Visualizing and Interpreting Multi-Group Confirmatory Factor Analysis

Erin M. Buchanan[1]

[1] Harrisburg University of Science and Technology

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Erin M. Buchanan: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Erin M. Buchanan, 326 Market St., Harrisburg, PA, USA. E-mail: ebuchanan@harrisburgu.edu

Abstract

Latent variable modeling as a lens for psychometric theory is a popular tool for social scientists to examine measurement of constructs (Beaujean, 2014). Journals such as *Assessment* regularly publish articles supporting new or previously established measures of latent constructs (e.g., depression, anxiety) wherein a measurement model is established for the scale in question. These measurement models designate the relationship between the measured, observed variables, and the underlying construct, with the assumption that these relations hold in many samples. Confirmatory factor analysis can be used to investigate the replicability and generalizability of the measurement model in new samples, while multi-group confirmatory factor analysis is used to examine the measurement model across groups within samples (Brown, 2015). With the rise of the replication crisis and "psychology's renaissance" (Nelson, Simmons, & Simonsohn, 2018), interest in divergence in measurement has increased, often focused on small parameter differences within the latent model. While the statistical procedure for examining measurement invariance is moderately well established, it is clear that the toolkit for inspecting these results is lacking. This manuscript will outline ways to visualize potential non-invariance, to supplement large numbers of tables that often overwhelm a reader within these published reports. Further, given these visualizations, readers will learn how to interpret the impact and size of the proposed non-invariance in models. While it is tempting to suggest that problems with replication and generalizability are simply issues with measurement, it is crucial to remember that all models have variability and error, even those models estimating the differences between item functioning, such as multi-group confirmatory factor analysis. This manuscript will help provide a framework for researchers interested in registered reports in this area.

Keywords: multigroup confirmatory factor analysis, measurement invariance, visualization, effect size

Visualizing and Interpreting Multi-Group Confirmatory Factor Analysis

Outline

- talk about LVM
- talk about cfa
- talk about mgcfa
- how are measurement and replication/crisis related
- why it is a bad idea to say that replication/crisis are *because* bad measurement
- how can we visualize and interpret MGCFA to help us understand the impact of measurement differences

By the end of this tutorial manuscript, readers will:

1. Be able to create visualizations for common steps to multi-group confirmatory factor analysis.
2. Be able to interpret the impact and size of potential non-invariance on measurement.
3. Understand the impact of measurement variability on replication and generalizability.

Method

Design and Analysis

Data was simulated using the `simulateData` function in the *R* package `lavaan` (Rosseel, 2012) assuming multivariate normality using a μ of 0 and σ of 1 for the data. This function allows you to write `lavaan` syntax for your model with estimated values to generate data for observed variables. The data included two groups of individuals ("Group 1", "Group 2") for a multi-group confirmatory factor analysis (n_{group} = 250, N = 500). The latent variables were assumed to be continuous normal. The model consisted of five observed items predicted by one latent variable (`lv =~ q1 + q2 + q3 + q4 + q5`); however, the demonstration in this manuscript extends to multiple latent variables and other combinations of observed variables. Each item was assumed to be related to the latent variable with loadings approximately equal to .40 to .80, except when cases of non-invariance on the loadings was assumed.

The Brown (2015) steps of testing measurement invariance are demonstrated in this manuscript for illustration purposes, but in line with Stark, Chernyshenko, and Drasgow (2006) suggestions, the visualizations show the impact of loadings and intercepts together. The configural model was analyzed nesting both groups into the same CFA model requiring that both groups show the same model structure, but all other parameters are free to vary between groups. The metric model constrained the factor loadings of each group to be equal within the model. The scalar model then constrained the item intercepts (i.e., item mean) to be equal across groups. Finally, the strict model constrained the item variances (i.e., error variances) to be equal for each item across groups. These models are normally tested sequentially, and a convenience function `mgcfa` is provided in the supplemental documents for this manuscript.

The data was then simulated to represent invariance across all model steps, small, medium, and large invariance using d_{MACS} estimated sizes from Nye, Bradburn, Olenick, Bialko, and Drasgow (2019). While d_{MACS} is used primarily for an effect size of the (non)-invariance for intercepts and loadings together, a similar approach was taken for the estimation of small, medium, and large effects on the residuals. The effect size is presented for all models, calculated from the *dmacs* package Nye & Drasgow (2011). Only one item in each model was manipulated from the invariant model to create the non-invariant models.

Results

Discussion

Conclusions:

- framework for submitted/interpreting reports

References

Beaujean, A. A. (2014). *Latent variable modeling using r: A step by step guide*. New York: Routledge/Taylor & Francis Group.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition). New York ; London: The Guilford Press.

Dueber, D. (2023). *Dmacs*. Retrieved from https://github.com/ddueber/dmacs

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology*, *69*(1), 511–534. https://doi.org/10.1146/annurev-psych-122216-011836

Nye, C. D., Bradburn, J., Olenick, J., Bialko, C., & Drasgow, F. (2019). How Big Are My Effects? Examining the Magnitude of Effect Sizes in Studies of Measurement Equivalence. *Organizational Research Methods*, *22*(3), 678–709. https://doi.org/10.1177/1094428118761122

Nye, C. D., & Drasgow, F. (2011). Effect size indices for analyses of measurement equivalence: Understanding the practical importance of differences between groups. *Journal of Applied Psychology*, *96*(5), 966–980. https://doi.org/10.1037/a0022955

Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling. *Journal of Statistical Software*, *48*(1), 1–36. https://doi.org/10.18637/jss.v048.i02

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, *91*(6), 1292–1306. https://doi.org/10.1037/0021-9010.91.6.1292

