

Statistical Analysis Using Structural Equation Models

EPsy 8266

Christopher David Desjardins

Research Methodology Consulting Center

4/9/19

Motivation

- ▶ You measure executive functioning (EF) in two groups of children:
a) children living in homeless shelters and b) children not living in homeless shelters.
- ▶ You administer several EF tasks (e.g., Stroop, peg tapping)
- ▶ Ultimately, you'd like to be able to test if children not living in homeless shelters have greater EF than children in homeless shelters.
- ▶ Can you just sum up the scores on the tasks and compare the children using a t-test? Why? Why not?

Motivation

- ▶ You want to measure problem behaviors from childhood to early adulthood.
- ▶ You collect data on problem behaviors at 7 years, 10 years, 14 years, 18, and 22 years.
- ▶ Ultimately, you'd like to examine how problem behaviors develop from childhood to adulthood (does it decrease? increase? is it linear? nonlinear?)
- ▶ Can you just fit a latent growth curve model?

Multi-samples SEM

- ▶ In a multi-samples SEM, we want to know do the parameters of interest vary appreciable across samples.
- ▶ What are these *parameters*?

Multi-samples SEM

- ▶ In a multi-samples SEM, we want to know do the parameters of interest vary appreciable across samples.
- ▶ What are these *parameters*?
- ▶ This is analogous to whether group (living in a homeless shelter or not) or time (problem behaviors at 7, 10, 14, 18, and 22) **moderate** the relations in our SEM.
- ▶ This means that the group/time affects the estimated parameters of our model and we can't collapse over group/time and these parameters must be estimated separately by group/time.

How might you examine this?

One approach

- ▶ For each sample (group or time), fit a separate SEM model.
- ▶ Compare the unstandardized pattern and structural coefficients.
- ▶ If the values are quite different, conclude that these parameters are not the same.

A better approach

- ▶ Use a single SEM to estimate the samples simultaneously (i.e., that estimates all the parameters for all the samples).
- ▶ Progressively constrain unstandardized parameters systematically across the groups to test if equality reduces model fit (**free baseline approach**).
- ▶ Compare the models using chi-square test of difference, RMSEA, and/or CFI
- ▶ If constraining doesn't deteriorate fit significantly (or appreciably) conclude that parameters could be equal in the population.

Multiple-samples CFA

Goal: To assess whether the suite indicators are **invariant**.

How:

1. Specify a measurement model (CFA) with a mean structure,
2. Constrain to equality over the samples the unstandardized parameters,
3. Compare a series of models with differing levels of constraints for the samples.

Measurement invariance

- ▶ **Measurement invariance (MI)** - scores from the way a construct is operationalized have the same meaning under different conditions.
- ▶ Conditions could be time (month 12, month 24, and so on), methods (paper/pencil vs. Internet-delivered), or group memberships (USA sample vs. Russia sample).
- ▶ **Longitudinal measurement invariance** - refers specifically to time. The methods are similar.
 - ▶ Without this can't ascertain if change over time is because of real change on the factor or the factor changed.
 - ▶ Using the same test over multiple occasions doesn't guarantee longitudinal MI.
 - ▶ See **Little's** or **Newsom's** longitudinal book.

Types of MI

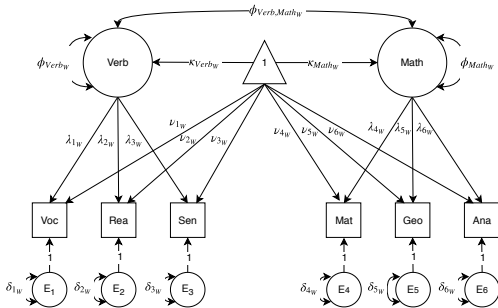
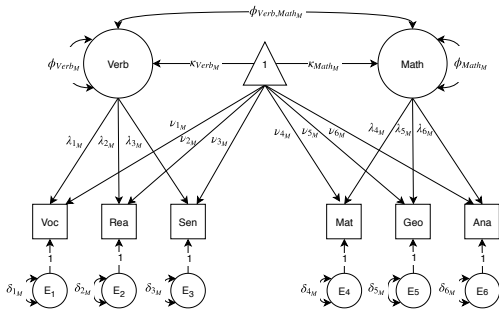
Pre-MI - Does your proposed factor structure actually fit? If yes, then proceed ...

- ▶ Configural - Does the same number of factors and relationship between factors and indicators hold in each group?
 - ▶ No parameters are constrained.
- ▶ Weak - Assuming configural invariance holds, constrain the unstandardized pattern coefficients across the groups.
- ▶ Strong - Assuming weak invariance holds, constrain the unstandardized intercepts.
- ▶ Strict - Assuming strong invariance holds, constrain the error variances and covariances across the groups.

Interest inventory

- ▶ Data from a cognitive, personality, and vocational interest inventory
- ▶ 33 variables on 250 participants.
- ▶ vocab (Vocabulary test), reading (Reading comprehension), sentcomp (Sentence completion) - verbal IQ
- ▶ mathmtcs (Mathematics), geometry (Geometry), analyrea (Analytical reasoning) - math IQ
- ▶ sex: coded 1 for male and 2 for female
- ▶ educ: Years of education
- ▶ age: Age, in years

Are there sex differences in verbal and math IQ?



