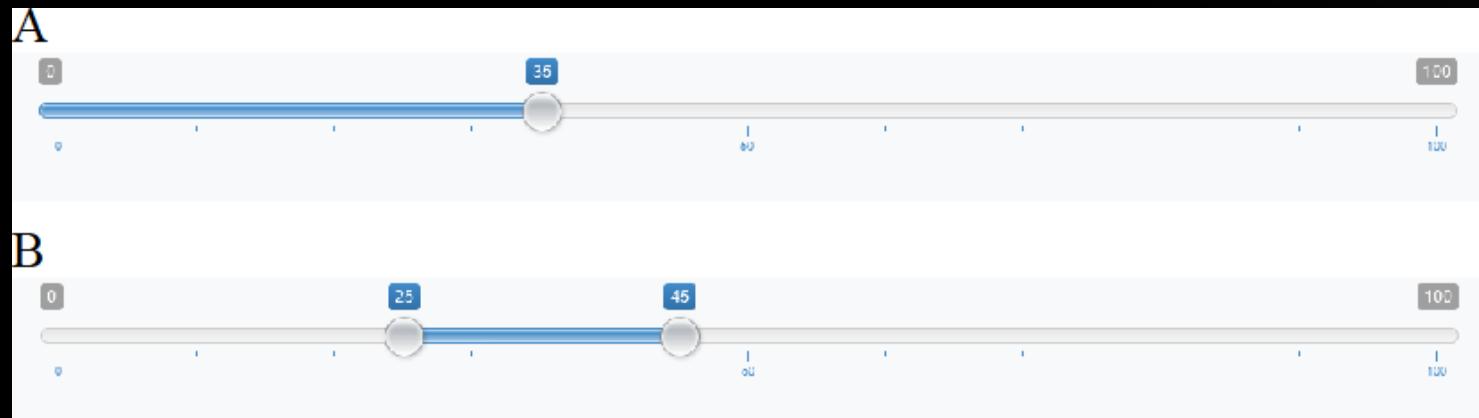


INTERVAL RESPONSES

Single-range slider / visual analog scale (**VAS**)

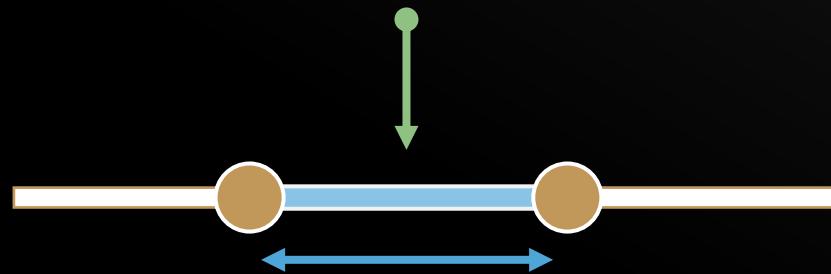


Dual-range slider (**DRS**)



INTERVAL RESPONSES

DRS Location: $\frac{y^{(L)}+y^{(U)}}{2}$



DRS Width: $y^{(U)} - y^{(L)}$

INTERVAL RESPONSES

Variability / Plausible Range:

- Self-ratings, stimuli

Uncertainty / Expertise:

- Estimation (forecasting, risks)

Ambiguity:

- Item content unclear
- No clear cut true answer (Verbal Quantifiers, e.g., **seldom**)

INTERVAL RESPONSES

DRS Location:

- Works **similar to VAS** (Kloft et al, in 2023a)

DRS Width:

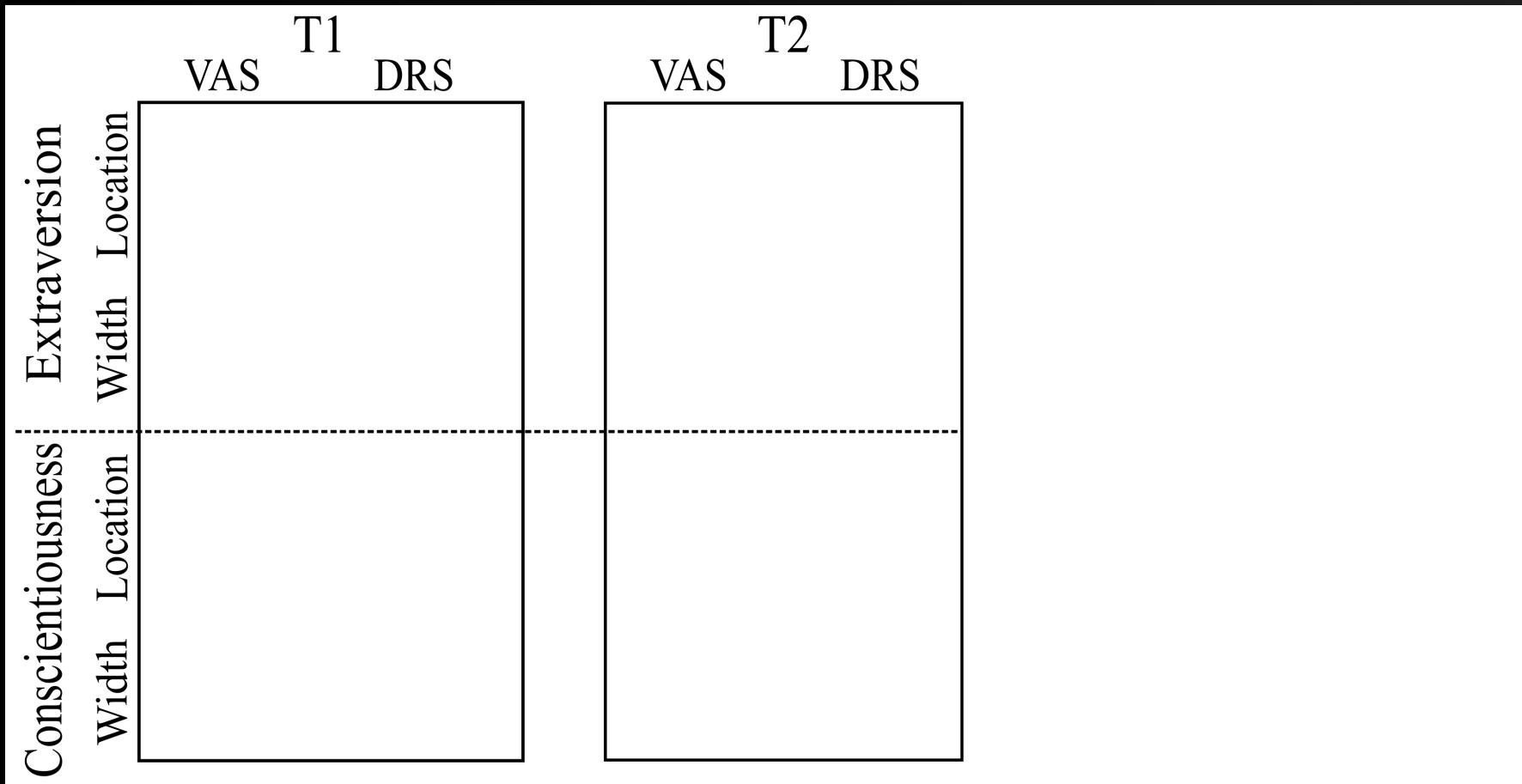
- Psychometric properties?
- Study 1: Longitudinal MTMT
- Study 2: Mutliple Applications, cross-sectional
 - Online surveys in student samples