Appendix

Table 1

Model Fit for Invariant Model

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,505.72	7,547.87	0.99	0.99	0.02	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,751.95	3,787.17	0.98	0.96	0.04	0.04
configural	7,527.70	7,654.14	0.98	0.96	0.04	0.03
metric	7,529.39	7,638.97	0.95	0.93	0.06	0.05
scalar	7,522.90	7,615.62	0.96	0.96	0.04	0.05
strict	7,519.51	7,591.16	0.96	0.96	0.04	0.06

Table 2

Model Fit for Small Differences in Loadings

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,527.67	7,569.81	0.98	0.96	0.04	0.03
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,767.83	3,803.05	0.98	0.96	0.05	0.04
configural	7,543.58	7,670.02	0.98	0.95	0.05	0.03
metric	7,548.90	7,658.48	0.94	0.92	0.07	0.06
scalar	7,541.81	7,634.53	0.95	0.95	0.05	0.06
strict	7,541.66	7,613.31	0.93	0.94	0.05	0.07

Table 3

Model Fit for Medium Differences in Loadings

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,544.55	7,586.70	0.97	0.94	0.05	0.03
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,774.92	3,810.14	1.00	1.00	0.02	0.03
configural	7,550.67	7,677.11	0.99	0.98	0.04	0.03
metric	7,562.71	7,672.29	0.93	0.89	0.07	0.06
scalar	7,556.86	7,649.58	0.93	0.93	0.06	0.06
strict	7,558.05	7,629.70	0.91	0.92	0.06	0.08

Table 4

Model Fit for Large Differences in Loadings

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,652.99	7,695.14	0.98	0.97	0.04	0.03
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,847.21	3,882.42	0.97	0.94	0.08	0.04
configural	7,622.96	7,749.40	0.97	0.94	0.06	0.03
metric	7,659.19	7,768.77	0.85	0.79	0.12	0.08
scalar	7,652.60	7,745.32	0.86	0.85	0.10	0.09
strict	7,660.63	7,732.27	0.82	0.85	0.10	0.12

Table 5

Model Fit for Small Differences in Intercepts

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,509.69	7,551.83	1.00	0.99	0.02	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,760.41	3,795.63	0.93	0.86	0.08	0.05
configural	7,536.16	7,662.60	0.95	0.91	0.07	0.04
metric	7,531.36	7,640.94	0.96	0.94	0.05	0.04
scalar	7,531.34	7,624.06	0.94	0.93	0.06	0.05
strict	7,523.54	7,595.18	0.95	0.96	0.04	0.05

Table 6

Model Fit for Medium Differences in Intercepts

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,532.77	7,574.92	1.00	1.00	0.01	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,760.41	3,795.63	0.93	0.86	0.08	0.05
configural	7,536.16	7,662.60	0.95	0.91	0.07	0.04
metric	7,531.36	7,640.94	0.96	0.94	0.05	0.04
scalar	7,554.20	7,646.92	0.85	0.83	0.09	0.07
strict	7,546.38	7,618.03	0.86	0.88	0.08	0.07

Table 7

Model Fit for Large Differences in Intercepts

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,569.17	7,611.31	1.00	1.00	0.00	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,760.41	3,795.63	0.93	0.86	0.08	0.05
configural	7,536.16	7,662.60	0.95	0.91	0.07	0.04
metric	7,531.36	7,640.94	0.96	0.94	0.05	0.04
scalar	7,590.29	7,683.01	0.70	0.66	0.13	0.10
strict	7,582.47	7,654.12	0.71	0.75	0.11	0.10

Table 8

Model Fit for Small Differences in Residuals

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,439.49	7,481.64	1.00	1.01	0.00	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,683.32	3,718.53	1.00	1.01	0.00	0.02
configural	7,459.07	7,585.51	0.99	0.98	0.03	0.03
metric	7,461.41	7,570.99	0.97	0.95	0.05	0.05
scalar	7,455.85	7,548.58	0.97	0.97	0.04	0.05
strict	7,453.48	7,525.12	0.96	0.97	0.04	0.05

Table 9

Model Fit for Medium Differences in Residuals

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,368.57	7,410.71	1.00	1.00	0.00	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,587.77	3,622.99	1.00	1.03	0.00	0.02
configural	7,363.52	7,489.96	1.00	0.99	0.02	0.02
metric	7,366.63	7,476.21	0.97	0.96	0.05	0.05
scalar	7,360.15	7,452.87	0.98	0.98	0.03	0.05
strict	7,382.53	7,454.18	0.88	0.90	0.08	0.07

Table 10

Model Fit for Large Differences in Residuals

Model	AIC	BIC	CFI	TLI	RMSEA	SRMR
overall	7,284.21	7,326.36	1.00	1.01	0.00	0.02
group1	3,755.75	3,790.96	0.98	0.95	0.05	0.04
group2	3,443.47	3,478.69	0.95	0.90	0.07	0.04
configural	7,219.22	7,345.66	0.96	0.92	0.06	0.03
metric	7,216.38	7,325.96	0.96	0.94	0.05	0.04
scalar	7,210.65	7,303.37	0.96	0.96	0.04	0.05
strict	7,297.89	7,369.54	0.59	0.65	0.13	0.18