Male Model

$$Voc = \nu_{1_M} + \lambda_{1_M} Verb + \delta_{1_M}$$

$$Rea = \nu_{2_M} + \lambda_{2_M} Verb + \delta_{2_M}$$

$$Sen = \nu_{3_M} + \lambda_{3_M} Verb + \delta_{3_M}$$

$$Mat = \nu_{4_M} + \lambda_{4_M} Math + \delta_{4_M}$$

$$Geo = \nu_{5_M} + \lambda_{5_M} Math + \delta_{5_M}$$

$$Ana = \nu_{6_M} + \lambda_{6_M} Math + \delta_{6_M}$$

$$\phi Verb_M, \phi Math_M, \phi Verb, Math_M, \kappa Verb_M, \kappa Math_M$$

Female Model

$$Voc = \nu_{1_W} + \lambda_{1_W} Verb + \delta_{1_W}$$

$$Rea = \nu_{2_W} + \lambda_{2_W} Verb + \delta_{2_W}$$

$$Sen = \nu_{3_W} + \lambda_{3_W} Verb + \delta_{3_W}$$

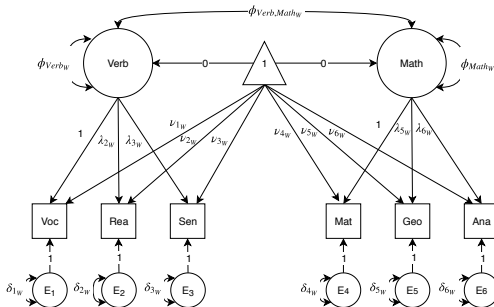
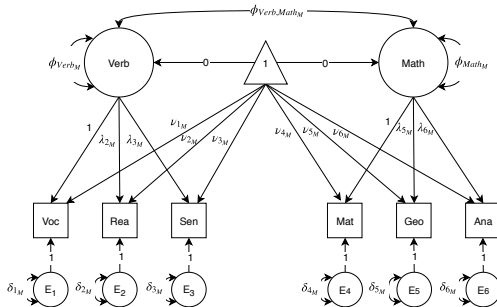
$$Mat = \nu_{4_W} + \lambda_{4_W} Math + \delta_{4_W}$$

$$Geo = \nu_{5_W} + \lambda_{5_W} Math + \delta_{5_W}$$

$$Ana = \nu_{6_W} + \lambda_{6_W} Math + \delta_{6_W}$$

$$\phi Verb_W, \phi Math_W, \phi Verb, Math_W, \kappa Verb_W, \kappa Math_W$$

- ▶ Is this model identified?
- ▶ How many unique elements do we have in our covariance matrix? In the mean matrix?
- ▶ How many by sex?
- ▶ Does it have a scale?
- ▶ By default for a multi-sample model, lavaan fixes the first indicator per factor to 1 (scale) and the factor mean to 0 (to identify the means).



What is the implicit assumption here?

Step 1: Configural model

```
library(lavaan)
interest <- read.csv("https://bit.ly/2IF3Hcm")
interest$sex <- ifelse(interest$sex == 2, "female", "male")
config.mod <- "
verb =~ vocab + reading + sentcomp
math =~ mathmtcs + geometry + analyrea
"

config.fit <- cfa(config.mod, data = interest, group = "sex")
fitMeasures(config.fit, c("chisq", "df", "pvalue", "cfi", "rmsea"))

##   chisq      df pvalue    cfi  rmsea
## 18.413 16.000  0.300  0.998  0.035
```

Step 1: Configural model

```
summary(config.fit, fit.measure = TRUE)
```

```
##      Number of free parameters              38
##
##      Number of observations per group
##      male                                122
##      female                              128
##
##      Estimator                             ML
##      Model Fit Test Statistic              18.413
##      Degrees of freedom                    16
##      P-value (Chi-square)                  0.300
##
##      Chi-square for each group:
##
##      female                                7.908
##      male                                  10.505
##
##      Model test baseline model:
##
##      Minimum Function Test Statistic       1303.205
##      Degrees of freedom                     30
##      P-value                               0.000
##
##      Comparative Fit Index (CFI)           0.998
##      RMSEA                                 0.035
##      90 Percent Confidence Interval         0.000  0.093
##
##      SRMR                                  0.014
```

Step 1: Configural model - female

```
## Group 1 [female]:  
##  
## Latent Variables:  
##  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   verb =~  
##     vocab           1.000  
##     reading         0.942     0.068    13.820     0.000  
##     sentcomp        0.951     0.070    13.514     0.000  
##   math =~  
##     mathmtcs        1.000  
##     geometry        0.857     0.069    12.462     0.000  
##     analyrea        0.953     0.068    13.936     0.000  
##  
## Covariances:  
##  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   verb ~~  
##     math           0.611     0.096     6.341     0.000  
##  
## Intercepts:  
##  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   .vocab           0.025     0.082     0.301     0.764  
##   .reading          0.054     0.086     0.625     0.532  
##   .sentcomp        -0.005     0.088    -0.056     0.955  
##   .mathmtcs         0.274     0.087     3.131     0.002  
##   .geometry         0.219     0.087     2.522     0.012  
##   .analyrea         0.275     0.091     3.018     0.003  
##   verb             0.000  
##   math             0.000
```

Step 1: Configural model - female

Variances:

##		Estimate	Std.Err	z-value	P(> z)
##	.vocab	0.098	0.031	3.178	0.001
##	.reading	0.269	0.043	6.251	0.000
##	.sentcomp	0.296	0.046	6.402	0.000
##	.mathmtcs	0.121	0.037	3.295	0.001
##	.geometry	0.334	0.050	6.673	0.000
##	.analyrea	0.281	0.048	5.913	0.000
##	verb	0.759	0.110	6.897	0.000
##	math	0.858	0.126	6.813	0.000

Step 1: Configural model - male

```
## Group 2 [male]:  
##  
## Latent Variables:  
##  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   verb =~  
##     vocab           1.000  
##     reading         0.856     0.057    14.994     0.000  
##     sentcomp        0.860     0.054    15.942     0.000  
##   math =~  
##     mathmtcs        1.000  
##     geometry        0.870     0.065    13.415     0.000  
##     analyrea        0.935     0.059    15.937     0.000  
##  
## Covariances:  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   verb ~~  
##     math           0.913     0.133     6.884     0.000  
##  
## Intercepts:  
##           Estimate   Std.Err   z-value   P(>|z|)  
##   .vocab           0.159     0.096     1.654     0.098  
##   .reading          0.220     0.091     2.433     0.015  
##   .sentcomp         0.156     0.089     1.755     0.079  
##   .mathmtcs        -0.071     0.098    -0.724     0.469  
##   .geometry         0.001     0.097     0.008     0.993  
##   .analyrea        0.070     0.097     0.725     0.468  
##   verb             0.000  
##   math             0.000
```

Step 1: Configural model - male

Variances:

##		Estimate	Std.Err	z-value	P(> z)
##	.vocab	0.119	0.031	3.835	0.000
##	.reading	0.262	0.041	6.454	0.000
##	.sentcomp	0.218	0.036	6.134	0.000
##	.mathmtcs	0.144	0.035	4.105	0.000
##	.geometry	0.360	0.054	6.699	0.000
##	.analyrea	0.240	0.041	5.792	0.000
##	verb	1.007	0.146	6.900	0.000
##	math	1.036	0.153	6.774	0.000

Step 1: Configural model - default lavaan method

```
config.mod.lav <- "  
# define factor/unstd. pattern coef  
verb =~ c(1, 1)*vocab + reading + sentcomp  
math =~ c(1, 1)*mathmtcs + geometry + analyrea  
  
# factor intercepts  
verb ~ c(0,0)*1  
math ~ c(0,0)*1  
  
# factor variances/covariance  
verb ~~ verb + math  
math ~~ math  
  
# intercepts  
vocab ~ 1  
reading ~ 1  
sentcomp ~ 1  
mathmtcs ~ 1  
geometry ~ 1  
analyrea ~ 1  
  
# residual variances  
vocab ~~ vocab  
reading ~~ reading  
sentcomp ~~ sentcomp  
mathmtcs ~~ mathmtcs  
geometry ~~ geometry  
analyrea ~~ analyrea  
"
```

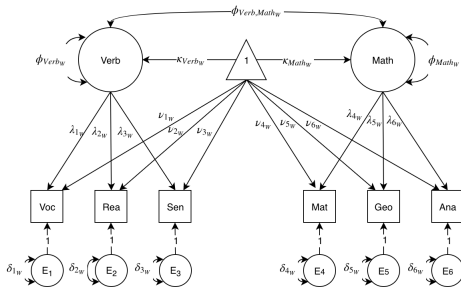
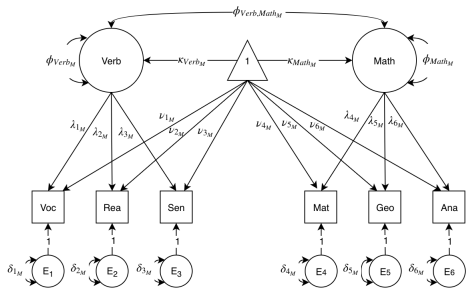

Step 1: Configural model - marker variable approach

```
config.mod.mark <- "  
# define factor/unstd. pattern coef  
verb  =~ c(1, 1)*vocab + reading + sentcomp  
math  =~ c(1, 1)*mathmtcs + geometry + analyrea  
  
# factor intercepts  
verb  ~ 1  
math  ~ 1  
  
# factor variances/covariance  
verb  ~~ verb + math  
math  ~~ math  
  
# intercepts  
vocab ~ c(0,0)*1  
reading ~ 1  
sentcomp ~ 1  
mathmtcs ~ c(0,0)*1  
geometry ~ 1  
analyrea ~ 1  
  
# residual variances  
vocab ~~ vocab  
reading ~~ reading  
sentcomp ~~ sentcomp  
mathmtcs ~~ mathmtcs  
geometry ~~ geometry  
analyrea ~~ analyrea  
"
```

A non-arbitrary approach to identifying a multi-samples CFA

Little et al., (2006) proposed the effects coding method.

- ▶ Constraining the average unstandardized pattern coefficient to 1.0 for each factor
 - ▶ $\sum_{i=1}^I \lambda_{if}^g = 1$. g = group, f = factor, i = indicator, I = # of indicators.
- ▶ Constraining the average intercept of the same indicators to 0 for each factor
 - ▶ $\sum_{i=1}^I \nu_{if}^g = 0$.
- ▶ This method results in estimates of the **latent variances that are the average of the indicators' variances accounted for by the construct**.
- ▶ **Latent means are estimated as optimally weighted averages** of the set of indicator means for a given construct.
- ▶ The estimated latent variances and means reflect the observed metric of the indicators, optimally weighted by the degree to which each indicator represents the underlying latent construct.



Step 1: Configural model - effects coding

```
config.mod <- "  
# define factor/unstd. pattern coef  
verb ~ c(lam1m, lam1w)*vocab + c(lam2m, lam2w)*reading + c(lam3m, lam3w)*sentcomp  
math ~ c(lam4m, lam4w)*mathmtcs + c(lam5m, lam5w)*geometry + c(lam6m, lam6w)*analyrea  
  
# factor intercepts  
verb ~ 1  
math ~ 1  
  
# factor variances/covariance  
verb ~~ verb + math  
math ~~ math  
  
# intercepts  
vocab ~ c(nu1m, nu1w)*1  
reading ~ c(nu2m, nu2w)*1  
sentcomp ~ c(nu3m, nu3w)*1  
mathmtcs ~ c(nu4m, nu4w)*1  
geometry ~ c(nu5m, nu5w)*1  
analyrea ~ c(nu6m, nu6w)*1  
  
# residual variances  
vocab ~~ c(e1m, e1w)*vocab  
reading ~~ c(e2m, e2w)*reading  
sentcomp ~~ c(e3m, e3w)*sentcomp  
mathmtcs ~~ c(e4m, e4w)*mathmtcs  
geometry ~~ c(e5m, e5w)*geometry  
analyrea ~~ c(e6m, e6w)*analyrea  
  
# define constraints  
lam1m == 3 - lam2m - lam3m  
lam1w == 3 - lam2w - lam3w  
lam4m == 3 - lam5m - lam6m  
lam4w == 3 - lam5w - lam6w  
nu1m == 0 - nu2m - nu3m  
nu1w == 0 - nu2w - nu3w  
nu4m == 0 - nu5m - nu6m  
nu4w == 0 - nu5w - nu6w  
"  
config.fit <- lavaan(config.mod, data = interest, group = "sex")
```

```
config.fit <- lavaan(config.mod, data = interest, group = "sex")
summary(config.fit)
```

```
## Group 1 [female]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb =				
##	vocab (lm1w)	1.037	0.040	25.627	0.000
##	reading (lm2w)	0.977	0.046	21.341	0.000
##	sentcmp (lm3w)	0.986	0.047	21.055	0.000
##	math =				
##	mthmtcs (lm4w)	1.067	0.042	25.117	0.000
##	geomtry (lm5w)	0.915	0.048	18.878	0.000
##	analyre (lm6w)	1.018	0.046	21.923	0.000

