Measurement Equivalence with Categorical Indicators

Thresholds as a First-Order Measurement Model

Terrence D. Jorgensen

Last updated: 18 June 2020

Overview

The Common Factor Model

Gotta start a presentation somewhere . . .

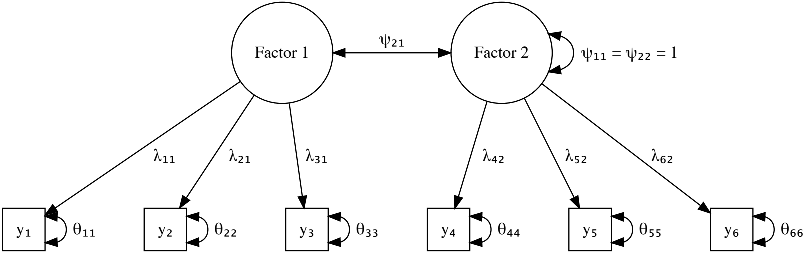


Figure 1: 2-factor CFA, 3 indicators per factor.

Outline

- Review identification constraints in CFA models
  - Trading identification constraints for invariance constraints
- Adding probit link to CFA
  - Threshold model as measurement model for latent item-response
  - Numerical examples with single indicators
- Measurement equivalence with multiple discrete indicators
  - Identify locations/scales of latent common factors **and** LIRs
  - Conflicting advice:
    - Millsap & Tein (2004)
    - Sofware defaults
    - Wu & Estabrook (2016)
  - Examples using `semTools::measEq.syntax()` function

Review Identification Constraints in CFA Models

A Single Context

3 statistically equivalent methods to identify the arbitrary location & scale of common factors:

- **Fixed-Factor** method: Set each construct's $\mu = 0$ and $\sigma^2 = 1$
- **Marker-Variable** method: For each construct, fix a *referent indicator*'s intercept $= 0$ and factor loading $= 1$
- **Effects-Coding** method: For each construct, constrain the *mean* of all intercepts $= 0$ and the *mean* of all loadings $= 1$

Additional constraints necessary with < 3 indicators:

- 1 indicator: Fix residual variance (to zero or $1 - \sigma_{xx}$)
- 2 indicators: Constrain loadings to equality (or both to 1), unless factor is substantially correlated with at least one other common factor

Counting Knowns & Unknowns: 1 Indicator

In a model with only a single indicator construct, we only observe a 1×1 covariance matrix

	X
X	2

The model-implied variance is a function of 3 parameters, only 1 of which can be estimated (resulting in *df* $= 1 - 1 = 0$)

	X
X	$\lambda_{1,1}\psi_{1,1}\lambda_{1,1} + \theta_{1,1}$

- Fix $\theta_{1,1} = 0$, estimate *either* $\lambda_{1,1}$ or $\psi_{1,1}$

Counting Knowns & Unknowns: 2 Indicators

With 2 indicators, we observe a 2×2 covariance matrix

	X	Y
X	2.00	0.87
Y	0.87	1.50

	X	Y
X	$\lambda_{1,1}\psi_{1,1}\lambda_{1,1} + \theta_{1,1}$	
Y	$\lambda_{2,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{2,1}\psi_{1,1}\lambda_{2,1} + \theta_{2,2}$

- Must set $\lambda_{1,1} = \lambda_{2,1}$
  - Optionally, both $= 1$ and estimate $\psi_{1,1}$
- Estimate both $\theta_{1,1}$ and $\theta_{2,2}$, $df = 3 - 3 = 0$

Counting Knowns & Unknowns: 3 Indicators

A 3-indicator model is just-identified ($df = 6 - 6 = 0$) once the latent scale is set

	X	Y	Z
X	2.00	0.87	0.67
Y	0.87	1.50	0.77
Z	0.67	0.77	2.50

	X	Y	Z
X	$\lambda_{1,1}\psi_{1,1}\lambda_{1,1} + \theta_{1,1}$		
Y	$\lambda_{2,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{2,1}\psi_{1,1}\lambda_{2,1} + \theta_{2,2}$	
Z	$\lambda_{3,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{3,1}\psi_{1,1}\lambda_{2,1}$	$\lambda_{3,1}\psi_{1,1}\lambda_{3,1} + \theta_{3,3}$

Counting Knowns & Unknowns: 4 Indicators

With ≥ 4 indicators, the model is over-identified

- Still need to set the latent scale for identification

$$df = 10 - 8 = 2$$

	W	X	Y	
W	$\lambda_{1,1}\psi_{1,1}\lambda_{1,1} + \theta_{1,1}$			
X	$\lambda_{2,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{2,1}\psi_{1,1}\lambda_{2,1} + \theta_{2,2}$		
Y	$\lambda_{3,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{3,1}\psi_{1,1}\lambda_{2,1}$	$\lambda_{3,1}\psi_{1,1}\lambda_{3,1} + \theta_{3,3}$	
Z	$\lambda_{4,1}\psi_{1,1}\lambda_{1,1}$	$\lambda_{4,1}\psi_{1,1}\lambda_{2,1}$	$\lambda_{4,1}\psi_{1,1}\lambda_{3,1}$	$\lambda_{4,1}\psi$

Multiple Contexts

Measure the same construct(s) in different contexts

- multiple populations or occasions, experimental conditions, members of dyads, etc.