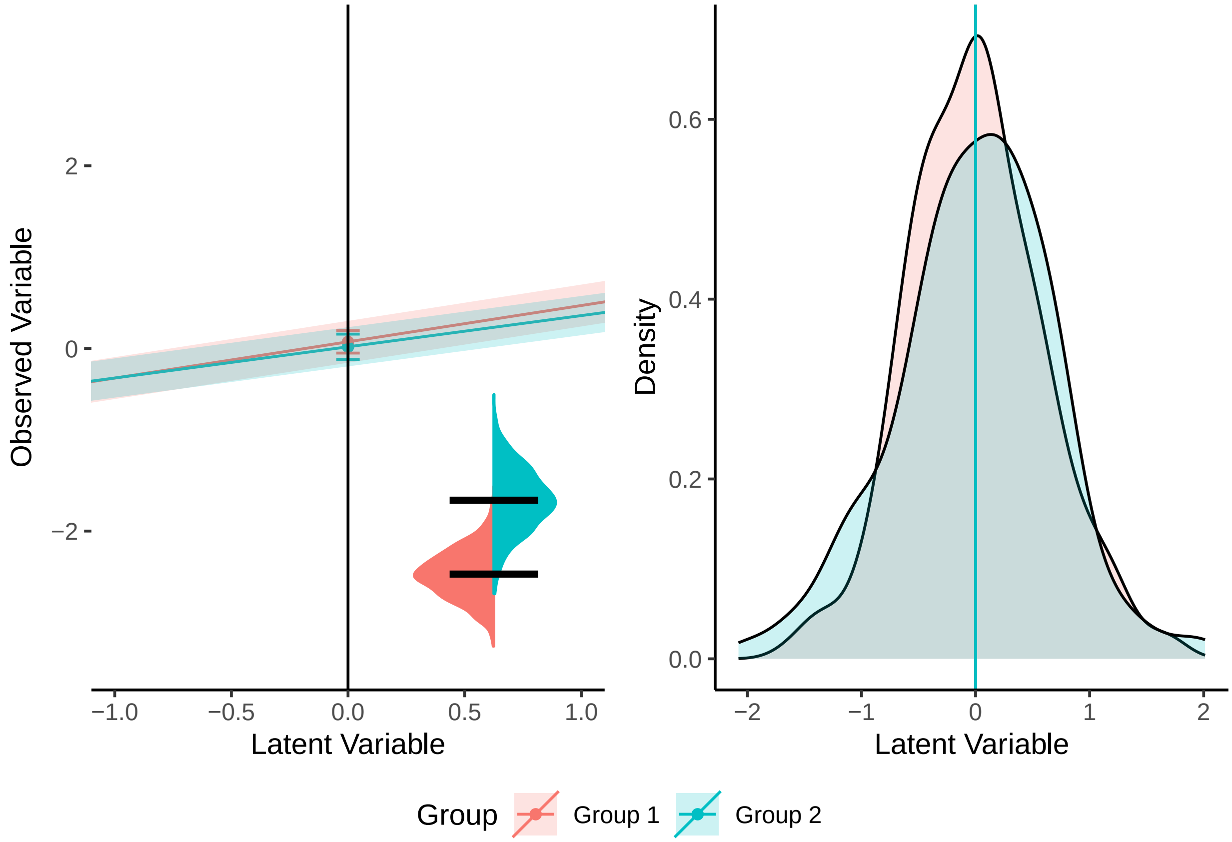

Figure 1. Invariant Model Visualization

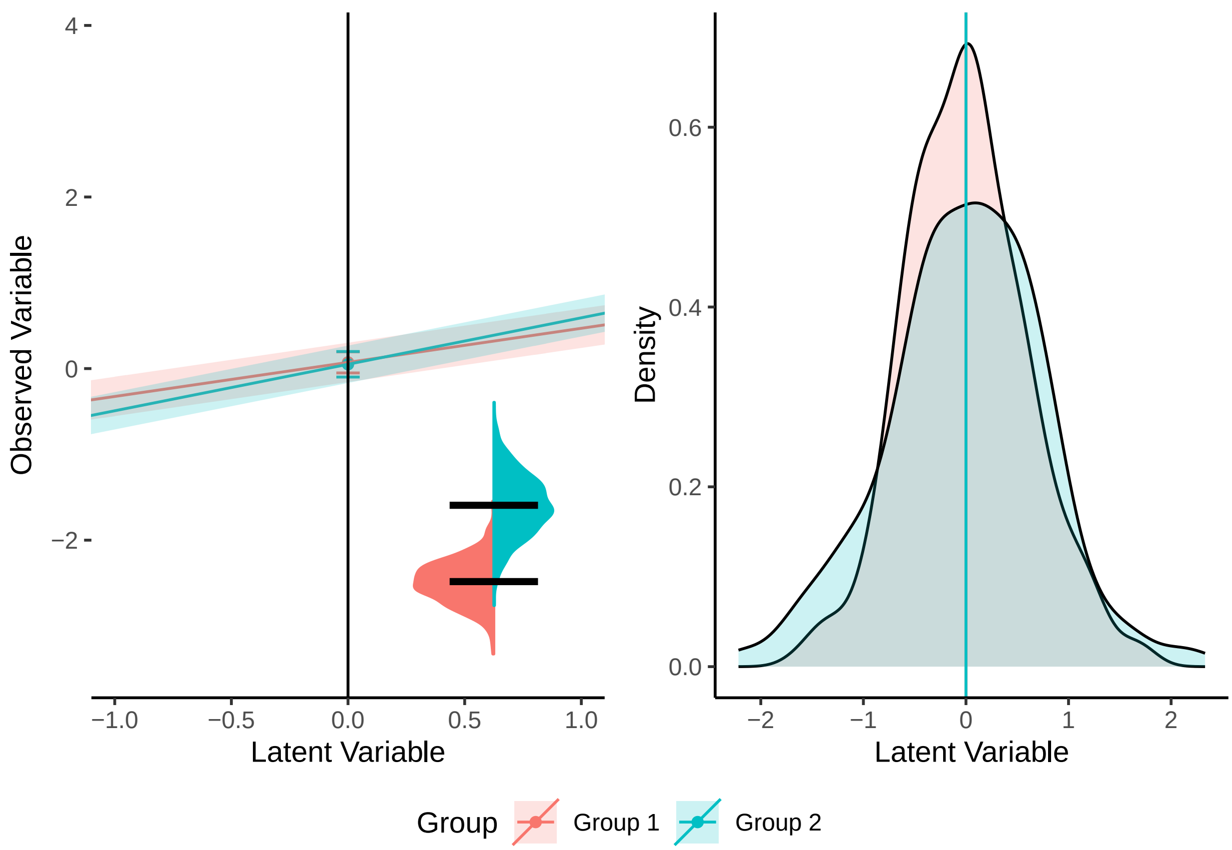

Figure 2. Small Loadings