

POSITION PAPER

ON THE (NON-)PRECISION OF
PSYCHOLOGICAL MEASURES IN
SECURITY AND PRIVACY RESEARCH

RELIABILITY BEYOND REPORTING
ALPHA

Nele Borgert¹, Tim Ulmann¹, and
Robin Merchel²

University of Bern, Institute of Psychology,
Psychology of Digitalisation, Switzerland

Ruhr University Bochum, Faculty of Psychology,
Psychological Assessment, Germany

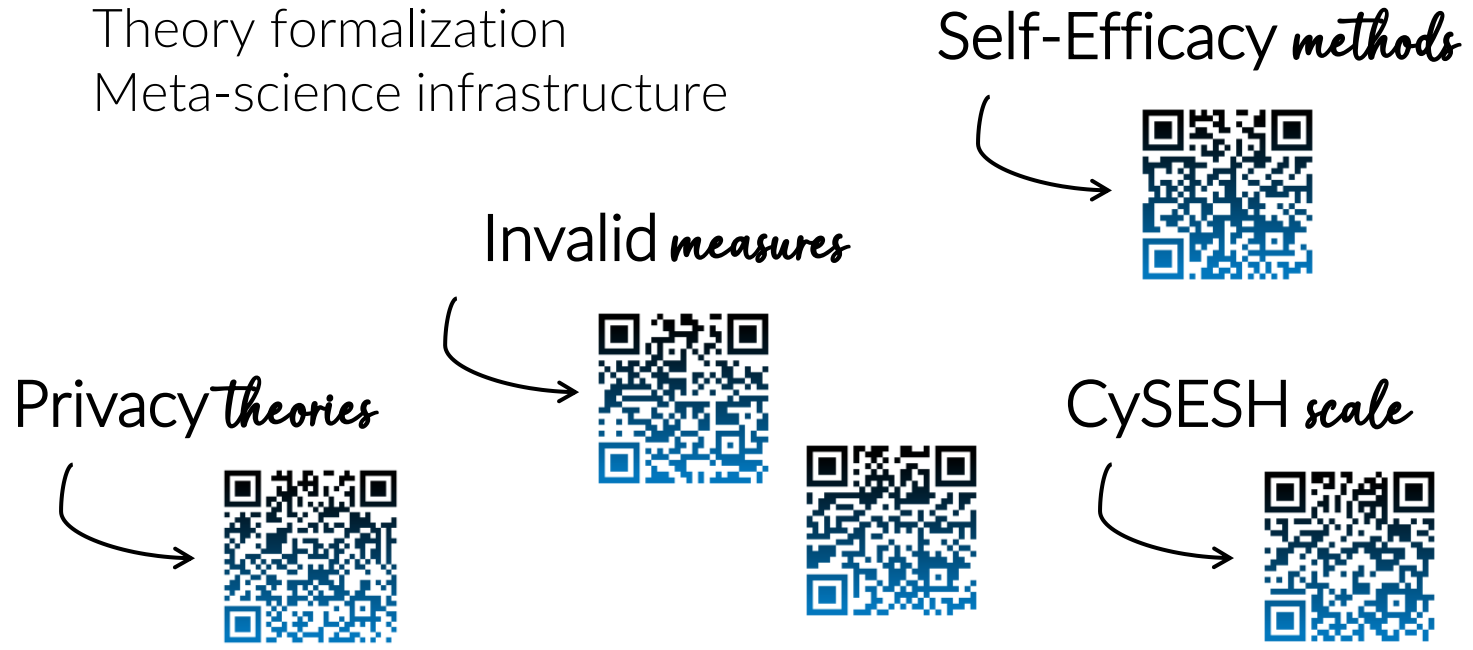


Why I'm Here.

I've rage-read one too many "reliable" scales that were duct-taped together the night before data collection and only report "alpha."

My Topics.

- Measurement reform
- Theory formalization
- Meta-science infrastructure

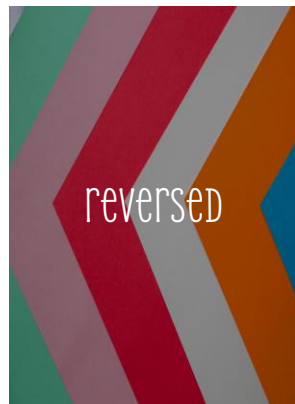


What's the Issue?