STUDY 1

LONGITUDINAL MTMM

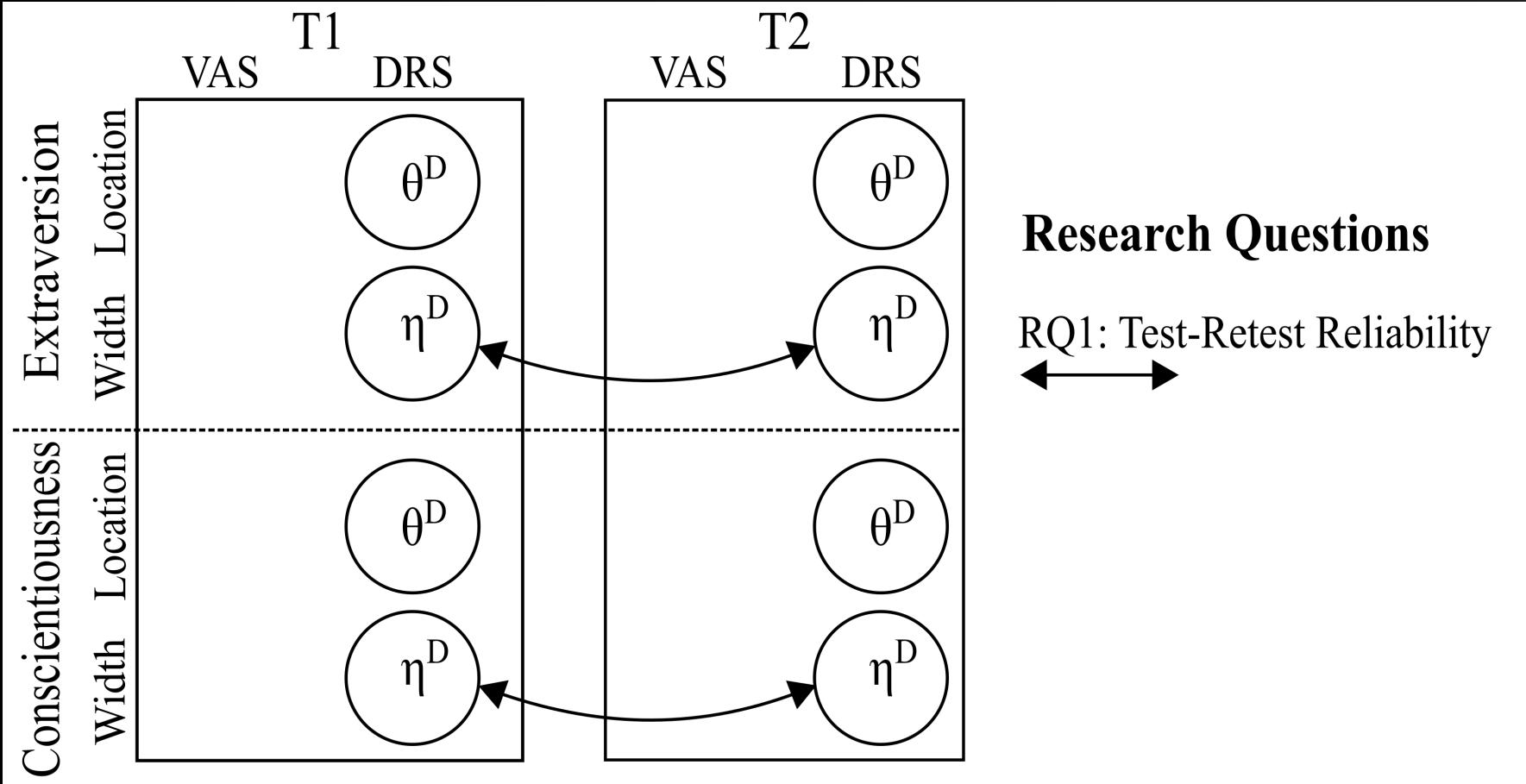
(Kloft et al., 2023b)

STUDY DESIGN



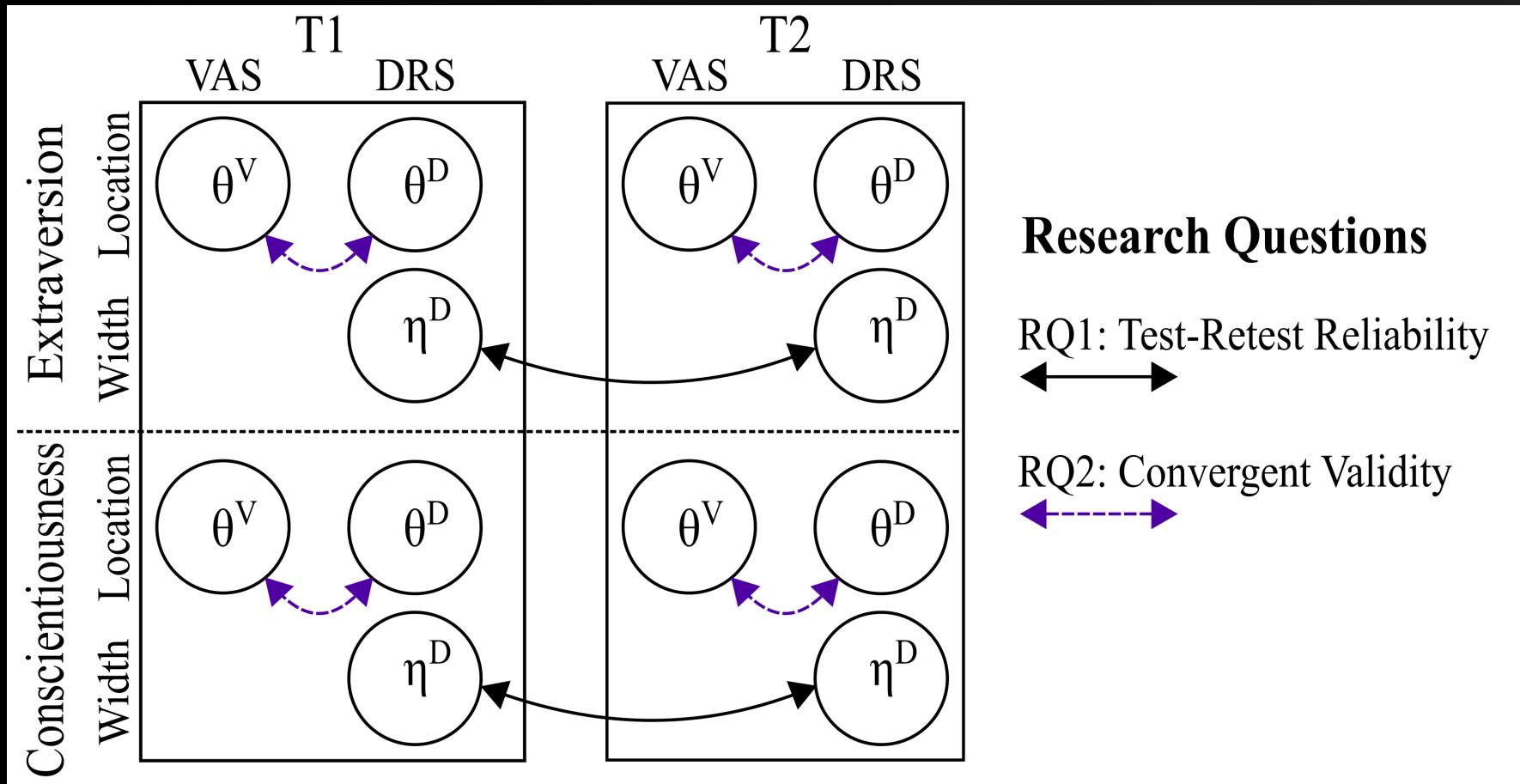
2 Timepoints, 2 Response Formats (VAS, DRS),
2 Scales (Extraversion & Conscientiousness)

RQ1: TEST-RETEST RELIABILITY



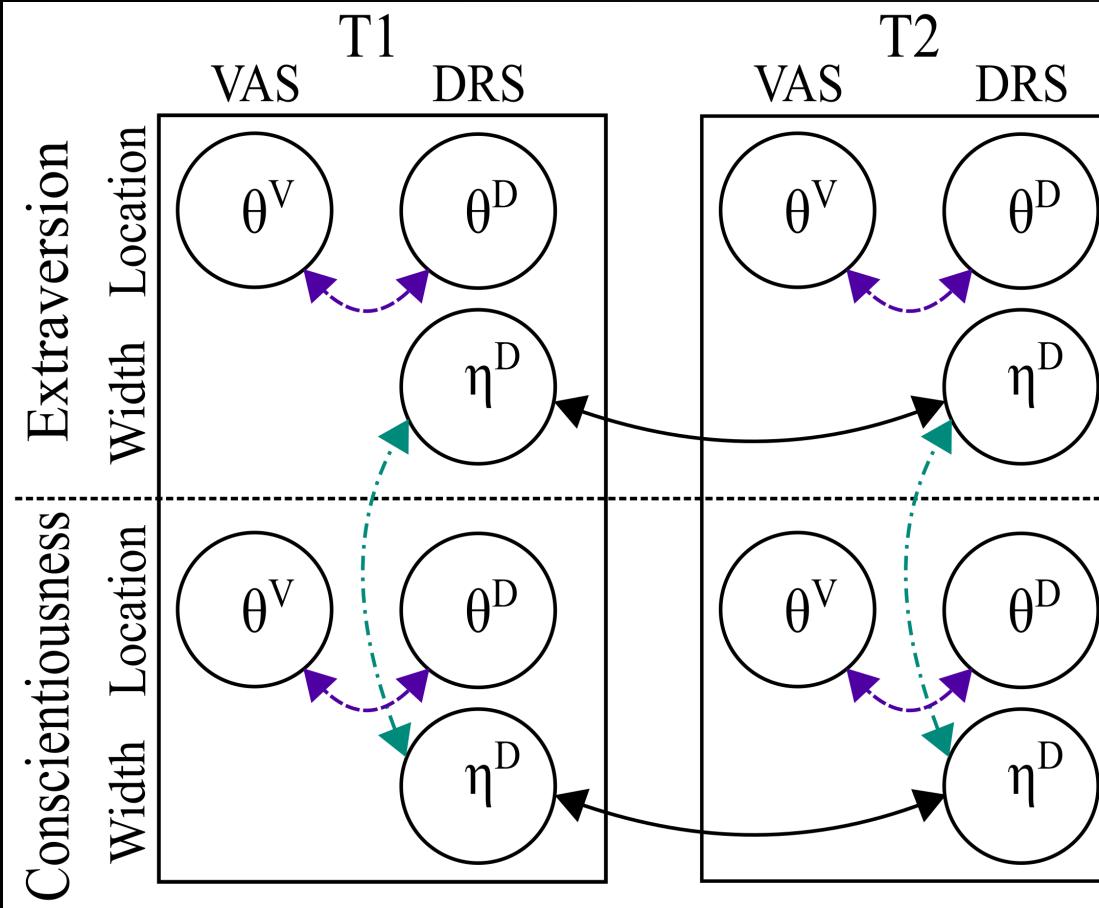
- **Correlations** of person scores
- Focus on DRS Width