Model Visualization

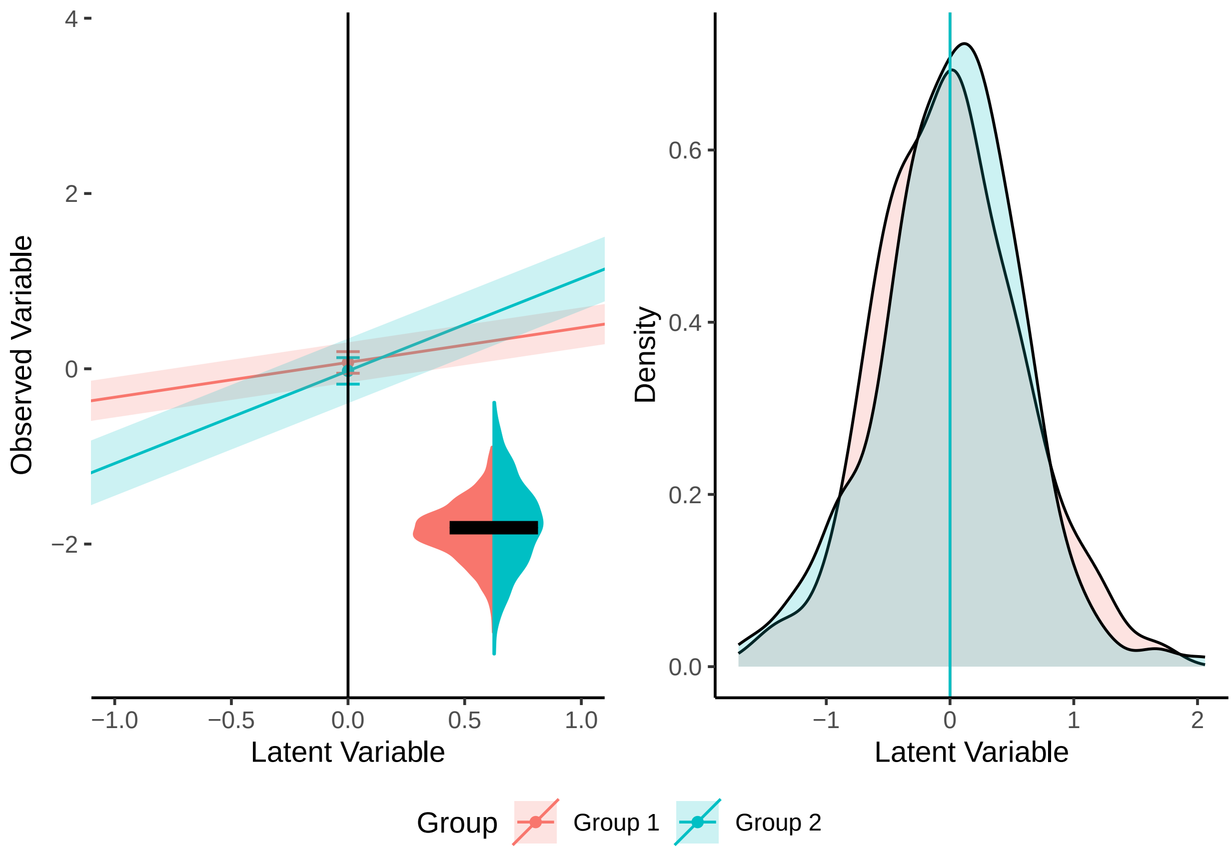

Figure 3. Medium Loadings Model Visualization

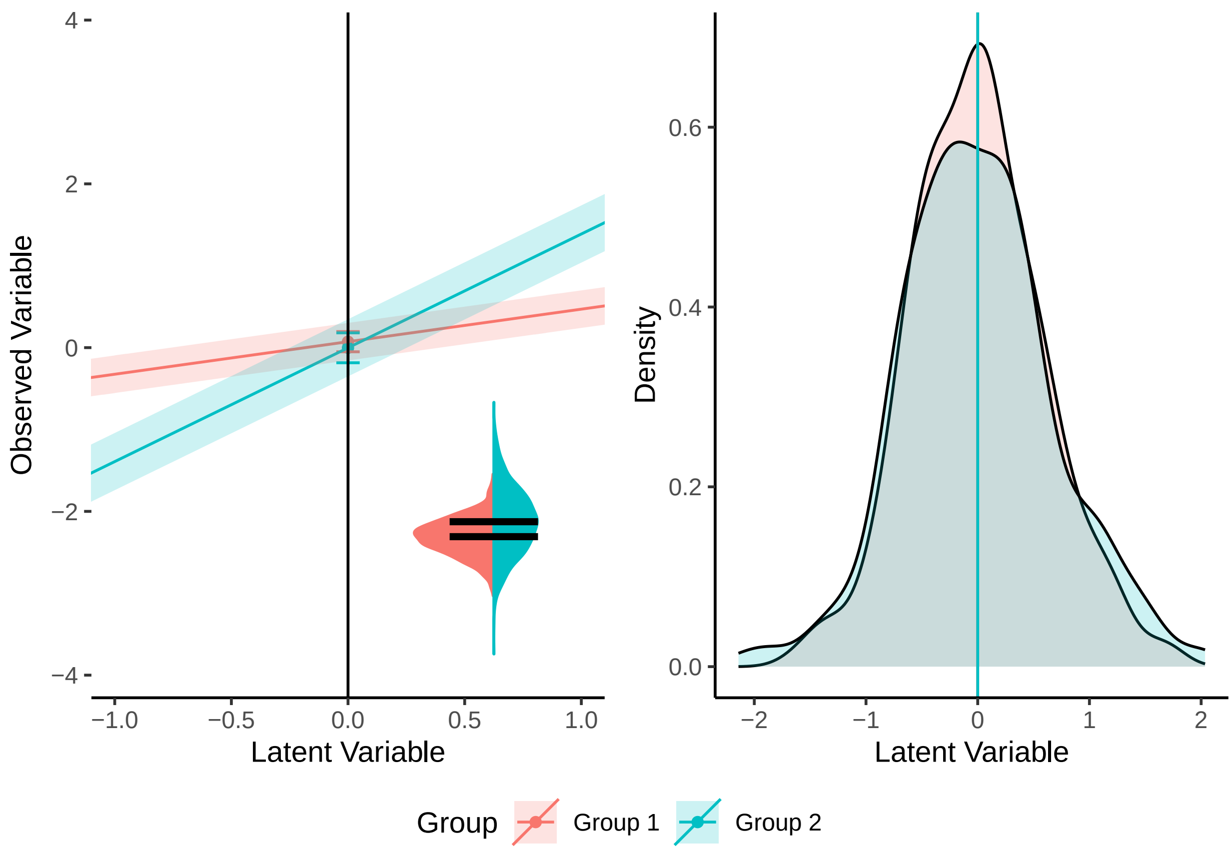