Imprecise measurement of latent psychological constructs (e.g., trust, risk taking, concerns).

Adds noise.

We may walk away claiming to know stuff when imprecise measures may have:



the very relationships we think we found.



How Can Measures Be Imprecise?



Reliability = extent to which scales measure the attribute without error

$$\frac{\chi}{\chi + \epsilon}$$

COMPOSITE RELIABILITY

Scale variance explained by a common factor among items

TEST-RETEST RELIABILITY

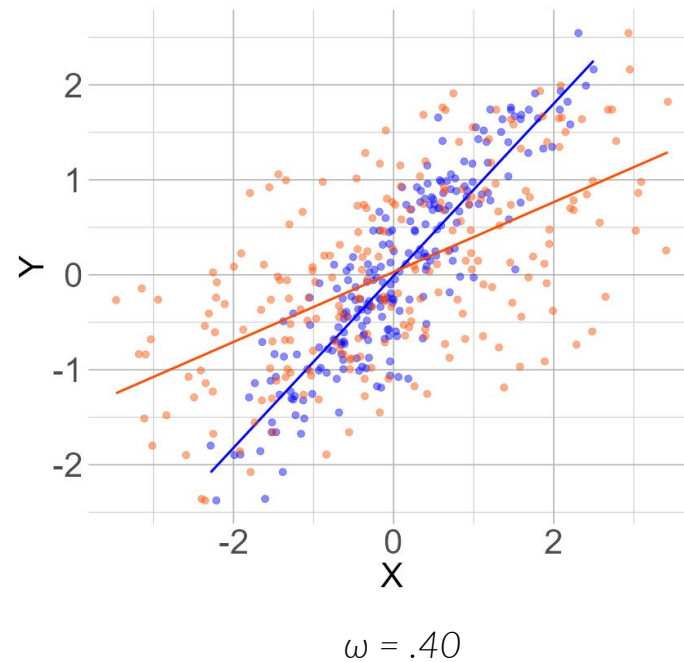
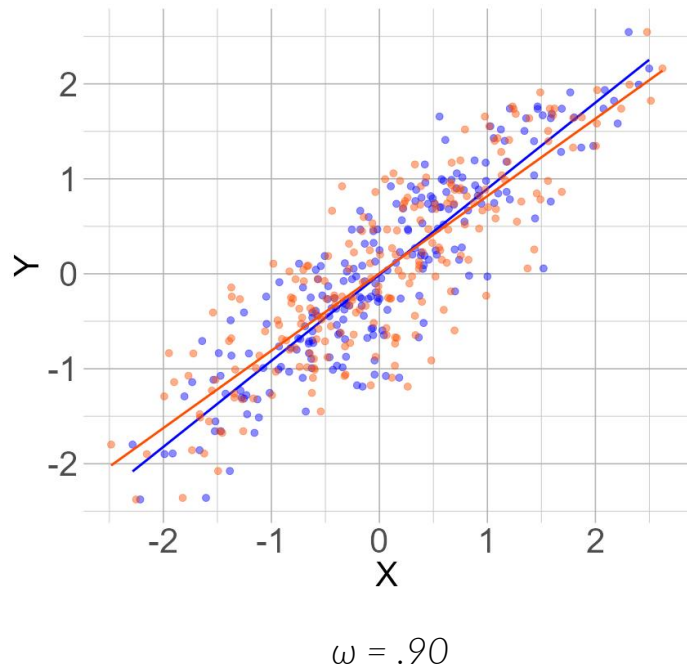
Expected scores at the *second measurement time*, given the first

MEASUREMENT PRACTICES

Reporting Practices.

Only 33-65% publications report composite reliability^[1,2]
All of them used a latent measure.