Compare distributional parameters

- latent means, variances, correlations, directed effects
- autoregressive/cross-lagged effects, growth trajectories

Requires homogenous measurement properties across contexts

- Inference about latent covariance structure requires metric equivalence (invariance factor loadings)
- Inference about latent mean structure requires scalar equivalence (invariance of loadings and intercepts)

Testing Measurement Equivalence in CFA

Assumptions are testable with latent variable models. Typically follows a sequence from least- to most-constrained model:

- **Configural invariance**
  - Same configuration of fixed/free parameters across contexts
- **Metric equivalence** (or *weak invariance*)
  - Additionally, equal factor loadings across contexts
- **Scalar equivalence** (or *strong invariance*)
  - Additionally, equal intercepts across contexts
- **Strict invariance**
  - Additionally, equal residual variances across contexts

Compare each adjacent pair of models using likelihood ratio test (LRT; $\Delta\chi^2$) to test each omnibus H_0

Violations of Invariance

If omnibus H_0 is rejected, individual indicators can be tested for *differential item functioning* (DIF)

- Only *partial invariance* is necessary to compare latent distributions, but DIF weakens inferences

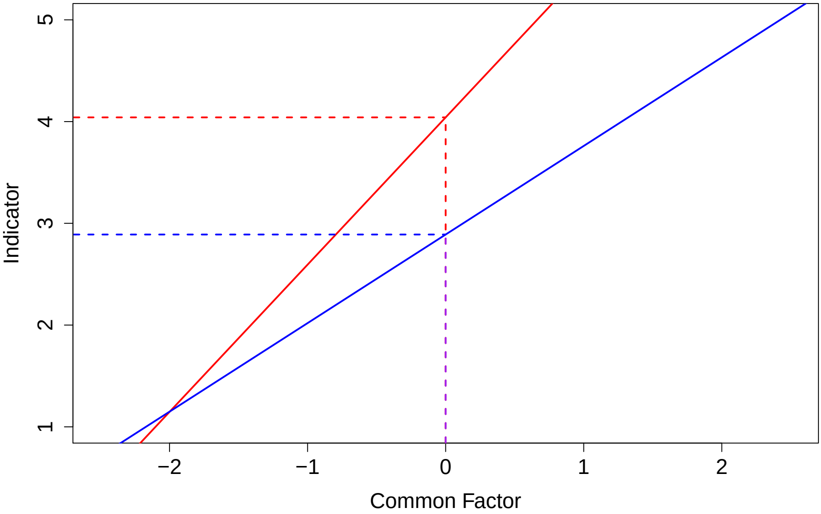
Analogous to regression of indicators (Y) on factors (η), potentially moderated by contextual variable (e.g., group G)

$$Y = \beta_0 + \beta_1 \eta + \beta_2 G + \beta_3 (\eta \times G) + \varepsilon$$

- Violation of metric invariance implies different loadings ($\beta_3 \neq 0$): *nonuniform* DIF
    - Don't constrain intercept for item with nonuniform DIF
- Violation of scalar invariance implies different intercepts ($\beta_3 = 0$, but $\beta_2 \neq 0$): *uniform* DIF

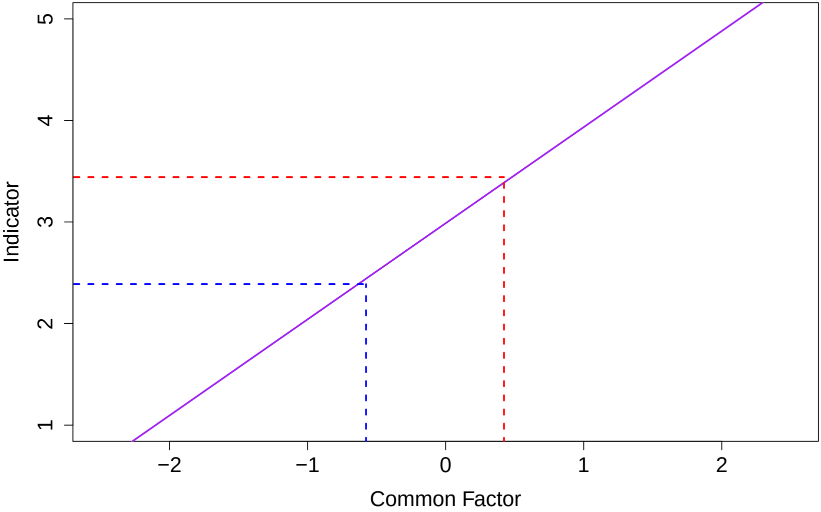
No Invariance Constraints (Each Latent $\mu = 0$)

Different indicator means, even holding factor constant



Invariance Holds

Different indicator means **must** be due to different factor means



Trading Constraints: Identification v. Invariance

Using a marker variable (or effects coding), parameters of latent distributions are already freely estimated without being constrained to equality across contexts

- conditional on equivalence of 1 (or average) loading & intercept
- not actually comparable until ≥ 2 indicators are constrained

Using the fixed-factor approach, latent distributions have fixed means and variances for identification