Figure 4. Large Loadings Model Visualization

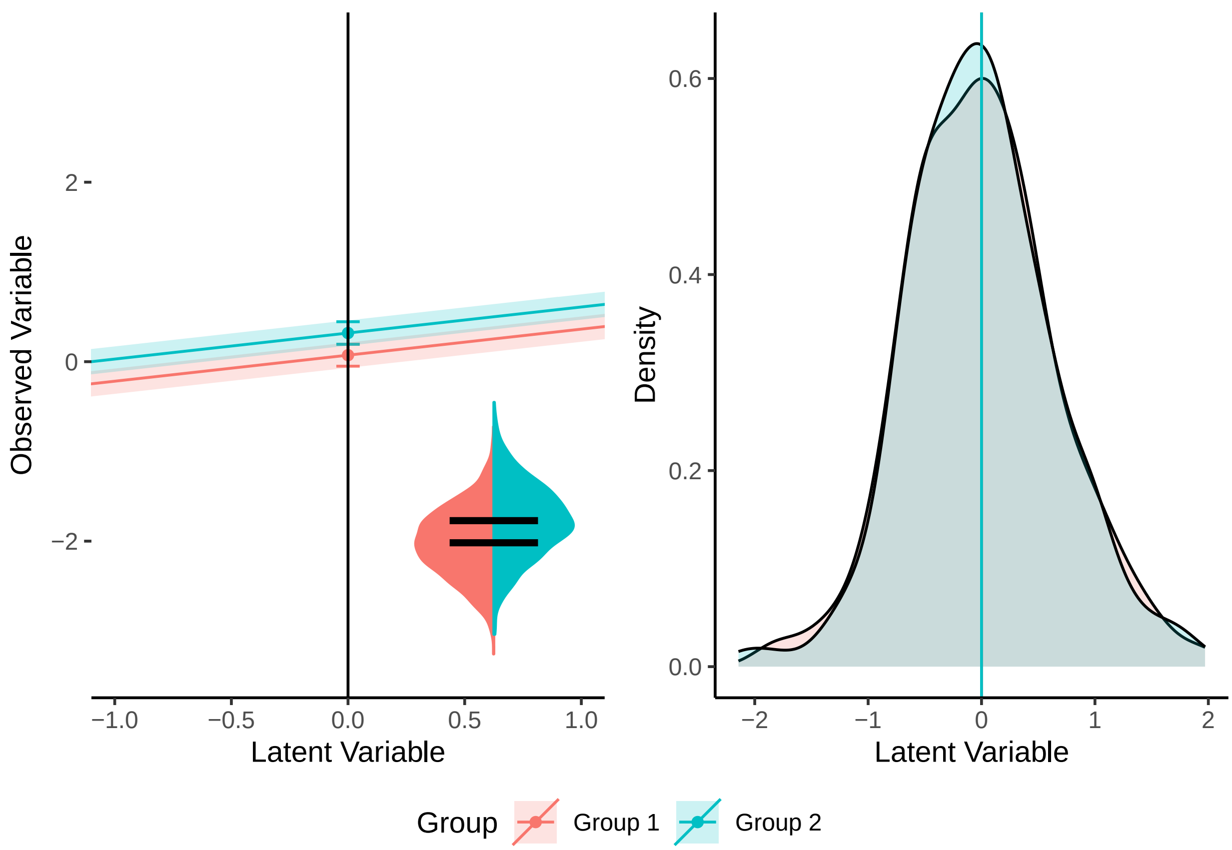

Figure 5. Small Intercepts Model Visualization