Monte-Carlo simulation ($N=230$; $r_{\text{true}} = .90$; ● observed):

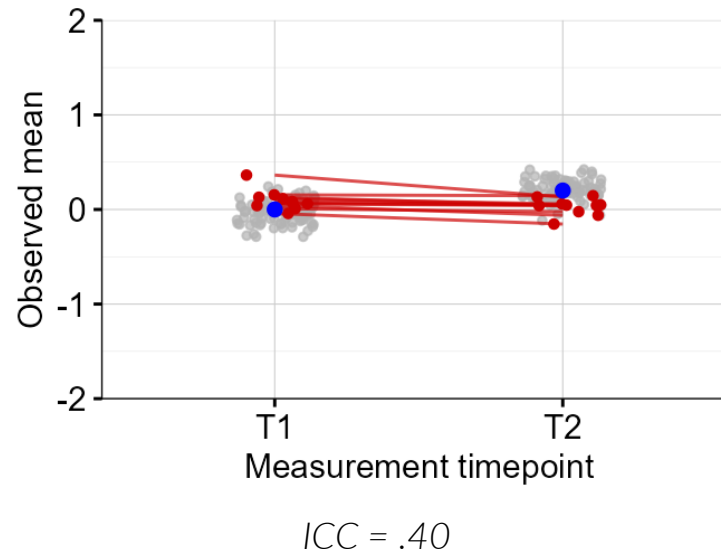
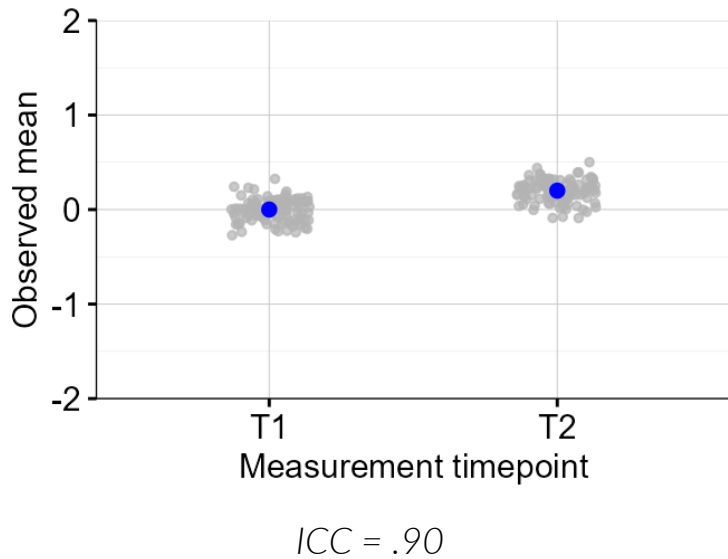


Reporting Practices.

0% publications report test-retest reliability [1,2]


There are at least some pre-post study designs in S&P.

Monte-Carlo simulation ($N=50$; $\Delta_{\text{true}} = .20$; ● reversed):




