Manuscript\_Draft

JoeNRoe

2023-03-14

### Introduction

Why is this work needed? Importance statement

##### Minimum Descriptive Length

Introduction of the concept of minimum descriptive length (Grunwald 2004)

##### Fit Propensity

Extensions of MDL concept into model fit propensity (cite Hansen & Yu 2001, Preacher 2007)

##### Wes Bunnifay’s Dissertation/Pub

A key limitation of the Wes Bunnifay’s 2017 work was the data simulation methods - specifically not restricting the simulated data to datasets that are plausible (rather than data that’s wholly random noise). Another limitation is that the work is mostly conducted in FlexMIRT and the analysis code was not published with the manuscript- making it hard or impossible to fully understand the work and/or replicate.

Extend by also generating data from the space of legitimate correlation matrices (instead of random noise)

### Methods

#### Replication

Random Vector sampling of 1 million cases of 7 variables, a la Bunnifay 2017

##### Data (Simulation)

##### Analysis

Analyses are conducted in R v4.2.2 (Cite R Core team) and with the MIRT package (Cite phil chalmers)

1. EIFA: Exploratory item factor analysis
2. Bifactor model
3. DINA: deterministic input noisy and-gate
4. DINO: deterministic input noisy or-gate
5. Unidimensional 3PL

#### New work

##### Data (Simulation)

Sampling of data (n=1m?) from plausible correlation matrices (e.g., fungible or cluster packages)

##### Analysis

1. EIFA: Exploratory item factor analysis
2. Bifactor model
3. DINA: deterministic input noisy and-gate
4. DINO: deterministic input noisy or-gate
5. Unidimensional 3PL

### Results

#### Replication

Y2/n

Local Dependence

Amoeba plots

#### New Work

Y2/n

Local Dependence

Amoeba plots

### Discussion

### References

### Acknowledgements and Disclosures

This work was funded by the letter Q and supported by lots of coffee/snacks. All code and manscript drafts (including preregistration) are available on OSF and Github, here: <https://github.com/zenit125/WesBunnifayReplication> [JOSEPH- would you be okay with putting code/etc on osf before pub, and maybe prereg? I can do the writing and management of that!]

OG Wes Bunnifay cite:  
<https://www.tandfonline.com/doi/abs/10.1080/00273171.2017.1309262>

Hansen & Yu MDL citation:  
Mark H Hansen & Bin Yu (2001) Model Selection and the Principle of Minimum Description Length, Journal of the American Statistical Association, 96:454, 746-774, DOI: 10.1198/016214501753168398

Grunwald MDL tutorial:  
<https://arxiv.org/abs/math/0406077>

Preacher MDL SEM:  
Preacher, K. J., Cai, L., & MacCallum, R. C. (2007). Alternatives to traditional model comparison strategies for covariance structure models. In T. D. Little, J. A. Bovaird, & N. A. Card (Eds.), Modeling contextual effects in longitudinal studies (pp. 33–62). Lawrence Erlbaum Associates Publishers.

R Core Team Cite:  
R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Phil Chalmers & MIRT Package:  
Chalmers RP (2012). “mirt: A Multidimensional Item Response Theory Package for the R Environment.” Journal of Statistical Software, 48(6), 1–29. <doi:10.18637/jss>