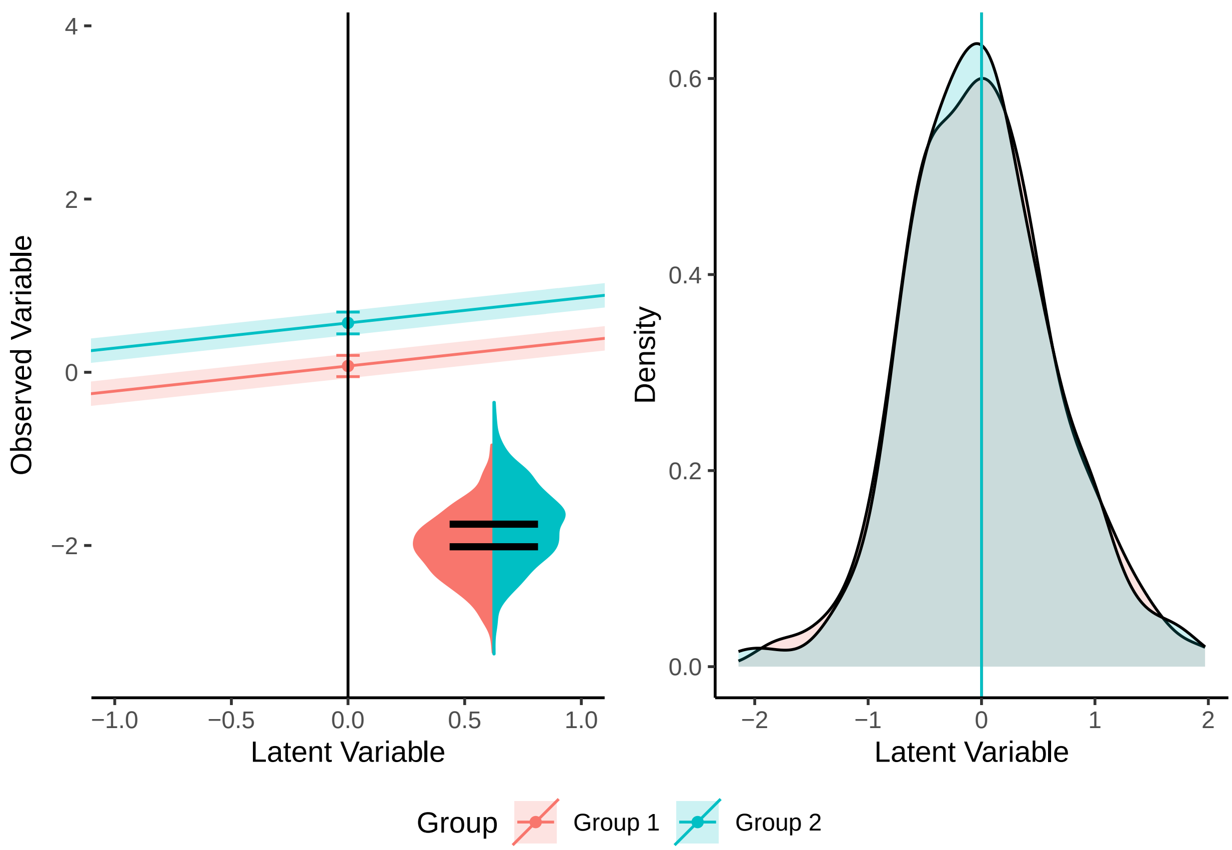

Figure 6. Medium Intercepts Model Visualization

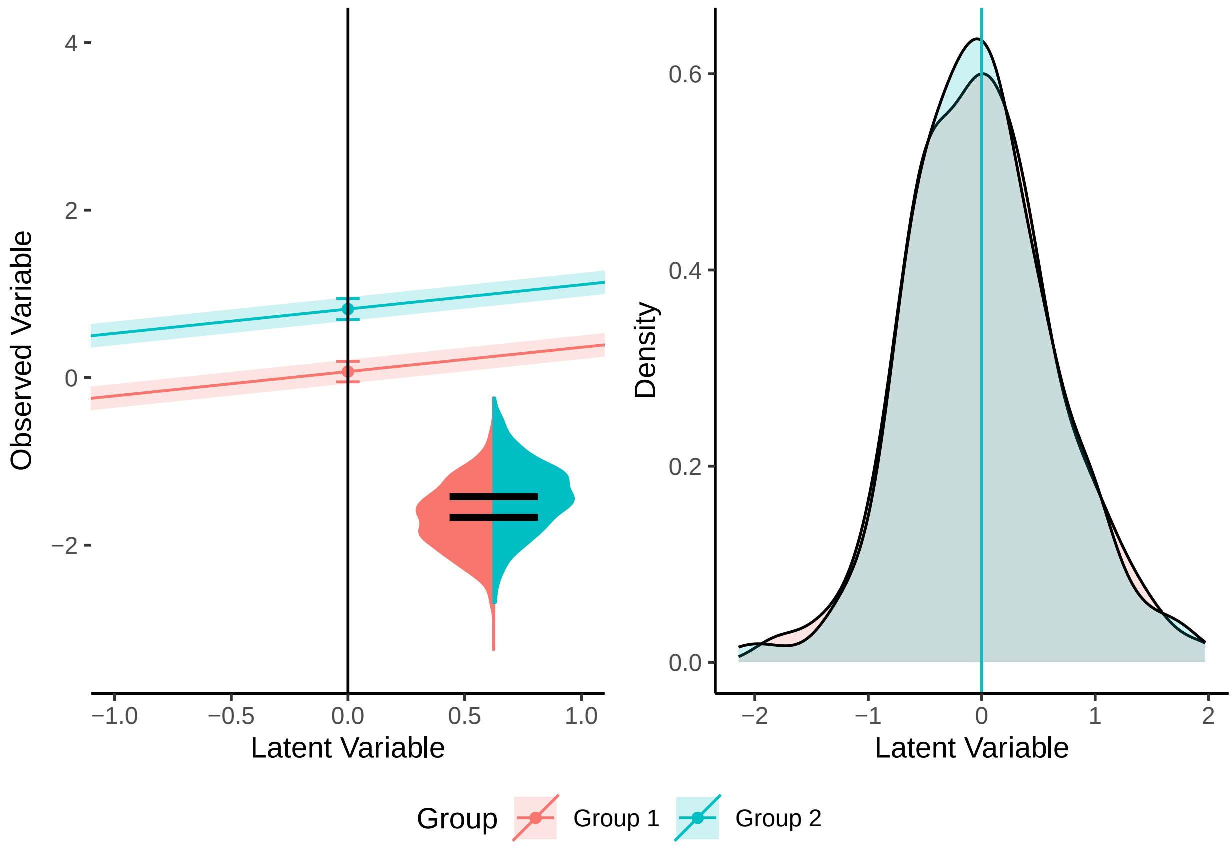