What Now? Hopes & Dreams

Reliability evidence as the first step

Composite reliability: ω 

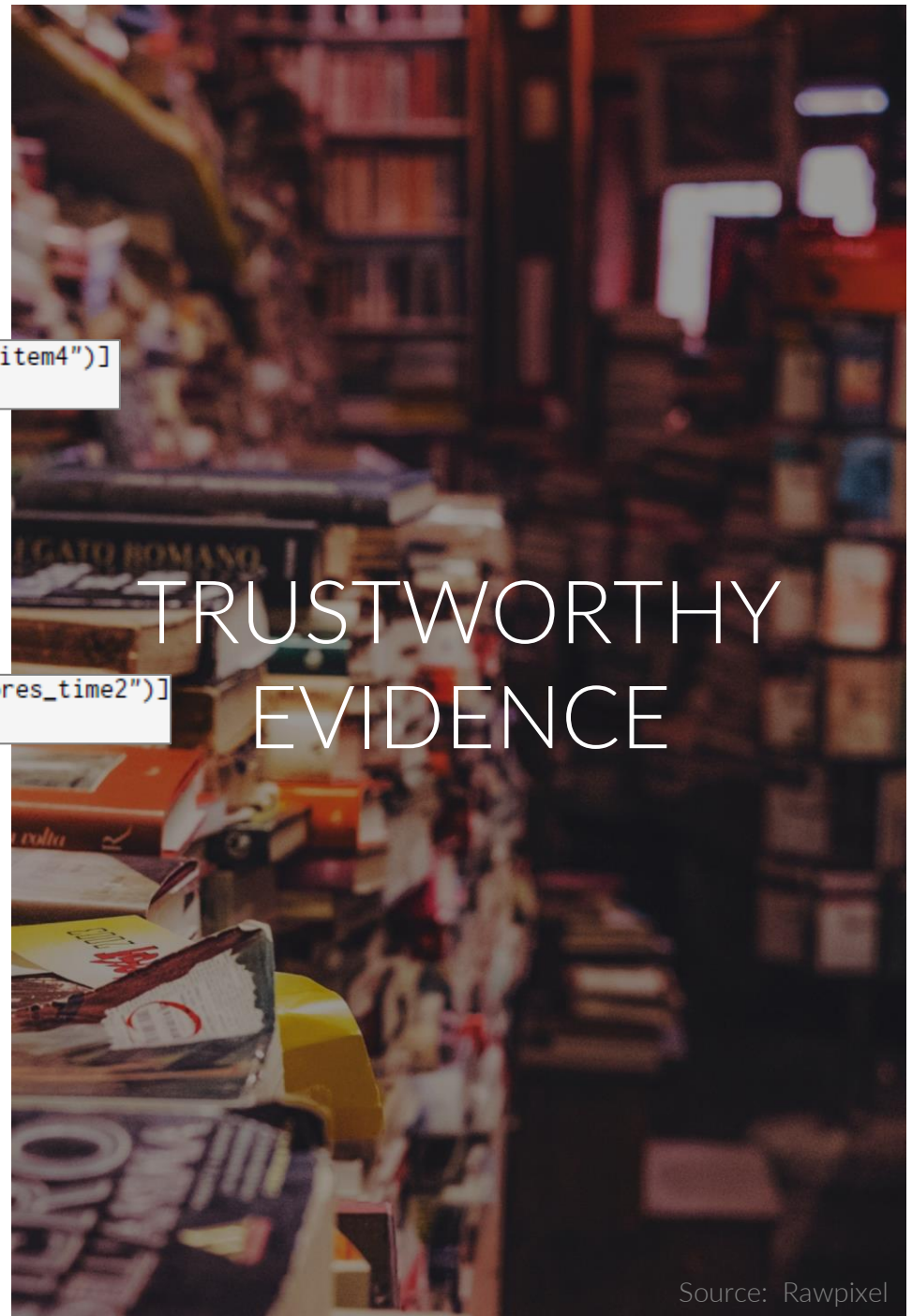
```
items <- dat[, c("item1","item2","item3","item4")]  
psych::omega(items, nfactors = 1)
```

~~Cronbach's α~~ (e.g., would need τ -equivalence evidence)

Test-retest reliability: ICC 

```
both_scores <- dat[, c("scores_time1","scores_time2")]  
psych::ICC(both_scores, alpha = .05)
```

For ALL research questions concerning change over time



We cannot trust the findings of human-centered security and privacy due to missing/low precision.



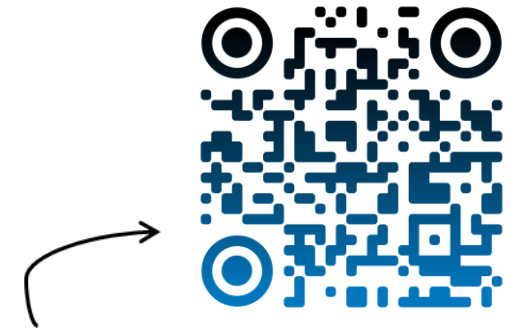
Join the “Validity by Design” Movement.

Talk to me today :-)

Or email us: nele.borgert@unibe.ch

Reference.

Nele Borgert, Tim Ulmann, and Robin Merchel. 2026. Position Paper: On the (Non-)Precision of Psychological Measures in Security and Privacy Research -- Reliability Beyond Reporting Alpha. *SPW'26 MetaCRiSP, Mai 21, 2026, CA, USA*. https://doi.org/10.31234/osf.io/87qmx_v1



Other References.

- [1] N. Borgert, L. Jansen, I. Böse, J. Friedauer, M. A. Sasse, and M. Elson, “Self-efficacy and security behavior: Results from a systematic review of research methods,” in Proceedings of the CHI Conference on Human Factors in Computing Systems, F. F. Mueller, P. Kyburz, J. R. Williamson, C. Sas, M. L. Wilson, P. T. Dugas, and I. Shklovski, Eds. New York, NY, USA: ACM, 2024, pp. 1–32.
- [2] S. A. C. Perrig, L. F. Aeschbach, N. Scharowski, N. von Felten, K. Opwis, and F. Brühlmann, “Measurement practices in user experience (UX) research: a systematic quantitative literature review,” *Frontiers in Computer Science*, vol. 6, 2024.