```
##
```

```
## Intercepts:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb	0.024	0.078	0.313	0.754
##	math	0.256	0.081	3.168	0.002
##	.vocab (nu1w)	-0.001	0.029	-0.025	0.980
##	.reading (nu2w)	0.030	0.036	0.832	0.405
##	.sentcmp (nu3w)	-0.029	0.037	-0.790	0.429
##	.mthmtcs (nu4w)	0.001	0.033	0.023	0.982
##	.geomtry (nu5w)	-0.015	0.041	-0.368	0.713
##	.analyre (nu6w)	0.014	0.039	0.371	0.711

```
config.fit <- lavaan(config.mod, data = interest, group = "sex")
summary(config.fit)
```

```
## Group 2 [male]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb = ~				
##	vocab (1m1m)	1.105	0.037	30.216	0.000
##	reading (1m2m)	0.945	0.041	22.834	0.000
##	sentcmp (1m3m)	0.950	0.039	24.114	0.000
##	math = ~				
##	mthmtcs (1m4m)	1.070	0.038	27.944	0.000
##	geomtry (1m5m)	0.931	0.046	20.435	0.000
##	analyre (1m6m)	1.000	0.041	24.488	0.000

```
##
```

```
## Intercepts:
```

```
##
```

##	.vocab (nu1m)	-0.038	0.031	-1.226	0.220
##	.reading (nu2m)	0.052	0.037	1.413	0.158
##	.sentcmp (nu3m)	-0.014	0.035	-0.391	0.696
##	.mthmtcs (nu4m)	-0.071	0.033	-2.132	0.033
##	.geomtry (nu5m)	0.001	0.041	0.021	0.983
##	.analyre (nu6m)	0.070	0.037	1.926	0.054

```
##
```

```
## Constraints:
```

	Slack
## lam1m - (3-lam2m-lam3m)	0.000
## lam1w - (3-lam2w-lam3w)	0.000
## lam4m - (3-lam5m-lam6m)	0.000
## lam4w - (3-lam5w-lam6w)	0.000
## nu1m - (0-nu2m-nu3m)	0.000
## nu1w - (0-nu2w-nu3w)	0.000
## nu4m - (0-nu5m-nu6m)	0.000
## nu4w - (0-nu5w-nu6w)	0.000

Configural model - summary

- ▶ Configural model has good fit (based on chi-square and fit statistics).
- ▶ Conclude this model is reasonable for both groups.
- ▶ Move onto weak invariance.

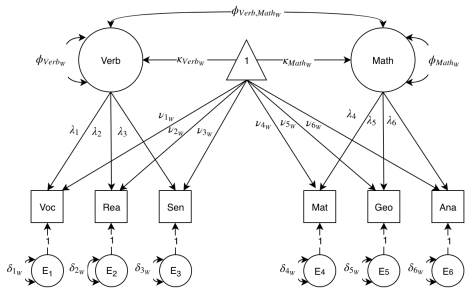
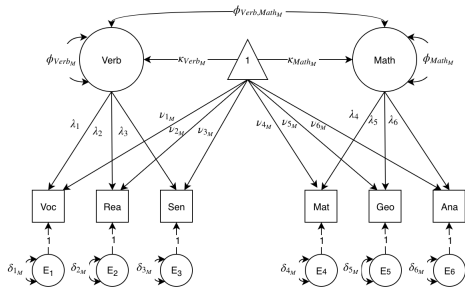
Weak Invariance

- ▶ Adds the constraint that unstandardized pattern coefficients are the same across groups.
- ▶ Weak invariance model is nested within configural model, can use chi-square test of difference.
 - ▶ Also, should consider change in RMSEA and CFI.
- ▶ If weak invariance not meaningful reduction in fit, constructs are manifested in the same way in each group
- ▶ Factor scores would be calculated using the same weighting scheme in all groups.

What if we fail to obtain weak invariance?

Gregorich (2006) provides two reasons

- ▶ The factors, or a subset of the items, have different meaning over the groups.
- ▶ **Extreme response style** - Low ERS, avoidance of extremes, high ERS, favoring extremes.
 - ▶ Maybe one group favors/discourages extremes or a group changes in how it views extremes over time.



Male Model

$$Voc = \nu_{1_M} + \lambda_1 Verb + \delta_{1_M}$$

$$Rea = \nu_{2_M} + \lambda_2 Verb + \delta_{2_M}$$

$$Sen = \nu_{3_M} + \lambda_3 Verb + \delta_{3_M}$$

$$Mat = \nu_{4_M} + \lambda_4 Math + \delta_{4_M}$$

$$Geo = \nu_{5_M} + \lambda_5 Math + \delta_{5_M}$$

$$Ana = \nu_{6_M} + \lambda_6 Math + \delta_{6_M}$$

$$\phi_{Verb_M}, \phi_{Math_M}, \phi_{Verb_M, Math_M}, \kappa_{Verb_M}, \kappa_{Math_M}$$

Female Model

$$Voc = \nu_{1_W} + \lambda_1 Verb + \delta_{1_W}$$

$$Rea = \nu_{2_W} + \lambda_2 Verb + \delta_{2_W}$$

$$Sen = \nu_{3_W} + \lambda_3 Verb + \delta_{3_W}$$

$$Mat = \nu_{4_W} + \lambda_4 Math + \delta_{4_W}$$

$$Geo = \nu_{5_W} + \lambda_5 Math + \delta_{5_W}$$

$$Ana = \nu_{6_W} + \lambda_6 Math + \delta_{6_W}$$

$$\phi_{Verb_W}, \phi_{Math_W}, \phi_{Verb_W, Math_W}, \kappa_{Verb_W}, \kappa_{Math_W}$$

Step 2: Weak invariance

```
weak.mod <- "  
# define factor/unstd. pattern coef  
verb ~ c(lam1, lam1)*vocab + c(lam2, lam2)*reading + c(lam3, lam3)*sentcomp  
math ~ c(lam4, lam4)*mathmtcs + c(lam5, lam5)*geometry + c(lam6, lam6)*analyrea  
  