Figure 7. Large Intercepts Model Visualization

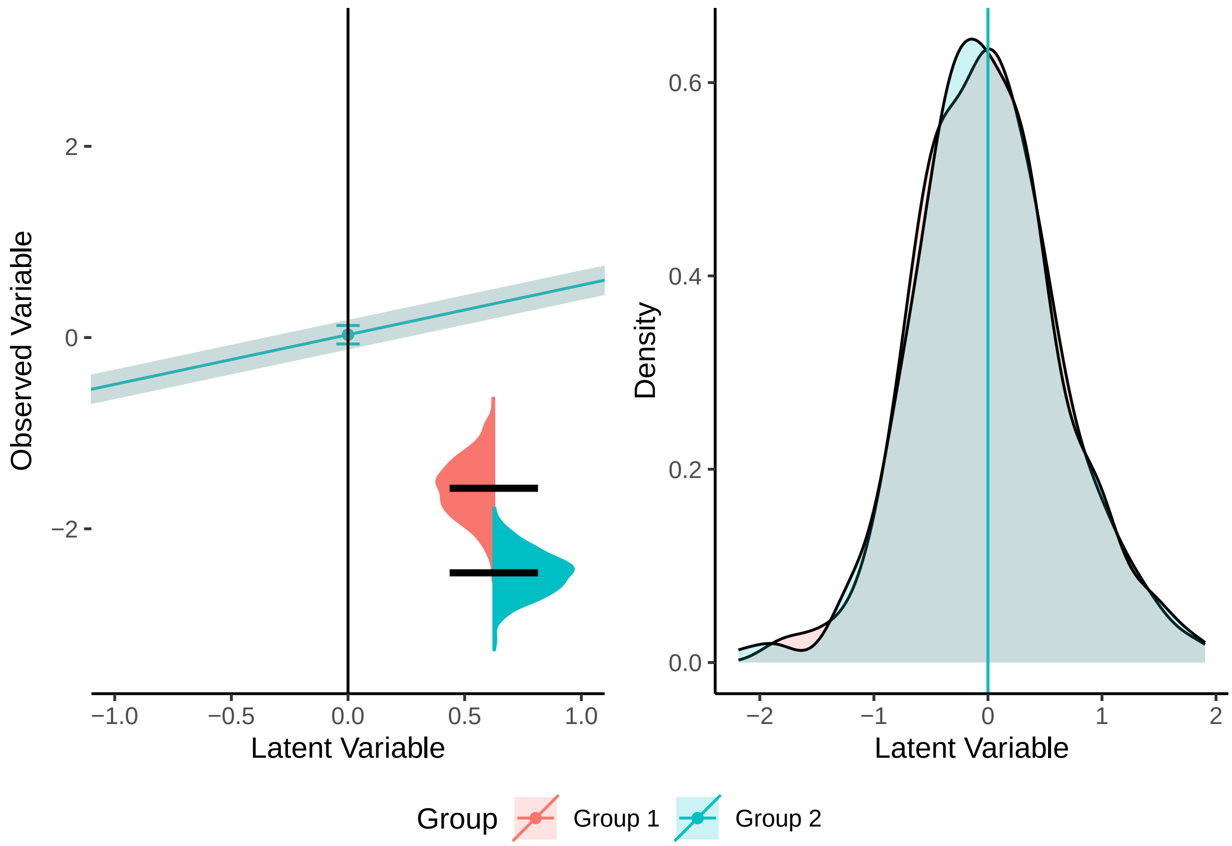

Figure 8. Small Residuals Model Visualization

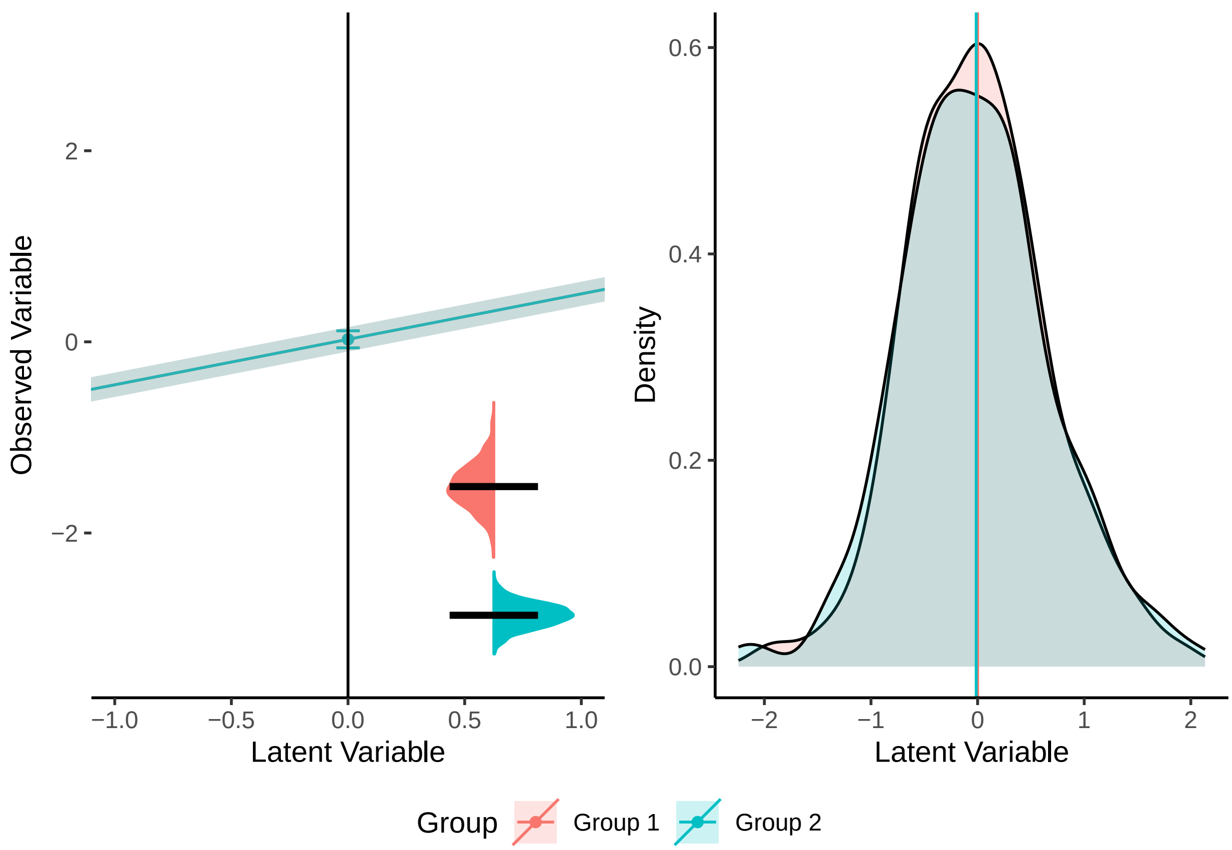