- When loadings are constrained, must free unnecessary identification constraints on latent variances
  - If not, test conflates equal loadings $+$ latent variances
- When intercepts are constrained, must free unnecessary identification constraints on latent means
  - If not, test conflates equal intercepts $+$ latent means

Item Factor Analysis

Extended Common Factor Model

Accommodate discrete data using a probit link

- Assume the linear effects in the common-factor model apply to multivariate normal **latent item-responses** (LIRs)
- LIRs linked to observed indicators via a **threshold model**

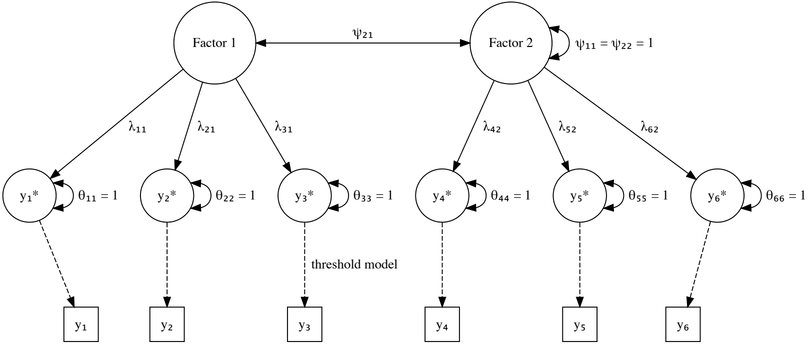


Figure 2: 2-factor CFA model with discrete indicators.

Item Factor Analysis

$$y^* = \nu + \lambda\eta + \theta$$

$$y = \begin{cases} 0 \text{ if } -\infty < y^* \leq \tau_1 \\ 1 \text{ if } \tau_1 < y^* \leq \tau_2 \\ \dots \\ k \text{ if } \tau_k < y^* \leq \infty \end{cases}$$

$$E(y^*) = \nu + \Lambda\alpha$$

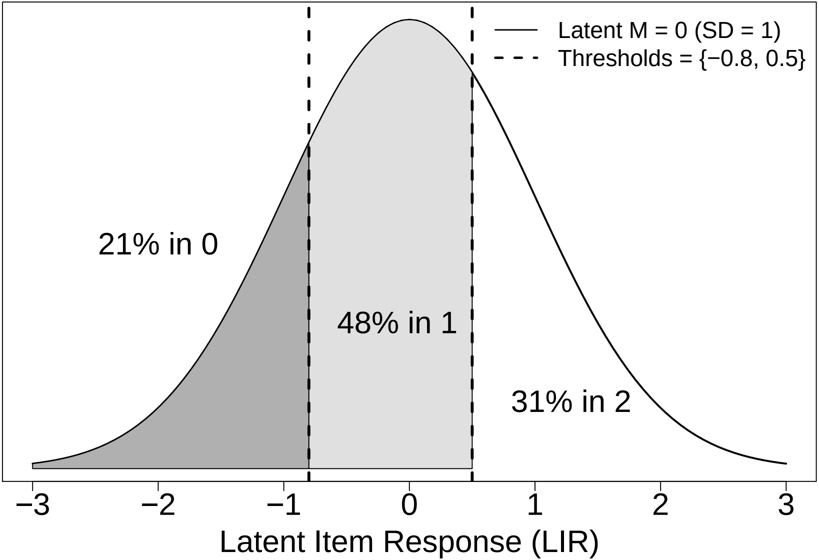
$$\text{Var}(y^*) = \Lambda\Psi\Lambda' + \Theta$$

Linear factor model reproduces the estimated **polychoric correlations** among the LIRs (y^*)

- Polychorics estimated assuming LIR $\mu = 0$ and $\sigma = 1$

Threshold Model for 3-Category Indicator

Threshold $=$ z score beyond which subjects enter a higher category



Interpret Thresholds

Suppose this 0–2 scale was used to measure perceived pain among patients with a chronic condition

- 0 (tolerable), 1 (uncomfortable), 2 (unbearable)

If this variable were treated as continuous, the mean would be a weighted sum of the values of the categories:

$$M = 0.21 \times 0 + 0.48 \times 1 + 0.31 \times 2 = 1.1$$

Imagine after experience with their condition, patients' threshold for pain increases, resulting in less perceived pain

- **Higher thresholds** imply **lower means** (more subjects in lower categories)

Interpret Thresholds

Treating the variable as ordinal, the probit model posits a normally distributed LIR underlying the observed discrete response

- If we fix the 2 thresholds (e.g., to 0 and 1, as in LISREL), we can estimate the LIR's μ and σ
- If we fix $\mu = 0$ and $\sigma = 1$ (as in M*plus* and `lavaan`), we can estimate the thresholds

These are alternative identification constraints, but statistically equivalent models

- i.e., same observed proportions in each category, implied by location(s) of threshold(s) in the latent distribution

Thresholds as a Measurement Model

In CFA, individual differences on latent factors are extrapolated from multiple observed indicators

- μ and σ are fixed or set relative to an indicator

Analogously, individual differences on an observed discrete indicator are used to extrapolate about individual differences on the underlying LIR