# factor intercepts  
verb ~ 1  
math ~ 1  
  
# factor variances/covariance  
verb ~~ verb + math  
math ~~ math  
  
# intercepts  
vocab ~ c(nu1m, nu1w)*1  
reading ~ c(nu2m, nu2w)*1  
sentcomp ~ c(nu3m, nu3w)*1  
mathmtcs ~ c(nu4m, nu4w)*1  
geometry ~ c(nu5m, nu5w)*1  
analyrea ~ c(nu6m, nu6w)*1  
  
# residual variances  
vocab ~~ c(e1m, e1w)*vocab  
reading ~~ c(e2m, e2w)*reading  
sentcomp ~~ c(e3m, e3w)*sentcomp  
mathmtcs ~~ c(e4m, e4w)*mathmtcs  
geometry ~~ c(e5m, e5w)*geometry  
analyrea ~~ c(e6m, e6w)*analyrea  
  
# define constraints  
lam1 == 3 - lam2 - lam3  
lam4 == 3 - lam5 - lam6  
  
nu1m == 0 - nu2m - nu3m  
nu1w == 0 - nu2w - nu3w  
nu4m == 0 - nu5m - nu6m  
nu4w == 0 - nu5w - nu6w  
"
```

```

weak.fit <- lavaan(weak.mod, data = interest, group = "sex")
anova(weak.fit, config.fit)

## Chi Square Difference Test
##
##           Df      AIC      BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## config.fit 16 3081.8 3215.6 18.413
## weak.fit   20 3075.4 3195.2 20.048      1.6348      4      0.8025

fit.stats <- rbind(fitmeasures(config.fit, c("chisq", "df", "pvalue", "cfi", "rmsea", "srmr")),
  fitmeasures(weak.fit, c("chisq", "df", "pvalue", "cfi", "rmsea", "srmr")))
rownames(fit.stats) <- c("Configural", "Weak")
fit.stats

##           chisq df    pvalue      cfi      rmsea      srmr
## Configural 18.41320 16 0.3002620 0.9981046 0.034736151 0.01433416
## Weak       20.04805 20 0.4549276 0.9999623 0.004384013 0.02119742

```

```
summary(weak.fit)
```

```
## Group 1 [female]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb =~				
##	vocab (lam1)	1.075	0.027	39.261	0.000
##	reading (lam2)	0.961	0.031	31.128	0.000
##	sentcmp (lam3)	0.964	0.030	31.807	0.000
##	math =~				
##	mthmtcs (lam4)	1.069	0.028	37.574	0.000
##	geomtry (lam5)	0.924	0.033	27.818	0.000
##	analyre (lam6)	1.008	0.031	32.894	0.000

```
## Group 2 [male]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb =~				
##	vocab (lam1)	1.075	0.027	39.261	0.000
##	reading (lam2)	0.961	0.031	31.128	0.000
##	sentcmp (lam3)	0.964	0.030	31.807	0.000
##	math =~				
##	mthmtcs (lam4)	1.069	0.028	37.574	0.000
##	geomtry (lam5)	0.924	0.033	27.818	0.000
##	analyre (lam6)	1.008	0.031	32.894	0.000

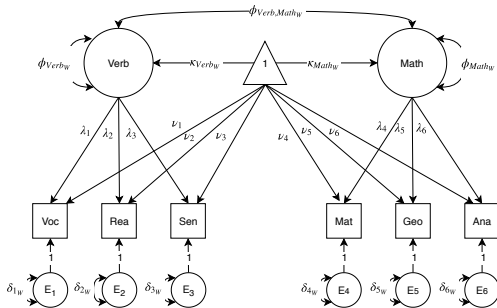
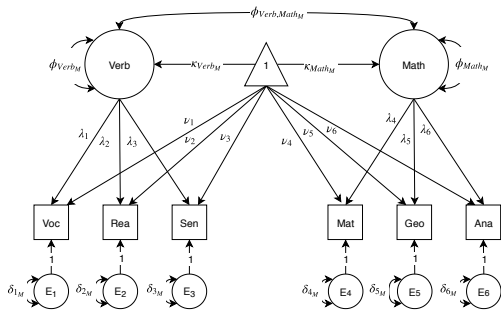
Strong Invariance

- ▶ Adds the constraint that unstandardized intercepts are the same across groups.
- ▶ Strong invariance model is nested within weak invariance model, can use chi-square test of difference.
- ▶ If strong invariance not meaningful reduction in fit, groups use the response scale of the indicator in a similar manner.
- ▶ Two people from two different groups with the same score on the factor *expected* to obtain the same observed score on the indicator.
- ▶ Ensures that ...
 1. Group differences in estimated factor means will be unbiased
 2. Group differences in indicator means or estimated factor scores will be related to the factor means.
- ▶ Minimum level required for group comparisons.

What if we fail to obtain strong invariance?

Systematic influences unrelated to the factor(s) decrease/increase the overall level of responding on an indicator in a given group (**differential additive response style**).

- ▶ Could be caused by cultural differences, cohorts effects, differences in how data are collected.
 - ▶ Measuring response time to a certain computer task. One sample uses a new computer and records response instantly; another computer is old, and response time adds a couple of deciseconds.
- ▶ **Differential item functioning (DIF)** - unequal pattern coefficients or intercepts over groups.
 - ▶ Score depends on membership in a group.
 - ▶ During instrument development, items flagged for DIF need to be altered or removed, so identifying indicators responsible for violating weak/strong invariance critical.



Male Model

$$Voc = \nu_1 + \lambda_1 Verb + \delta_{1_M}$$

$$Rea = \nu_2 + \lambda_2 Verb + \delta_{2_M}$$

$$Sen = \nu_3 + \lambda_3 Verb + \delta_{3_M}$$

$$Mat = \nu_4 + \lambda_4 Math + \delta_{4_M}$$

$$Geo = \nu_5 + \lambda_5 Math + \delta_{5_M}$$

$$Ana = \nu_6 + \lambda_6 Math + \delta_{6_M}$$

$$\phi_{Verb_M}, \phi_{Math_M}, \phi_{Verb_M, Math_M}, \kappa_{Verb_M}, \kappa_{Math_M}$$

Female Model

$$Voc = \nu_1 + \lambda_1 Verb + \delta_{1_W}$$

$$Rea = \nu_2 + \lambda_2 Verb + \delta_{2_W}$$

$$Sen = \nu_3 + \lambda_3 Verb + \delta_{3_W}$$

$$Mat = \nu_4 + \lambda_4 Math + \delta_{4_W}$$

$$Geo = \nu_5 + \lambda_5 Math + \delta_{5_W}$$

$$Ana = \nu_6 + \lambda_6 Math + \delta_{6_W}$$

$$\phi_{Verb_W}, \phi_{Math_W}, \phi_{Verb_W, Math_W}, \kappa_{Verb_W}, \kappa_{Math_W}$$

Step 3: Strong invariance