RQ2: CONVERGENT VALIDITY



**Equivilancy of VAS & DRS Location:
Central Tendency**

RQ3: DISCRIMINANT VALIDITY



Research Questions

RQ1: Test-Retest Reliability
↔

RQ2: Convergent Validity
↔

RQ3: Discriminant Validity
↔

Different Variability for Different Traits?

IRT MODEL

VAS: Beta Response Model

- (Noel & Dauvier, 2007)

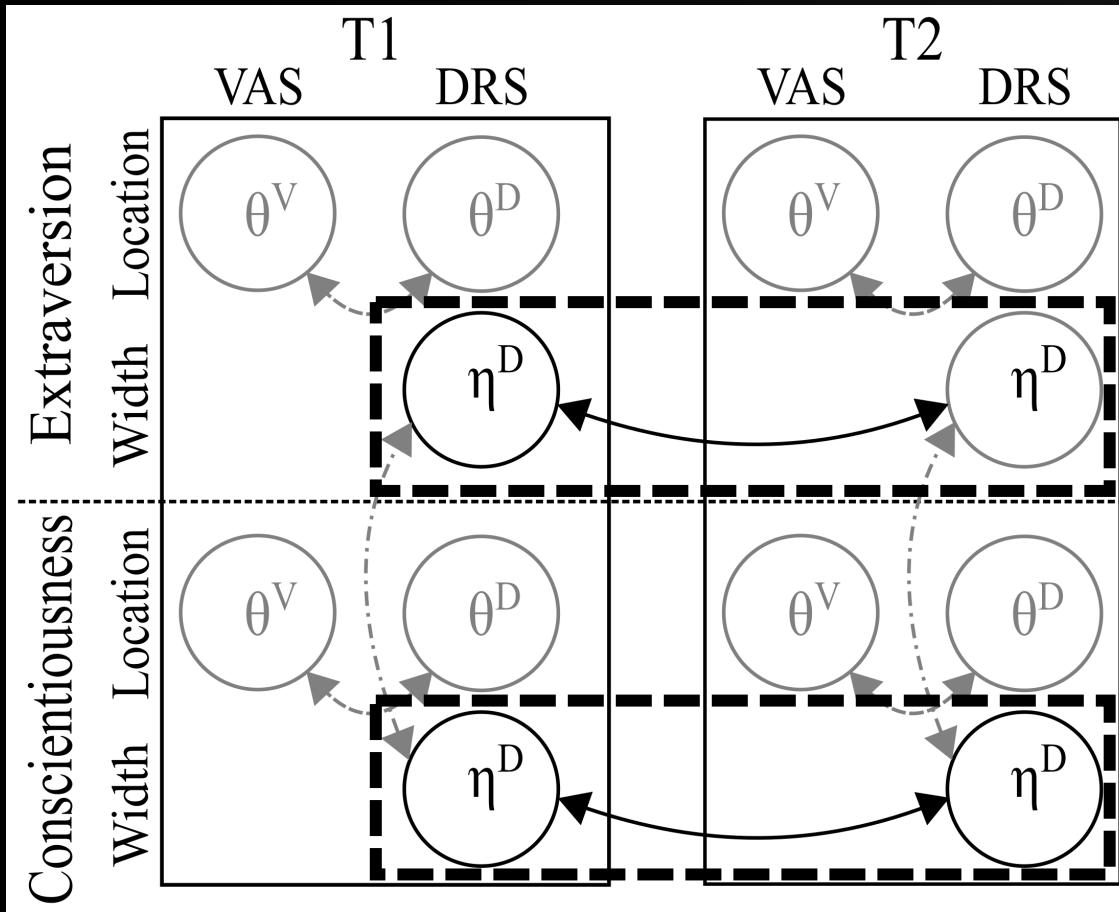
DRS: Dirichlet Dual Response Model

- (Kloft et al., 2023a)

Correlations of person scores:

- Bayesian hierarchical model with MVN prior

RQ1: TEST-RETEST RELIABILITY

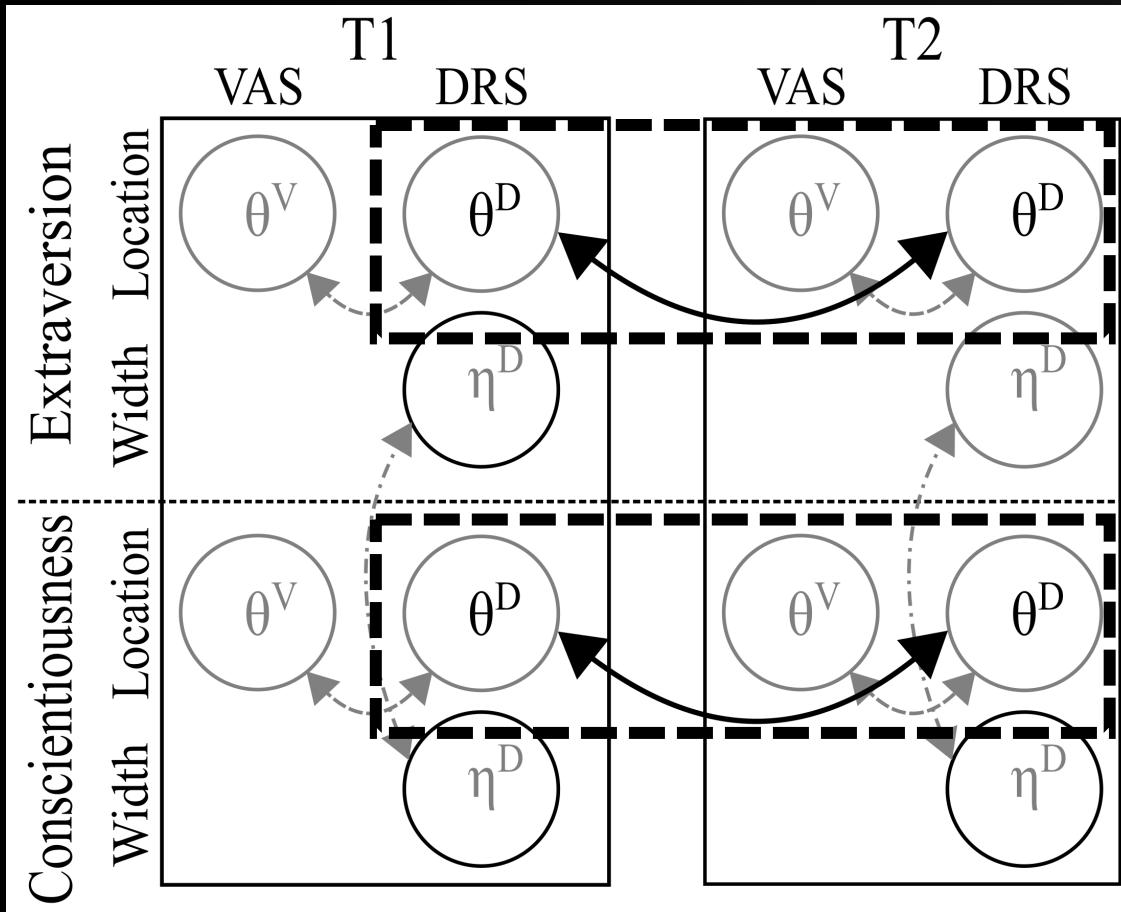


Width:
 $\hat{\rho} = .81$

$\hat{\rho} = .73$

Good reliability for DRS **Width** scores

RQ1: TEST-RETEST RELIABILITY



Location:
 $\hat{\rho} = .92$

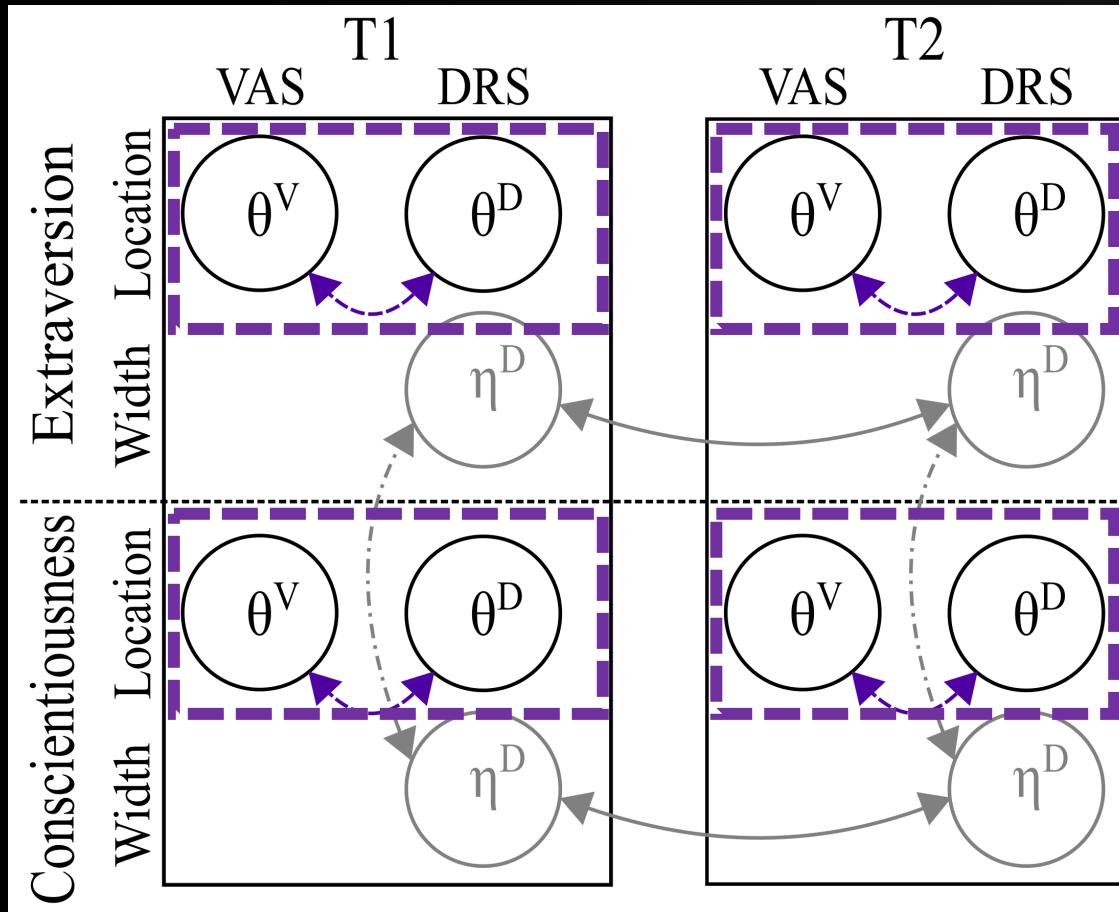
Width:
 $\hat{\rho} = .81$

$\hat{\rho} = .87$

$\hat{\rho} = .73$

High reliability for DRS Location scores

RQ2: CONVERGENT VALIDITY



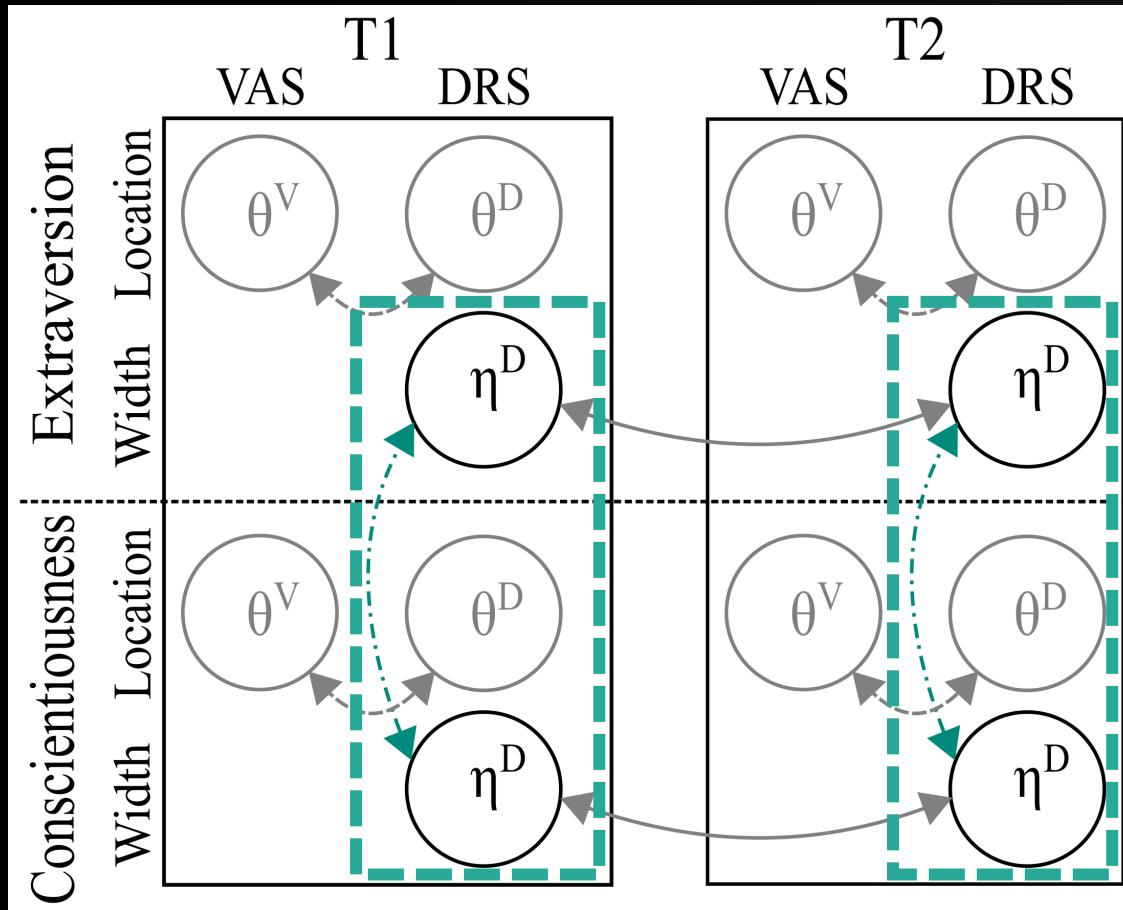
T1: T2:

$$\hat{\rho} = .93 \quad \hat{\rho} = .96$$

$$\hat{\rho} = .88 \quad \hat{\rho} = .90$$

VAS and DRS Location could be used **equivalently**

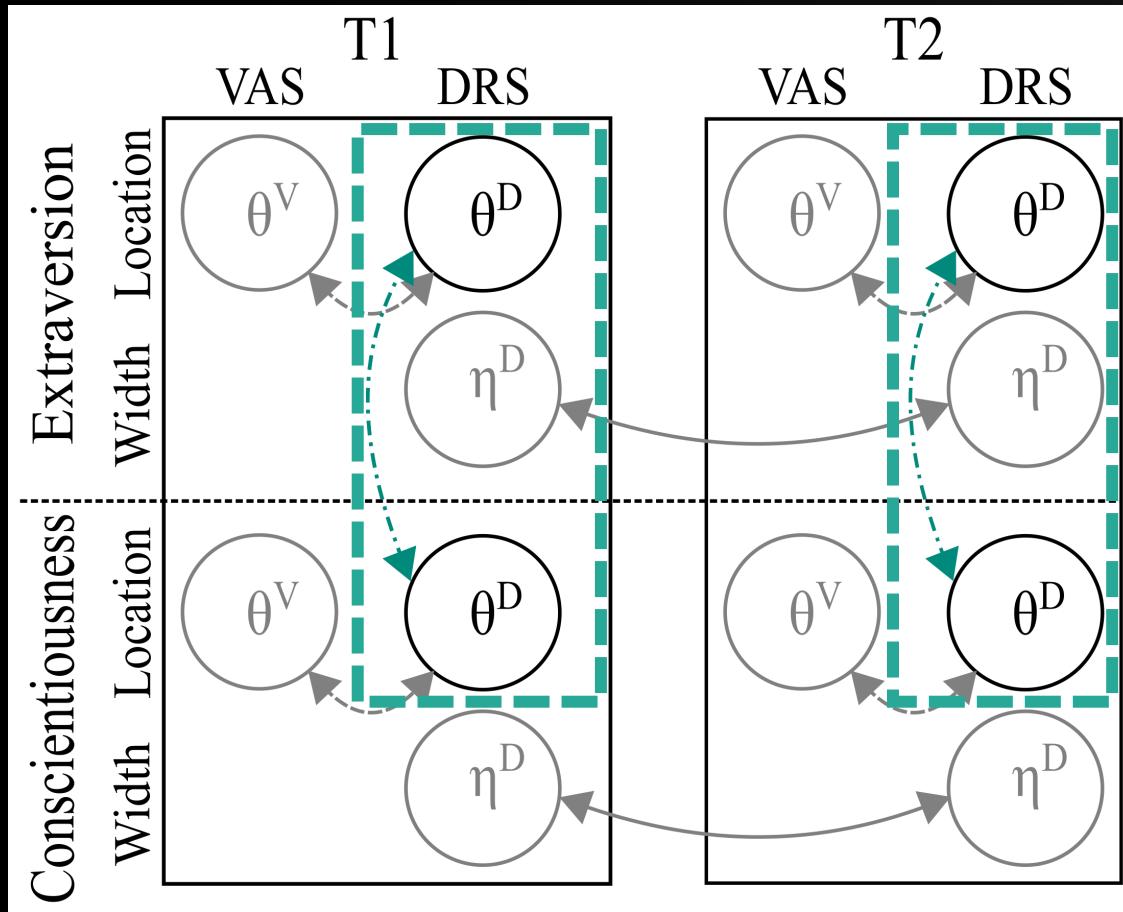
RQ3: DISCRIMINANT VALIDITY



Width:
 $\hat{\rho} = .94$ $\hat{\rho} = .96$

Poor discrimination of DRS Width scores

RQ3: DISCRIMINANT VALIDITY



Location:

$$\hat{\rho} = .30 \quad \hat{\rho} = .31$$

Width:

$$\hat{\rho} = .94 \quad \hat{\rho} = .96$$

Reasonable correlation of DRS **Location** scores

RQ3: DISCRIMINANT VALIDITY

What do DRS **Widths** measure?

- Personal **preference** for a particular width
 - Preference response style
- **Global trait** of perceived variability in personality?

STUDY 2: MULTIPLE APPLICATIONS

(Preliminary Results)

RESEARCH QUESTIONS

Focus on DRS **Widths**:

- **Sensitivity** to different tasks?
- Just a preference response style?
- **Dimensionality**?

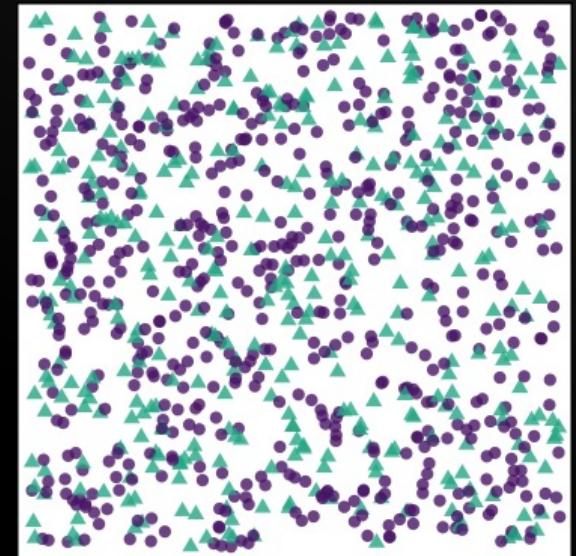
Factor Analyses: EFA, CFA

- Multivariate-logit transformation to account for boundedness and location

STUDY DESIGN: APPLICATIONS / TASKS

Extraversion & Conscientiousness:
Talkative = [75% – 93%] applicable

Color Dot Estimation:
% purple dots = [46% – 73%]
(true = 62%)



Election Forecasting: outcomes for 6 parties

- *Green party* = [10% – 19%]

STUDY DESIGN: APPLICATIONS / TASKS

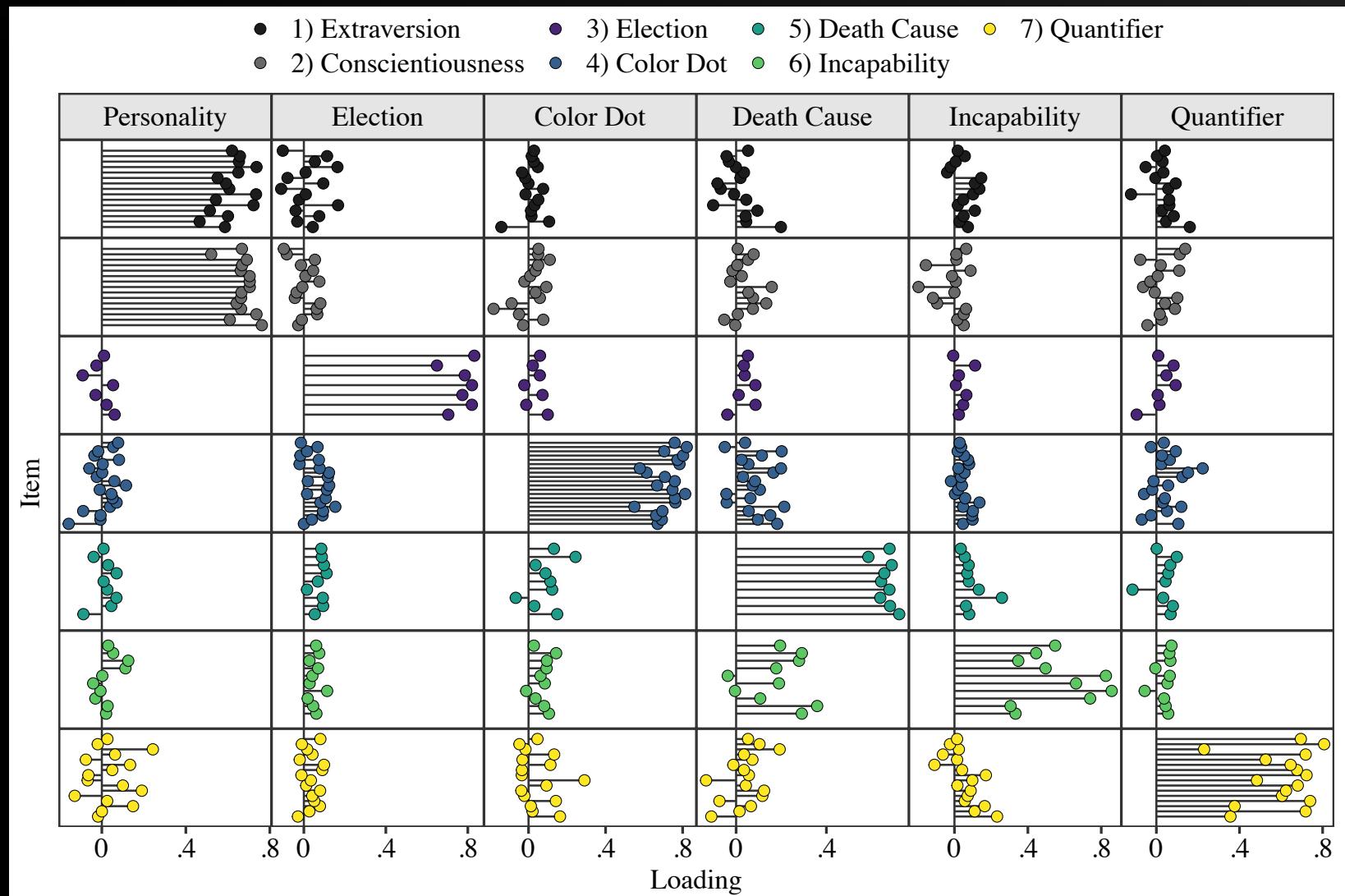
Estimation of percentages:

- Death Causes: *heart diseases*
- Reasons for Incapability for work:
mental health

Rating of verbal quantifiers as probabilities
that a so described event would occur:

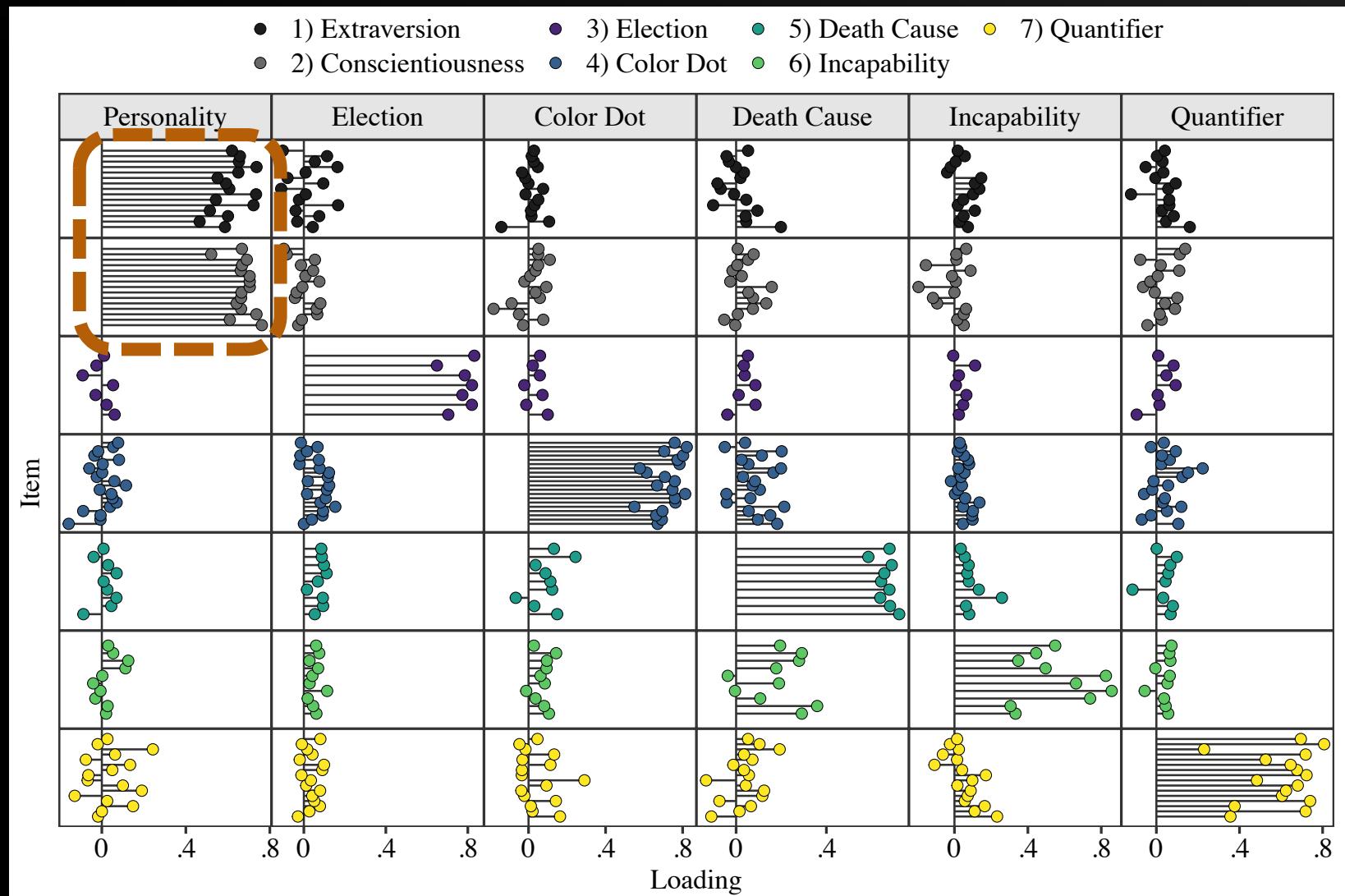
- *Seldom* = [5% – 30%]

INTERVAL WIDTH: EFA LOADINGS



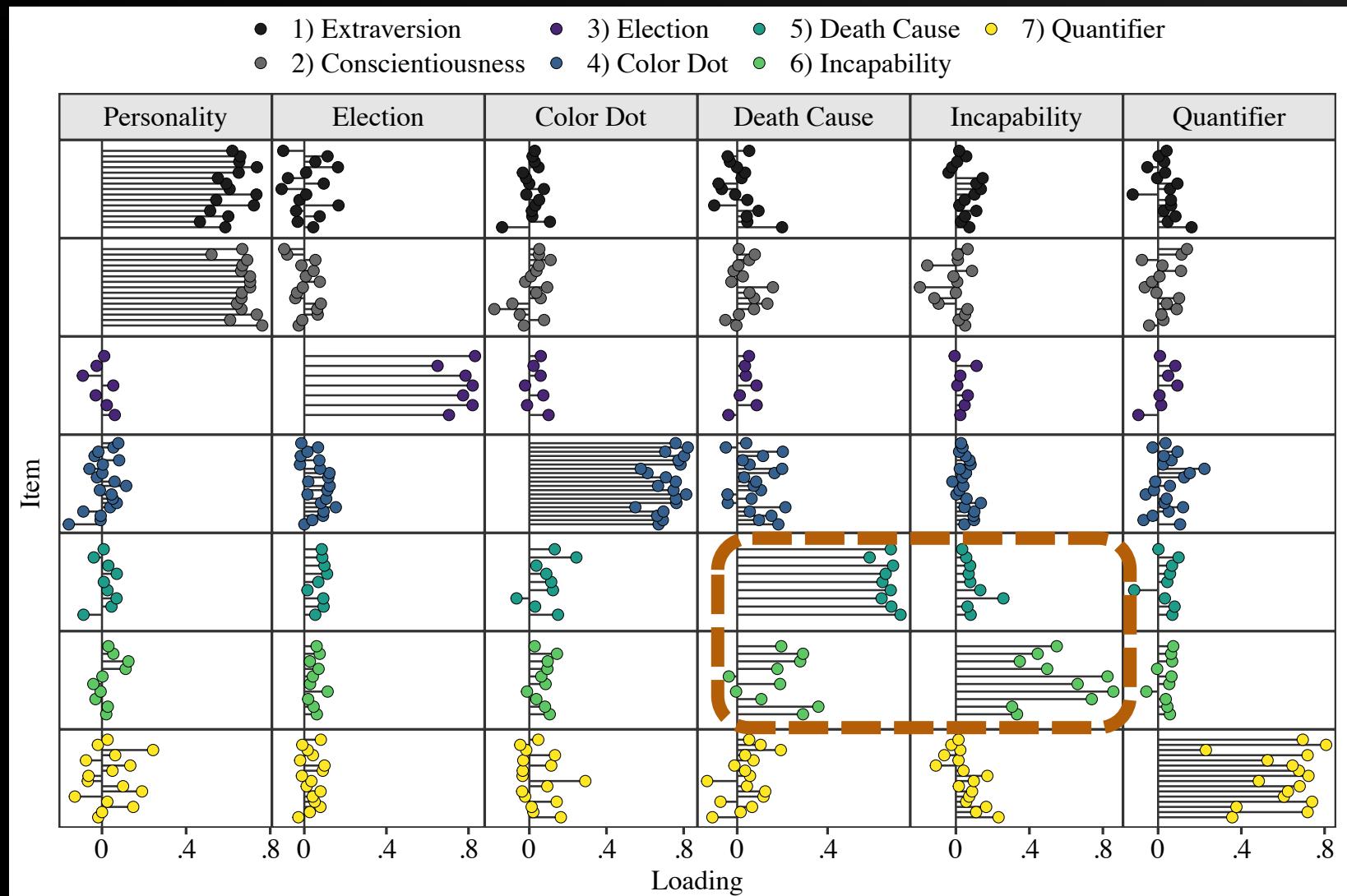
DRS Width only!