- Likewise, μ and σ are fixed or set relative to thresholds

The more categories in our scale, the more information we have about how individuals (co)vary on the underlying LIR(s)

- As in CFA, the number of parameters we can estimate is limited by the number of summary statistics we observed (i.e., thresholds)

Identifying the LIR Underlying a Binary Item

With ≥ 2 thresholds, we could estimate both μ and σ, but a binary item requires additional constraints for identification

- Similar to measuring a latent factor with only 1 or 2 indicators, there is not enough information from only 1 threshold to estimate both μ and σ

If we fix the threshold (e.g., $\tau = 0$), we can estimate **either** μ **or** σ, but not both

- Fixing either $\tau = 0$ or $\mu = 0$ to estimate the other parameter, the estimates would be the same magnitude but different signs
- To estimate σ, we must fix $\tau \neq \mu$ because σ is extrapolated from the chosen distance between τ and μ

Numerical Examples

Generate Longitudinal Items from a Bivariate LIR

```
N <- 1000
# mean difference = 0.8
mu <- c(wave1 = 0, wave2 = 0.8)
# autocorrelation = 0.7 / sqrt(2) = 0.495
Sigma <- matrix(c(1, .7, .7, 2), nrow = 2)
set.seed(123)
dat <- data.frame(MASS::mvrnorm(N, mu, Sigma))
# binary (1 threshold)
dat$y2w1 <- as.numeric(dat$wave1 > -0.5)
dat$y2w2 <- as.numeric(dat$wave2 > -0.5)
# ternary (2 thresholds)
dat$y3w1 <- dat$y2w1 + (dat$wave1 > 0.5)
dat$y3w2 <- dat$y2w2 + (dat$wave2 > 0.5)
# polytomous (3 thresholds)
dat$y4w1 <- dat$y3w1 + (dat$wave1 > 1)
dat$y4w2 <- dat$y3w2 + (dat$wave2 > 1)
```

Specify Model Parameters for Binary Items

```
library(lavaan)
mod2 <- '
## LIR means
  y2w1 ~ mean1*1
  y2w2 ~ mean2*1
## LIR (co)variances
  y2w1 ~~ var1*y2w1 + y2w2
  y2w2 ~~ var2*y2w2
## thresholds link LIRs to observed items
  y2w1 | thr1*t1
  y2w2 | thr2*t1
'
```

Each parameter is labeled so that different identification constraints can be applied when fitting the model to data

Fit Model with Standard-Normal LIR Distributions

The default method in M*plus* and `lavaan` is to fix the LIR intercepts to 0 and the *marginal* (total) LIR variance to 1

- The default is called the "delta" parameterization
  - Δ scaling factor = the reciprocal of LIR's *SD*
- Request the alternative "theta" parameterization to fix the *conditional* (residual) LIR variance to 1

```
constr2z <- '
## Wave 1
  mean1 == 0 ; var1 == 1
## Wave 2
  mean2 == 0 ; var2 == 1
'
fit2z <- lavaan(mod2, data = dat, constraints = constr2z,
                ordered = c("y2w1","y2w2"),
                parameterization = "theta")
```

Fit Model with Standard-Normal LIR Distributions

```
##
## Intercepts:
##                    Estimate
##     y2w1    (men1)    0.000
##     y2w2    (men2)    0.000
##
## Thresholds:
##                    Estimate
##     y2w1|t1 (thr1)   -0.468
##     y2w2|t1 (thr2)   -0.927
##
## Variances:
##                    Estimate
##     y2w1    (var1)    1.000
##     y2w2    (var2)    1.000
```

Fix Thresholds to $\tau = 0$ Instead

In `lavaan` it is possible to free the LIR intercepts

- In M*plus* phantom constructs would need to be specified

```
constr2t <- '
## Wave 1
  thr1 == 0 ; var1 == 1
## Wave 2
  thr2 == 0 ; var2 == 1
'
fit2t <- lavaan(mod2, data = dat, constraints = constr2t,
                ordered = c("y2w1","y2w2"),
                parameterization = "theta")
```

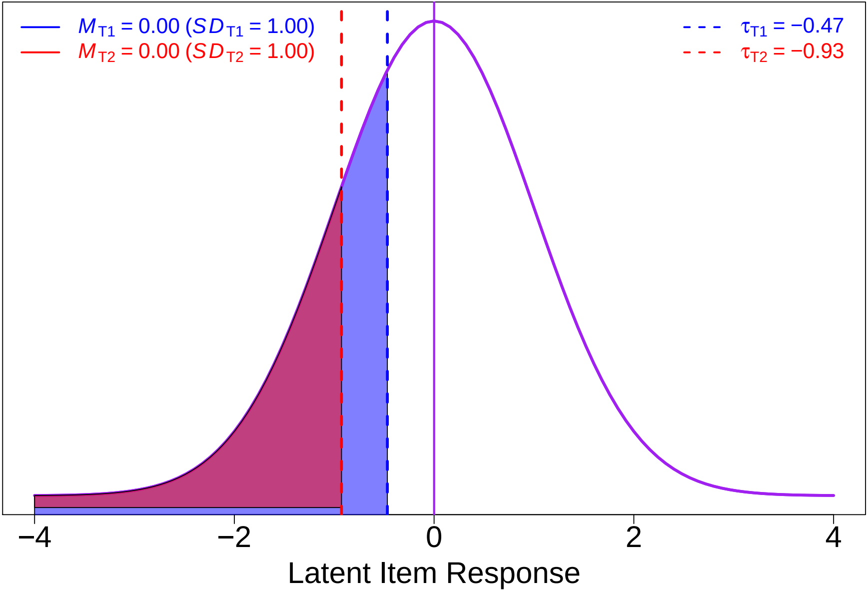
Notice the estimated means are the same magnitude (but opposite sign) as the previously estimated thresholds