Figure 9. Medium Residuals Model Visualization

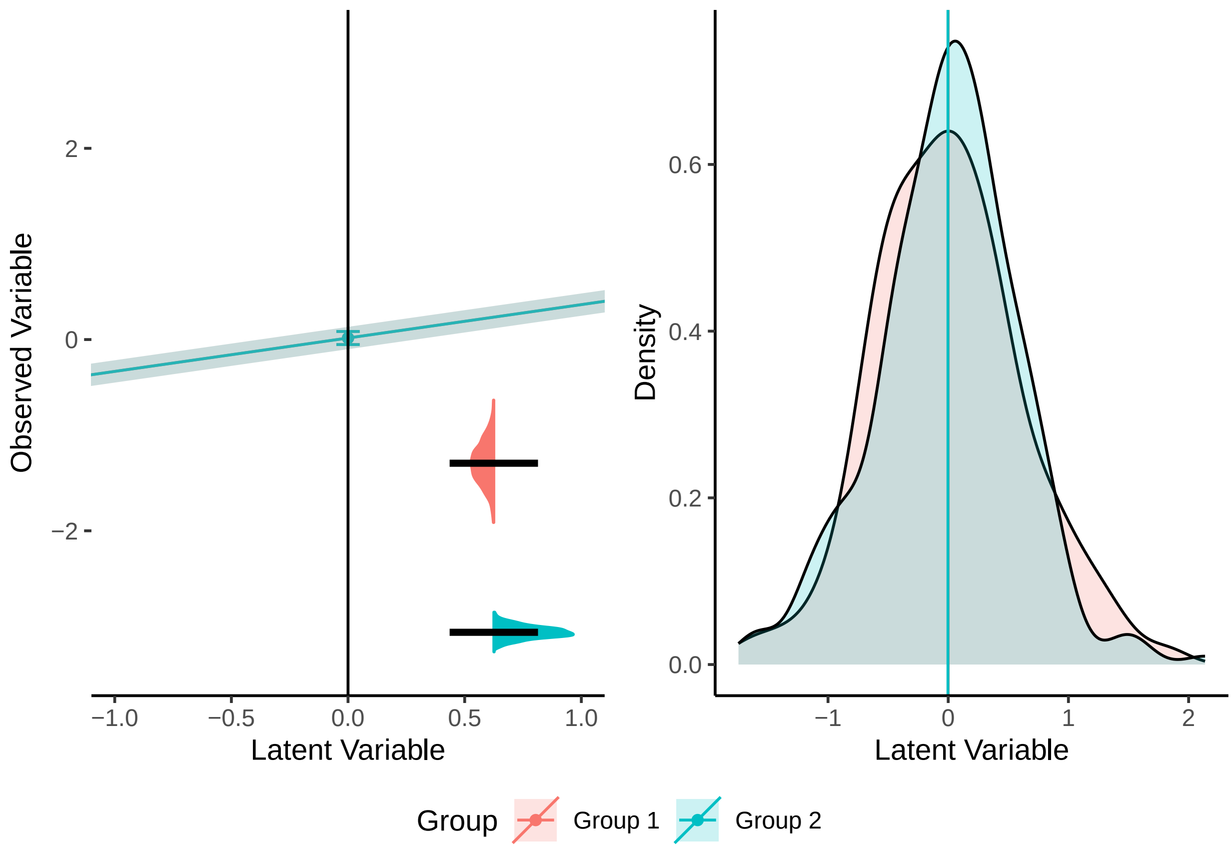

Figure 10. Large Residuals Model Visualization