INTERVAL WIDTH: EFA LOADINGS



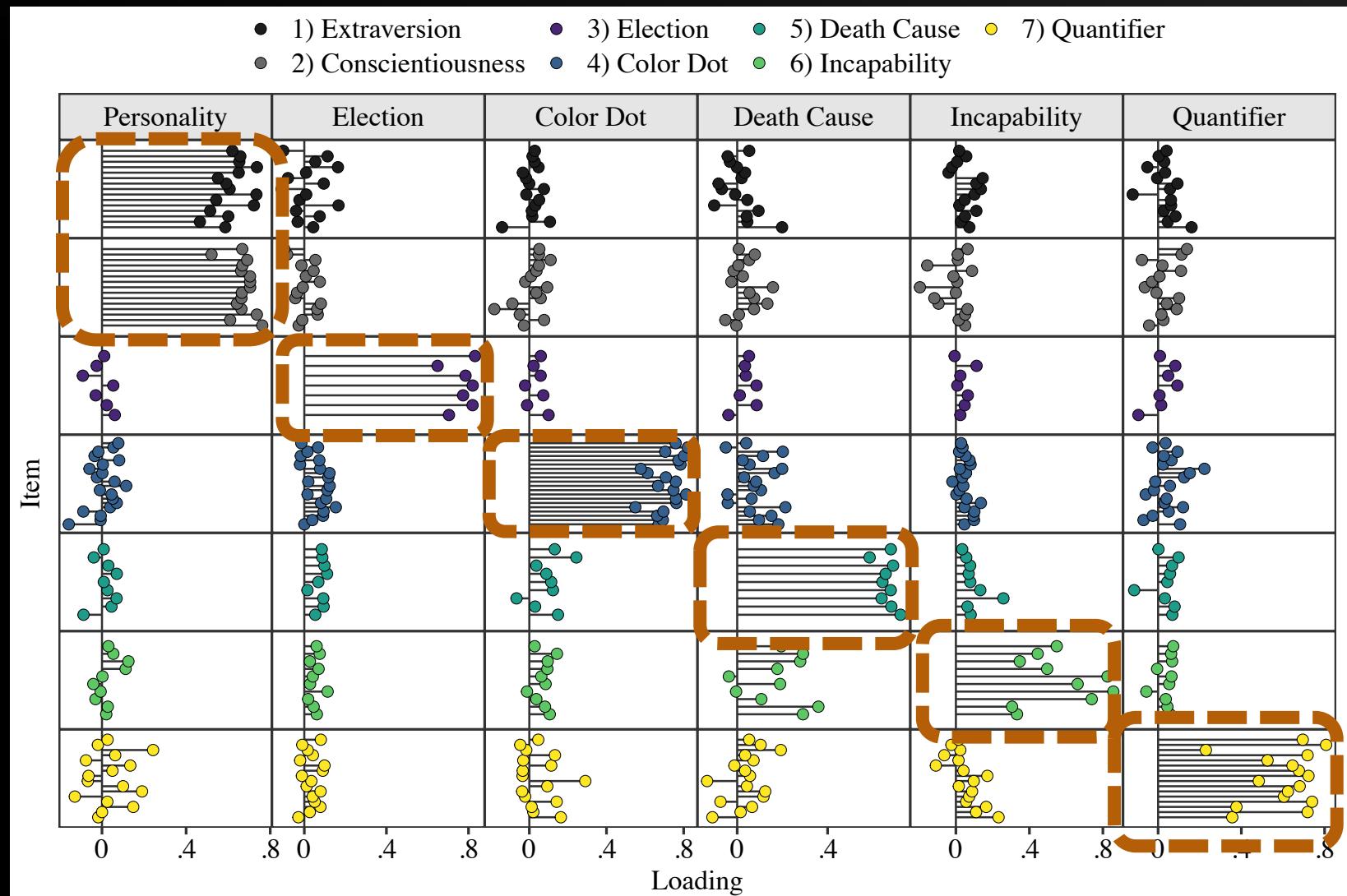
Common factor for Extraversion & Conscientiousness