- Recall these are statistically equivalent interpretations

Fix Thresholds to $\tau = 0$ Instead

```
##
## Intercepts:
##                    Estimate
##     y2w1    (men1)    0.468
##     y2w2    (men2)    0.927
##
## Thresholds:
##                    Estimate
##     y2w1|t1 (thr1)    0.000
##     y2w2|t1 (thr2)    0.000
##
## Variances:
##                    Estimate
##     y2w1    (var1)    1.000
##     y2w2    (var2)    1.000
```

Compare Fixed Means to Fixed Thresholds



Equality Constraints on Thresholds

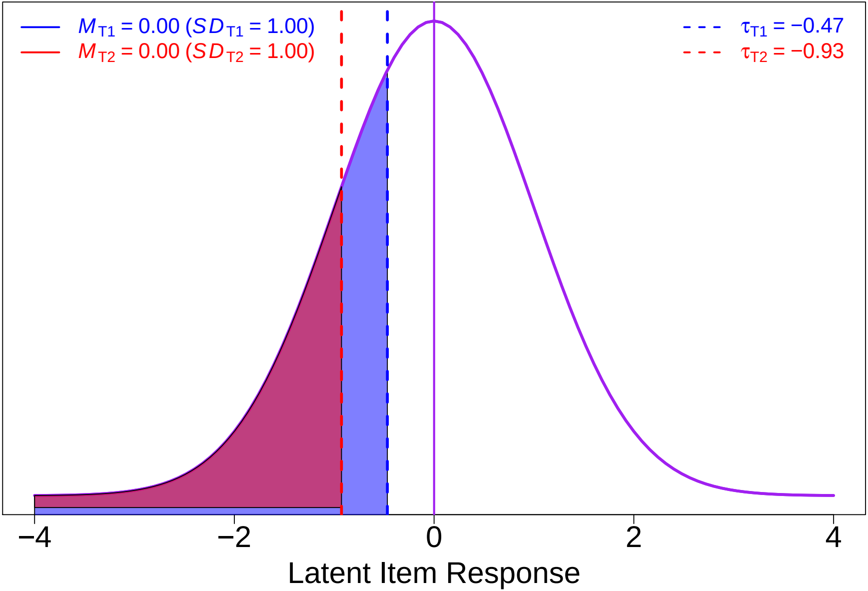
Fixing an LIR's $\mu = 0$ and $\sigma = 1$ is analogous to fixed-factor identification constraints. Imposing equality constraints on thresholds (the LIR's measurement model) would only represent a hypothesis of *measurement* equivalence if any unnecessary identification constraint(s) were freed.

```
constr2e <- '
thr1 == thr2  # measurement equivalence constraint
mean1 == 0    # identification constraints at Wave 1
var1 == 1
var2 == 1     # identification constraint  at Wave 2
'
fit2e <- lavaan(mod2, data = dat, constraints = constr2e,
                ordered = c("y2w1","y2w2"),
                parameterization = "theta")
```

Equality Constraints on Thresholds

```
##
## Intercepts:
##                    Estimate
##     y2w1    (men1)   -0.000
##     y2w2    (men2)    0.459
##
## Thresholds:
##                    Estimate
##     y2w1|t1 (thr1)   -0.468
##     y2w2|t1 (thr2)   -0.468
##
## Variances:
##                    Estimate
##     y2w1    (var1)    1.000
##     y2w2    (var2)    1.000
```

Interpret Time-2 Mean Relative to Wave 1



Fix Thresholds **and** Means to Estimate Variances

We have seen it is equivalent to interpret the difference in observed distributions (fewer zeros / more ones at Wave 2) as:

- A **lower threshold** or a **higher mean** at Wave 2

With only 2 categories, another equivalent interpretation would be that subjects are **more homogenous** at Wave 2