```
strong.mod <- "  
# define factor/unstd. pattern coef  
verb ~ c(lam1, lam1)*vocab + c(lam2, lam2)*reading + c(lam3, lam3)*sentcomp  
math ~ c(lam4, lam4)*mathmtcs + c(lam5, lam5)*geometry + c(lam6, lam6)*analyrea  
  
# factor intercepts  
verb ~ 1  
math ~ 1  
  
# factor variances/covariance  
verb ~~ verb + math  
math ~~ math  
  
# intercepts  
vocab ~ c(nu1, nu1)*1  
reading ~ c(nu2, nu2)*1  
sentcomp ~ c(nu3, nu3)*1  
mathmtcs ~ c(nu4, nu4)*1  
geometry ~ c(nu5, nu5)*1  
analyrea ~ c(nu6, nu6)*1  
  
# residual variances  
vocab ~~ c(e1m, e1w)*vocab  
reading ~~ c(e2m, e2w)*reading  
sentcomp ~~ c(e3m, e3w)*sentcomp  
mathmtcs ~~ c(e4m, e4w)*mathmtcs  
geometry ~~ c(e5m, e5w)*geometry  
analyrea ~~ c(e6m, e6w)*analyrea  
  
# define constraints  
lam1 == 3 - lam2 - lam3  
lam4 == 3 - lam5 - lam6  
  
nu1 == 0 - nu2 - nu3  
nu4 == 0 - nu5 - nu6  
"
```

```

strong.fit <- lavaan(strong.mod, data = interest, group = "sex")
anova(strong.fit, weak.fit)

## Chi Square Difference Test
##
##           Df      AIC      BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## weak.fit   20 3075.4 3195.2 20.048
## strong.fit 24 3070.6 3176.3 23.230      3.1816      4      0.5279

fit.stats <- rbind(fit.stats,
  fitmeasures(strong.fit, c("chisq", "df", "pvalue", "cfi", "rmsea", "srmr")))
rownames(fit.stats)[3] <- c("Strong")
fit.stats

##           chisq df    pvalue      cfi      rmsea      srmr
## Configural 18.41320 16 0.3002620 0.9981046 0.034736151 0.01433416
## Weak       20.04805 20 0.4549276 0.9999623 0.004384013 0.02119742
## Strong     23.22963 24 0.5062782 1.0000000 0.000000000 0.02371500

```

```
summary(strong.fit)
```

```
## Group 1 [female]:
```

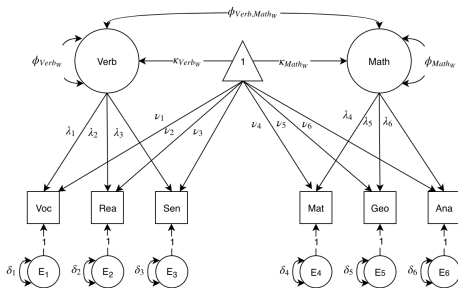
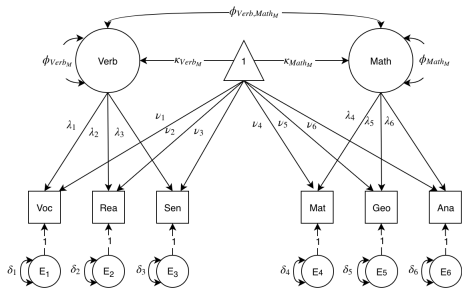
	Estimate	Std.Err	z-value	P(> z)
## verb = ~				
## vocab (lam1)	1.073	0.027	39.362	0.000
## reading (lam2)	0.962	0.031	31.298	0.000
## sentcmp (lam3)	0.965	0.030	31.957	0.000
## math = ~				
## mthmtcs (lam4)	1.075	0.028	38.100	0.000
## geomtry (lam5)	0.922	0.033	28.009	0.000
## analyre (lam6)	1.002	0.030	32.944	0.000
##				
##	Estimate	Std.Err	z-value	P(> z)
## verb	0.174	0.086	2.029	0.042
## math	-0.009	0.089	-0.105	0.916
## .vocab (nu1)	-0.016	0.021	-0.768	0.443
## .reading (nu2)	0.038	0.026	1.500	0.133
## .sentcomp (nu3)	-0.022	0.025	-0.870	0.384
## .mathmtcs (nu4)	-0.035	0.023	-1.491	0.136
## .geometry (nu5)	-0.011	0.029	-0.365	0.715
## .analyrea (nu6)	0.045	0.026	1.718	0.086
##				