INTERVAL WIDTH: EFA LOADINGS



High cross loadings: similar tasks

INTERVAL WIDTH: EFA LOADINGS

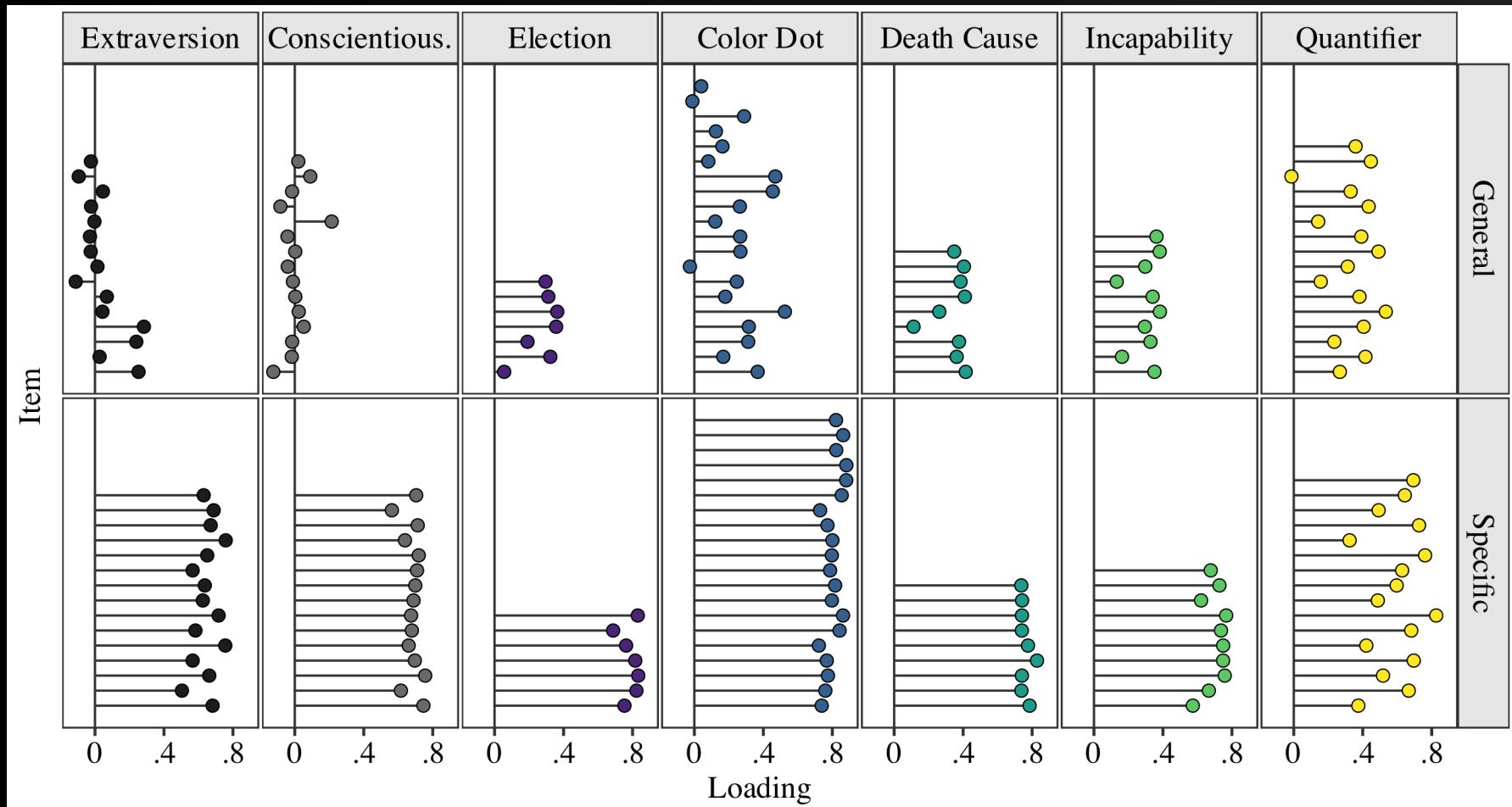


Factor structure **follows** task structure

CFA: BIFACTOR MODEL

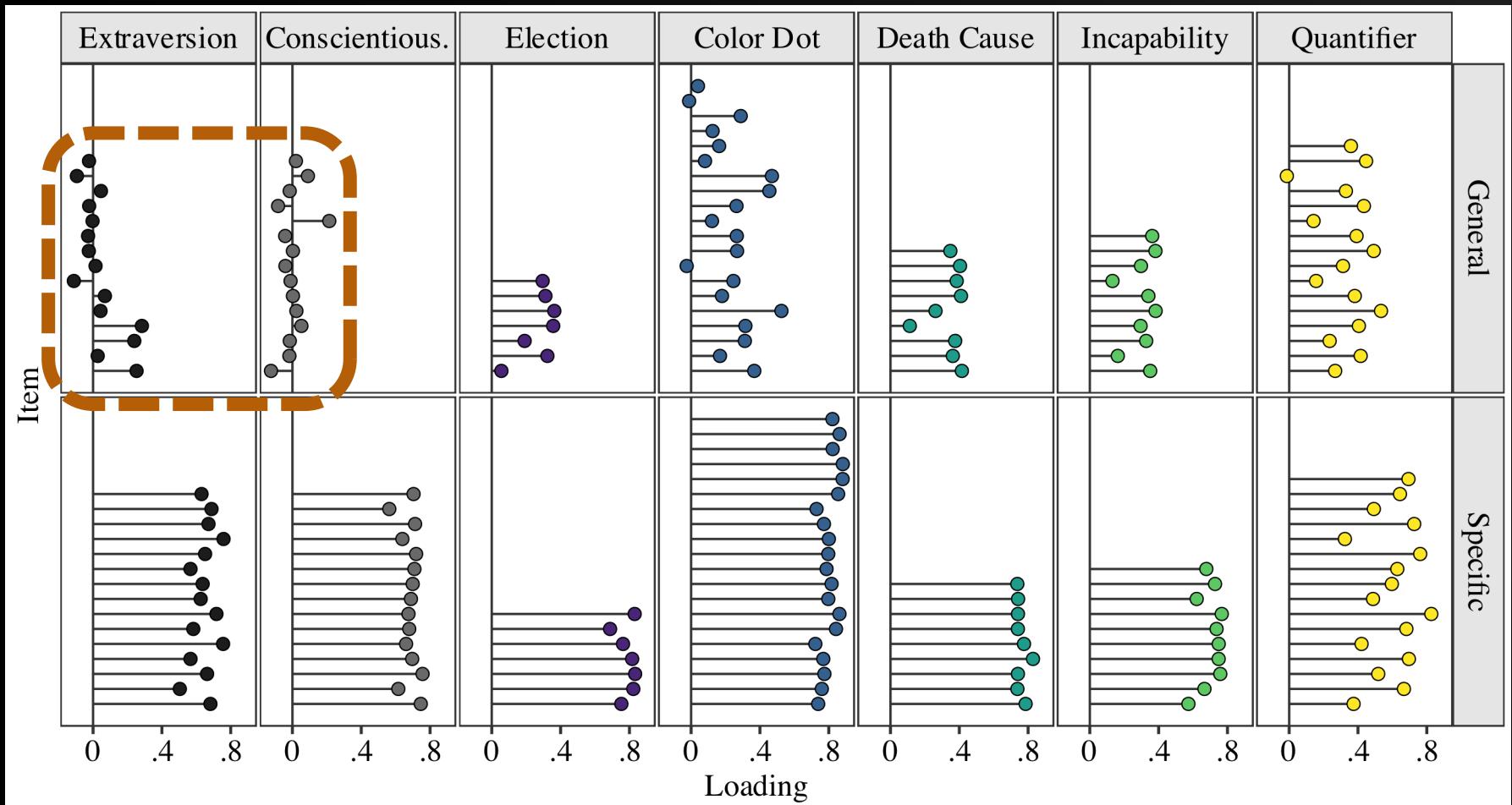
- Again only DRS **Width**
- 6 **Specific Factors (S-Factor)**:
 - Assignment by theoretical domain / task
- **General Factor (G-Factor)**: preferred width

LOADINGS: G-FACTOR VS. S-FACTOR



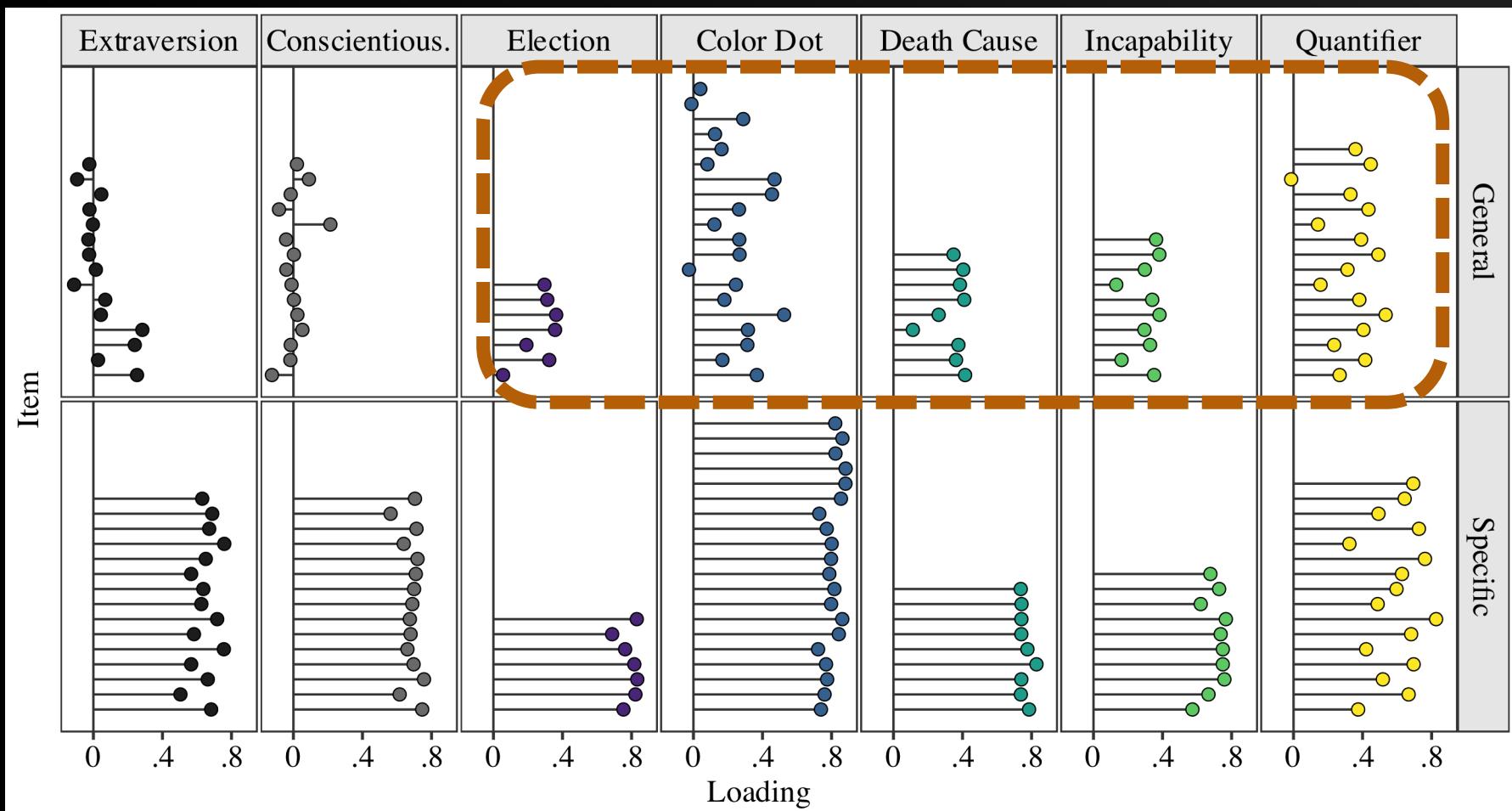
Top: G-Factor
 Bottom: S-Factor

LOADINGS: G-FACTOR VS. S-FACTOR



Personality: mostly small to zero loadings on General Factor

LOADINGS: G-FACTOR VS. S-FACTOR



Estimation tasks: mostly small to medium loadings on General Factor

TAKE HOME POINTS

- High **convergent validity** of response formats (VAS & DRS Location)
- Good **test-retest reliability** of DRS factor scores
- Poor **discriminant validity** of the DRS **Width** in personality measurement
- Good **discriminant validity** of the DRS **Width** regarding dissimilar tasks

THANKS TO:



- Prof. Dr. Daniel W. Heck



- Dr. Jean-Paul Snijder

Contact:

kloft@uni-marburg.de

Slides:

<https://github.com/matthiaskloft/>

REFERENCES

- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011–1027. APA PsycArticles. <https://doi.org/10.1037/0022-3514.80.6.1011>
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, 7(1), 7–28.
- Kloft, M., Hartmann, R., Voss, A., & Heck, D. W. (2023a). The Dirichlet dual response model: An item response model for continuous bounded interval responses. *Psychometrika*. <https://doi.org/10.1007/s11336-023-09924-7>
- Kloft, M., Snijder, J., & Heck, D. W. (2023b). *Measuring the variability of personality traits with interval responses: Psychometric properties of the dual-range slider response format*. PsyArXiv. <https://psyarxiv.com/pa4m3/>
- Noel, Y., & Dauvier, B. (2007). A beta item response model for continuous bounded responses. *Applied Psychological Measurement*, 31(1), 47–73. <https://doi.org/10.1177/0146621605287691>