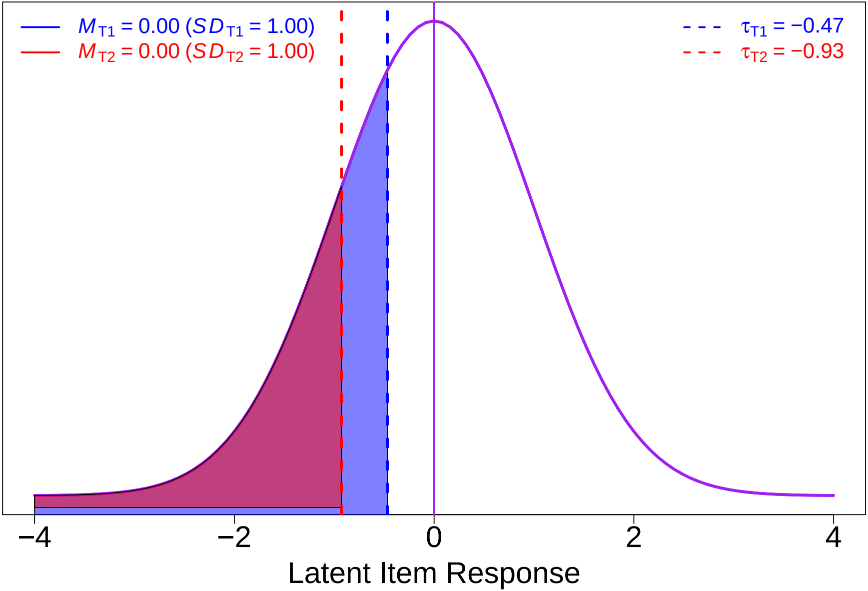
- Fix both τ and μ to different arbitrary values to estimate σ^2

```
constr2v <- '
## Wave 1
  thr1 == -0.5 ; mean1 == 0.5
## Wave 2
  thr2 == -0.5 ; mean2 == 0.5
'
fit2v <- lavaan(mod2, data = dat, constraints = constr2v,
                ordered = c("y2w1","y2w2"),
                parameterization = "theta")
```

Fix Thresholds **and** Means to Estimate Variances

```
##
## Intercepts:
##                    Estimate
##     y2w1    (men1)    0.500
##     y2w2    (men2)    0.500
##
## Thresholds:
##                    Estimate
##     y2w1|t1 (thr1)   -0.500
##     y2w2|t1 (thr2)   -0.500
##
## Variances:
##                    Estimate
##     y2w1    (var1)    4.572
##     y2w2    (var2)    1.164
```

Compare Estimated Thresholds to Estimated Variances



Specify Model Parameters for Ternary Items

The same principles apply to items with > 2 categories, but there are enough thresholds to trade for both μ and σ

```
mod3 <- '
## LIR means
  y3w1 ~ mean1*1
  y3w2 ~ mean2*1
## LIR (co)variances
  y3w1 ~~ var1*y3w1 + y3w2
  y3w2 ~~ var2*y3w2
## thresholds link LIRs to observed items
  y3w1 | thr1.w1*t1 + thr2.w1*t2
  y3w2 | thr1.w2*t1 + thr2.w2*t2
'
```

Fit Models with Free and Equated Thresholds

```
constr3z <- '
mean1 == 0 ; var1 == 1  # Wave 1 identification constraints
mean2 == 0 ; var2 == 1  # Wave 2 identification constraints
'
fit3z <- lavaan(mod3, data = dat, constraints = constr3z,
                ordered = c("y3w1","y3w2"),
                parameterization = "theta")
constr3e <- '
## Wave 1 identification constraints
  mean1 == 0 ; var1 == 1
## measurement equivalence constraints
  thr1.w1 == thr1.w2
  thr2.w1 == thr2.w2
'
fit3e <- lavaan(mod3, data = dat, constraints = constr3e,
                ordered = c("y3w1","y3w2"),
                parameterization = "theta")
```

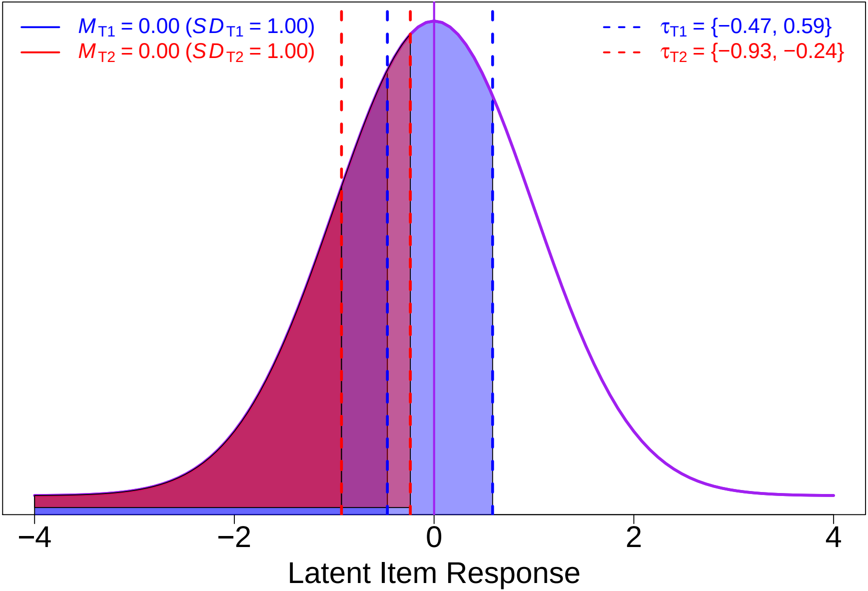
Free Thresholds (Standard-Normal LIRs)

```
##
## Intercepts:
##                    Estimate
##     y3w1    (men1)    0.000
##     y3w2    (men2)    0.000
##
## Thresholds:
##                    Estimate
##     y3w1|t1 (t1.1)   -0.468
##     y3w1|t2 (t2.1)    0.586
##     y3w2|t1 (t1.2)   -0.927
##     y3w2|t2 (t2.2)   -0.238
##
## Variances:
##                    Estimate
##     y3w1    (var1)    1.000
##     y3w2    (var2)    1.000
```