```
## Group 2 [male]:
```

	Estimate	Std.Err	z-value	P(> z)
## verb = ~				
## vocab (lam1)	1.073	0.027	39.362	0.000
## reading (lam2)	0.962	0.031	31.298	0.000
## sentcmp (lam3)	0.965	0.030	31.957	0.000
## math = ~				
## mthmtcs (lam4)	1.075	0.028	38.100	0.000
## geomtry (lam5)	0.922	0.033	28.009	0.000
## analyre (lam6)	1.002	0.030	32.944	0.000
##				
##	Estimate	Std.Err	z-value	P(> z)
## verb	0.031	0.077	0.409	0.683
## math	0.268	0.080	3.340	0.001
## .vocab (nu1)	-0.016	0.021	-0.768	0.443
## .reading (nu2)	0.038	0.026	1.500	0.133
## .sentcomp (nu3)	-0.022	0.025	-0.870	0.384
## .mathmtcs (nu4)	-0.035	0.023	-1.491	0.136
## .geometry (nu5)	-0.011	0.029	-0.365	0.715
## .analyrea (nu6)	0.045	0.026	1.718	0.086

Strict invariance

- ▶ Adds the constraint that error variances/covariances are the same across groups.
- ▶ Strict invariance model is nested within strong invariance model, can use chi-square test of difference.
- ▶ If strict invariance not meaningful reduction in fit, indicators measure the same factor in each group with the same degree of error (inversely, precision).
- ▶ Disagreement as to whether it is necessary to obtain strict invariance to say factors are measured identically across groups. Why?
 - ▶ Recall, the residual (or unique variance) contains both random measurement error and systematic error.
 - ▶ Strict invariance means the sum of these components are the same across group.
 - ▶ Reasonable to think systematic error could be the same across groups.
 - ▶ But would we expect random measurement error to be the same?
- ▶ Similar disagreement as to whether it's necessary to have strict invariance to compare observed differences in variances and covariance.



Male Model

$$Voc = \nu_1 + \lambda_1 Verb + \delta_1$$

$$Rea = \nu_2 + \lambda_2 Verb + \delta_2$$

$$Sen = \nu_3 + \lambda_3 Verb + \delta_3$$

$$Mat = \nu_4 + \lambda_4 Math + \delta_4$$

$$Geo = \nu_5 + \lambda_5 Math + \delta_5$$

$$Ana = \nu_6 + \lambda_6 Math + \delta_6$$

$$\phi_{Verb_M}, \phi_{Math_M}, \phi_{Verb_M, Math_M}, \kappa_{Verb_M}, \kappa_{Math_M}$$

Female Model

$$Voc = \nu_1 + \lambda_1 Verb + \delta_1$$

$$Rea = \nu_2 + \lambda_2 Verb + \delta_2$$

$$Sen = \nu_3 + \lambda_3 Verb + \delta_3$$

$$Mat = \nu_4 + \lambda_4 Math + \delta_4$$

$$Geo = \nu_5 + \lambda_5 Math + \delta_5$$

$$Ana = \nu_6 + \lambda_6 Math + \delta_6$$

$$\phi_{Verb_W}, \phi_{Math_W}, \phi_{Verb_W, Math_W}, \kappa_{Verb_W}, \kappa_{Math_W}$$

Step 4: Strict invariance

```
strict.mod <- "  
# define factor/unstd. pattern coef  
verb  ~ c(lam1, lam1)*vocab + c(lam2, lam2)*reading + c(lam3, lam3)*sentcomp  
math  ~ c(lam4, lam4)*mathmtcs + c(lam5, lam5)*geometry + c(lam6, lam6)*analyrea  
  
# factor intercepts  
verb  ~ 1  
math  ~ 1  
  
# factor variances/covariance  
verb  ~~ verb + math  
math  ~~ math  
  
# intercepts  
vocab ~ c(nu1, nu1)*1  
reading ~ c(nu2, nu2)*1  
sentcomp ~ c(nu3, nu3)*1  
mathmtcs ~ c(nu4, nu4)*1  
geometry ~ c(nu5, nu5)*1  
analyrea ~ c(nu6, nu6)*1  
  
# residual variances  
vocab ~~ c(e1, e1)*vocab  
reading ~~ c(e2, e2)*reading  
sentcomp ~~ c(e3, e3)*sentcomp  
mathmtcs ~~ c(e4, e4)*mathmtcs  
geometry ~~ c(e5, e5)*geometry  
analyrea ~~ c(e6, e6)*analyrea  
  
# define constraints  
lam1 == 3 - lam2 - lam3  
lam4 == 3 - lam5 - lam6  
  
nu1 == 0 - nu2 - nu3  
nu4 == 0 - nu5 - nu6  
"
```

```

strict.fit <- lavaan(strict.mod, data = interest, group = "sex")
anova(strict.fit, strong.fit)

## Chi Square Difference Test
##
##           Df      AIC      BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## strong.fit 24 3070.6 3176.3 23.230
## strict.fit 30 3062.7 3147.2 27.264      4.0347      6      0.672

fit.stats <- rbind(fit.stats,
                   fitmeasures(strict.fit, c("chisq", "df", "pvalue", "cfi", "rmsea", "srmr")))
rownames(fit.stats)[4] <- c("Strict")
fit.stats

##           chisq df    pvalue      cfi      rmsea      srmr
## Configural 18.41320 16 0.3002620 0.9981046 0.034736151 0.01433416
## Weak       20.04805 20 0.4549276 0.9999623 0.004384013 0.02119742
## Strong     23.22963 24 0.5062782 1.0000000 0.000000000 0.02371500
## Strict     27.26431 30 0.6093608 1.0000000 0.000000000 0.03077196

```

```
summary(strict.fit)
```

```
## lavaan 0.6-3 ended normally after 33 iterations
```

```
##
```

```
## Optimization method NLMINB
```

```
## Number of free parameters 46
```

```
## Number of equality constraints 22
```

```
##
```

```
## Number of observations per group
```

```
## female 122
```

```
## male 128
```

```
##
```

```
## Estimator ML
```

```
## Model Fit Test Statistic 27.264
```

```
## Degrees of freedom 30
```

```
## P-value (Chi-square) 0.609
```

```
##
```

```
## Chi-square for each group:
```

```
##
```

```
## female 15.063
```

```
## male 12.201
```

```
##
```

```
## Parameter Estimates:
```

```
##
```

```
## Information Expected
```

```
## Information saturated (h1) model Structured
```

```
## Standard Errors Standard
```

```
summary(strict.fit)
```

```
## Group 1 [female]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb = ~				
##	vocab (lam1)	1.071	0.027	39.528	0.000
##	reading (lam2)	0.962	0.031	31.395	0.000
##	sentcmp (lam3)	0.968	0.030	31.808	0.000
##	math = ~				
##	mthmtcs (lam4)	1.076	0.028	38.237	0.000
##	geomtry (lam5)	0.922	0.033	28.018	0.000
##	analyre (lam6)	1.003	0.031	32.805	0.000

```
##
```

```
## Covariances:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb ~~				
##	math	0.780	0.111	7.038	0.000