Equated Thresholds to Estimate μ and σ^2 at Wave 2

```
##
## Intercepts:
##                    Estimate
##      y3w1    (men1)    0.000
##      y3w2    (men2)    0.949
##
## Thresholds:
##                    Estimate
##      y3w1|t1 (t1.1)   -0.468
##      y3w1|t2 (t2.1)    0.586
##      y3w2|t1 (t1.2)   -0.468
##      y3w2|t2 (t2.2)    0.586
##
## Variances:
##                    Estimate
##      y3w1    (var1)    1.000
##      y3w2    (var2)    2.338
```

Interpret Time-2 Distribution Relative to Wave 1



Specify Model Parameters for 4-Category Items

When we can constrain > 2 thresholds for an item:

- We can estimate both μ and σ
- There are *df* left over to **test** H_0 that thresholds are equivalent (without *additionally* assuming equivalence of μ and σ)

```
mod4 <- '
## LIR means
  y4w1 ~ mean1*1
  y4w2 ~ mean2*1
## LIR (co)variances
  y4w1 ~~ var1*y4w1 + y4w2
  y4w2 ~~ var2*y4w2
## thresholds link LIRs to observed items
  y4w1 | thr1.w1*t1 + thr2.w1*t2 + thr3.w1*t3
  y4w2 | thr1.w2*t1 + thr2.w2*t2 + thr3.w2*t3
'
```

Fit Models with Free and Equated Thresholds

```
constr4z <- '
mean1 == 0 ; var1 == 1  # Wave 1 identification constraints
mean2 == 0 ; var2 == 1  # Wave 2 identification constraints
'
fit4z <- lavaan(mod4, data = dat, constraints = constr4z,
                ordered = c("y4w1","y4w2"),
                parameterization = "theta")
constr4e <- ' ## Wave 1 identification constraints
  mean1 == 0 ; var1 == 1
## measurement equivalence constraints
  thr1.w1 == thr1.w2
  thr2.w1 == thr2.w2
  thr3.w1 == thr3.w2
'
fit4e <- lavaan(mod4, data = dat, constraints = constr4e,
                ordered = c("y4w1","y4w2"),
                parameterization = "theta")
```

Compare Models with Free and Equated Thresholds

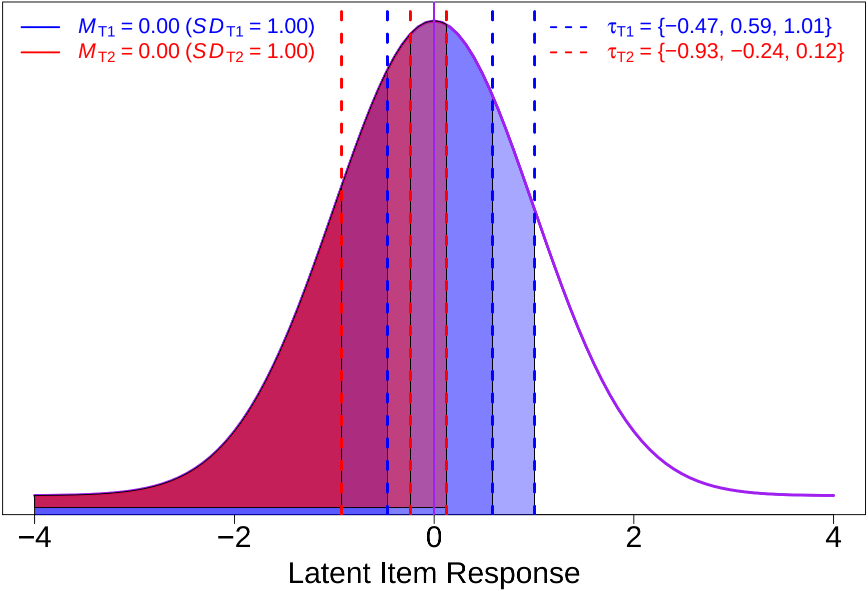
These models are not statistically equivalent

- The equal-thresholds model is more restrictive than the standard-normal model

```
lavTestLRT(fit4z, fit4e)
```

```
## Scaled Chi-Squared Difference Test (method = "satorra.20
##
## lavaan NOTE:
##     The "Chisq" column contains standard test statistics
##     robust test that should be reported per model. A rob
##     test is a function of two standard (not robust) stat
##
##       Df AIC BIC Chisq Chisq diff Df diff Pr(>Chisq)
## fit4z  0          0.00
## fit4e  1          0.91       3.46       1      0.063 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
```

Compare Models with Free and Equated Thresholds



Implications for Measurement Equivalence with *Multiple* Discrete Indicators

Previous Advice: Millsap & Tein (2004)

When LIRs are endogenous (e.g., indicators of common factors)

- μ and σ of each LIR becomes conditional on common factor
- i.e., an intercept ν and residual variance θ

Minimal identification constraints for configural model:

- Equate 1 threshold per item to identify residual variances
  - Fix residual variances to 1 in reference group
- Fix loading of marker variable to identify factor variance
- All intercepts fixed to zero
- Equate additional threshold of marker variable to identify common factor means
  - Fix common factor mean to 0 in reference group
  - If binary, instead fix $\theta = 1$ for all groups

Test invariance only by imposing equality constraints

Issues with Millsap & Tein's (2004) Advice

Handles all identification constraints as a single set

- More complicated than recognizing identification of LIR and common factor scales/locations as independent issues

Equality of loadings tested on the assumption of equal intercepts

- Violates hierarchical principle

Equality of intercepts assumed (never tested)

- Conflates equivalence of thresholds with strong invariance
- If equivalence of thresholds fails, it might only represent a constant shift in all thresholds (equivalent to different intercepts)

Previous Advice: Software Defaults

Researchers often rely on software defaults:

- LISREL fixes first 2 thresholds to 0 and 1, estimates ν and θ
  - For binary items, fix $\tau = 0$ and $\theta = 1$
  - Any estimated thresholds **assumed equal** across groups
- M*plus* fixes $\nu = 0$ and $\theta = 1$, estimates all thresholds
  - Analogous to **fixed-factor** approach
  - By default, equates thresholds across groups, fixes $\theta = 1$ only in reference group, so users must manually specify configural models
- `lavaan` defaults are similar to M*plus*, but without assuming equivalence across groups

Issues with Software Defaults

These are statistically equivalent approaches *only without equality constraints* imposed across contexts (e.g., groups or occasions)

- Thus, LISREL and M*plus* configural models will not be equivalent when items have > 3 categories

Some levels of invariance cannot be tested

- thresholds in LISREL
- intercepts in M*plus*, unless phantom constructs are specified for each LIR

Tricking the software is tedious, prone to human error

- See `semTools::measEq.syntax()` R function for help

Alternative Perspective

If we view a LIR as a *first-order* latent variable . . .

- measured by a single discrete item
- its thresholds are measurement-model parameters

. . . we can consider identification constraints *separately* for LIRs and common factors. Test equivalence of thresholds *first*

- Analogously, Chen et al. (2005) recommended testing equivalence of first-order loadings before second-order loadings
  - https://doi.org/10.1207/s15328007sem1203_7

Multiple LIRs then measure (higher-order) common factor(s)

- Test remaining measurement parameters in the usual sequence
- Same considerations about trading identification vs. invariance constraints

Trading Constraints: Identification v. Invariance

Concurs with Wu & Estabrook (2016), who recommended **freeing unnecessary identification constraints** when any measurement parameters were equated to test equivalence

- Otherwise, tests make overly strict assumptions about other parameters (e.g., equivalence of ν and θ)

Step 1: Test equivalence of thresholds

- Free identification constraints ($\nu = 0$ and $\theta = 1$), except for 1 reference category (e.g., first group/occasion)
  - analogous to **fixed factor** approach
- Requires ≥ 4 categories to do more than trade *df*
  - 2 thresholds used to identify LIR's ν and θ
  - any extras buy *df* when constrained to equality
- If binary: free ν, keep $\theta = 1$

Trading Constraints: Identification v. Invariance

Step 2: Test equivalence of loadings

- Free identification constraints on common factor variances, except for 1 reference category

Step 3: Test equivalence of intercepts

- Free identification constraints on common factor means, except for 1 reference category
- $\nu = 0$ in reference category for identification, so impose equivalence by fixing $\nu = 0$ in all contexts
- If binary: θ can now be freed (less constrained than Step 1)

Step 4: Test equivalence of residual variances

- $\theta = 1$ in reference category for identification, so impose equivalence by fixing $\theta = 1$ in all contexts

Special Considerations about Equivalent Models

Ternary items: Trade only 2 thresholds for ν and θ

- Threshold invariance is equivalent to configural invariance
- **Solution**: Consider threshold invariance to be Step 1

Binary items: Only 1 threshold to constrain

- First, free only ν
  - Equivalent to configural model
- After constraining λ then ν, free θ (*and* factor mean)
  - Less constrained than threshold invariance!
- **Solution**: Simultaneously constrain thresholds, loadings, and intercepts (and free θ) to test strong (vs. configural) invariance
  - Cannot distinguish between sources of violation

`semTools::measEq.syntax()` Examples

Install and Load R Package

```r
# stable version on CRAN
install.packages("semTools")

# development version on GitHub
devtools::install_github("simsem/semTools/semTools")

# load package
library(semTools)

# find help-page examples
?measEq.syntax
```

Thank you for your attention

Questions?

Some suggested reading:

- Muthen & Asparouhov (2002) Mplus web note 4
  - www.statmodel.com/download/webnotes/CatMGLong.pdf
- Mehta & Neale (2004)
  - http://dx.doi.org/10.1037/1082-989X.9.3.301
- Millsap & Tein (2004)
  - http://dx.doi.org/10.1207/S15327906MBR3903_4
- Kamata & Bauer (2008)
  - http://dx.doi.org/10.1080/10705510701758406
- **Wu & Estabrook (2016)**
  - http://dx.doi.org/10.1007/s11336-016-9506-0