```
##
```

```
## Intercepts:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb	0.173	0.086	2.004	0.045
##	math	-0.011	0.090	-0.125	0.900
##	.vocab (nu1)	-0.016	0.021	-0.769	0.442
##	.reading (nu2)	0.039	0.026	1.537	0.124
##	.sentcomp (nu3)	-0.023	0.025	-0.901	0.368
##	.mathmtcs (nu4)	-0.035	0.023	-1.519	0.129
##	.geometry (nu5)	-0.008	0.029	-0.284	0.776
##	.analyrea (nu6)	0.044	0.026	1.650	0.099

```
##
```

```
## Variances:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb	0.846	0.117	7.255	0.000
##	math	0.904	0.125	7.209	0.000
##	.vocab (e1)	0.110	0.022	5.024	0.000
##	.reading (e2)	0.265	0.030	8.984	0.000
##	.sentcomp (e3)	0.258	0.029	8.887	0.000
##	.mathmtcs (e4)	0.131	0.025	5.161	0.000
##	.geometry (e5)	0.348	0.037	9.479	0.000
##	.analyrea (e6)	0.265	0.032	8.383	0.000

```
summary(strict.fit)
```

```
## Group 2 [male]:
```

```
##
```

```
## Latent Variables:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb = "				
##	vocab (lam1)	1.071	0.027	39.528	0.000
##	reading (lam2)	0.962	0.031	31.395	0.000
##	sentcmp (lam3)	0.968	0.030	31.808	0.000
##	math = "				
##	mathmtcs (lam4)	1.076	0.028	38.237	0.000
##	geomtry (lam5)	0.922	0.033	28.018	0.000
##	analyre (lam6)	1.003	0.031	32.805	0.000

```
##
```

```
## Covariances:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb ~~				
##	math	0.545	0.084	6.462	0.000

```
##
```

```
## Intercepts:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb	0.030	0.077	0.386	0.700
##	math	0.267	0.080	3.324	0.001
##	.vocab (nu1)	-0.016	0.021	-0.769	0.442
##	.reading (nu2)	0.039	0.026	1.537	0.124
##	.sentcomp (nu3)	-0.023	0.025	-0.901	0.368
##	.mathmtcs (nu4)	-0.035	0.023	-1.519	0.129
##	.geometry (nu5)	-0.008	0.029	-0.284	0.776
##	.analyrea (nu6)	0.044	0.026	1.650	0.099

```
##
```

```
## Variances:
```

		Estimate	Std.Err	z-value	P(> z)
##	verb	0.688	0.094	7.322	0.000
##	math	0.748	0.103	7.277	0.000
##	.vocab (e1)	0.110	0.022	5.024	0.000
##	.reading (e2)	0.265	0.030	8.984	0.000
##	.sentcomp (e3)	0.258	0.029	8.887	0.000
##	.mathmtcs (e4)	0.131	0.025	5.161	0.000
##	.geometry (e5)	0.348	0.037	9.479	0.000
##	.analyrea (e6)	0.265	0.032	8.383	0.000

$$\hat{V}oc = -0.016 + 1.071Verb$$

$$\hat{R}ea = 0.039 + 0.962Verb$$

$$\hat{S}en = -0.023 + 0.968Verb$$

$$\hat{\sigma}_{Voc}^2 = 0.110, \hat{\sigma}_{Rea}^2 = 0.265, \hat{\sigma}_{Sen}^2 = 0.258$$

$$\hat{M}at = -0.035 + 1.076Math$$

$$\hat{G}eo = -0.008 + 0.922Math$$

$$\hat{A}na = 0.044 + 1.003Math$$

$$\hat{\sigma}_{Mat}^2 = 0.131, \hat{\sigma}_{Geo}^2 = 0.348, \hat{\sigma}_{Ana}^2 = 0.265$$

$$\hat{\phi}_{V_W} = 0.846, \hat{\phi}_{M_W} = 0.904, \hat{\phi}_{V_W, M_W} = 0.780, \hat{\kappa}_{V_W} = 0.173, \hat{\kappa}_{M_W} = -0.011$$

$$\hat{\phi}_{V_M} = 0.688, \hat{\phi}_{M_M} = 0.748, \hat{\phi}_{V_M, M_M} = 0.545, \hat{\kappa}_{V_M} = 0.030, \hat{\kappa}_{M_M} = 0.267$$

Now what?

- ▶ We concluded we have strict invariance.
 - ▶ Extremely rare!
- ▶ Compare the groups on IQ and calculate standardized mean differences (i.e., Cohen's d)

Comparing groups

```
params <- parameterestimates(strict.fit)
subset(params, (lhs %in% c("verb", "math") & op == "~1") | (lhs %in% c("verb", "math") & rhs %in% c("verb", "math")))
```

##	lhs	op	rhs	block	group	label	est	se	z	pvalue	ci.lower	ci.upper
## 7	verb	~1		1	1		0.173	0.086	2.004	0.045	0.004	0.342
## 8	math	~1		1	1		-0.011	0.090	-0.125	0.900	-0.187	0.164
## 9	verb	~~	verb	1	1		0.846	0.117	7.255	0.000	0.617	1.074
## 10	verb	~~	math	1	1		0.780	0.111	7.038	0.000	0.563	0.997
## 11	math	~~	math	1	1		0.904	0.125	7.209	0.000	0.658	1.150
## 30	verb	~1		2	2		0.030	0.077	0.386	0.700	-0.121	0.180
## 31	math	~1		2	2		0.267	0.080	3.324	0.001	0.109	0.424
## 32	verb	~~	verb	2	2		0.688	0.094	7.322	0.000	0.504	0.872
## 33	verb	~~	math	2	2		0.545	0.084	6.462	0.000	0.380	0.710
## 34	math	~~	math	2	2		0.748	0.103	7.277	0.000	0.547	0.949

- Do these differ? Perform a t-test.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}, \quad df \approx \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2(N_1-1)} + \frac{s_2^4}{N_2^2(N_2-1)}}$$

use `pt(t = t, df = df, lower.tail = FALSE) * 2` in R

- Calculate effect size.

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2}}$$