SEM & Lavaan

Zoë Hawks

10/27/2017

1. Fit a measurement model to your constructs at one time point. Try out the different types of scaling discussed in class. What changes what stays the same?

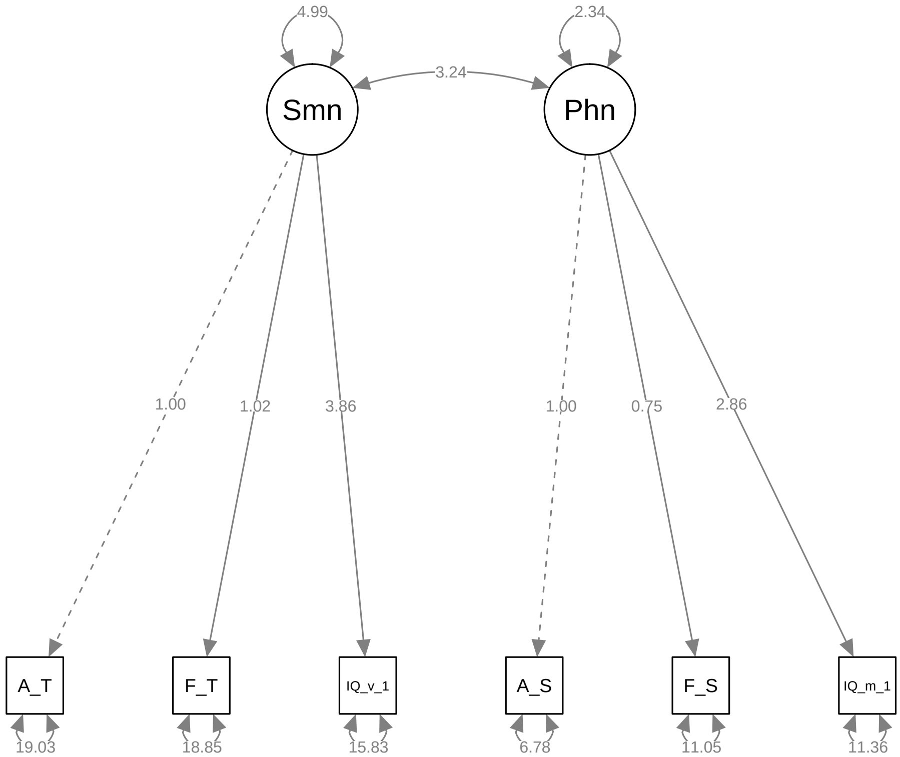
Measurement model was fit at timepoint 1. Under the marker method, parameter estimates were fixed at 1 for the first indicators (i.e., Animal_TW_Clusters_1 and Animal_Switches_1). Under the fixed factor method, in contrast, parameter estiamtes were fixed to 1 for the latent variables. In both cases, fit indices (e.g., logLikelihood, TLI, CFI, RMSEA) remain constant.

```
T1.mod <- '
            Semantic =~ Animal_TW_Clusters_1 + Food_TW_Clusters_1 + IQ_vocraw_1
            Phonemic =~ Animal_Switches_1 + Food_Switches_1 + IQ_mrraw_1
'

#Marker method
fit.marker <- cfa(T1.mod, data=mydata_wide)
summary(fit.marker, fit.measures = TRUE)
```

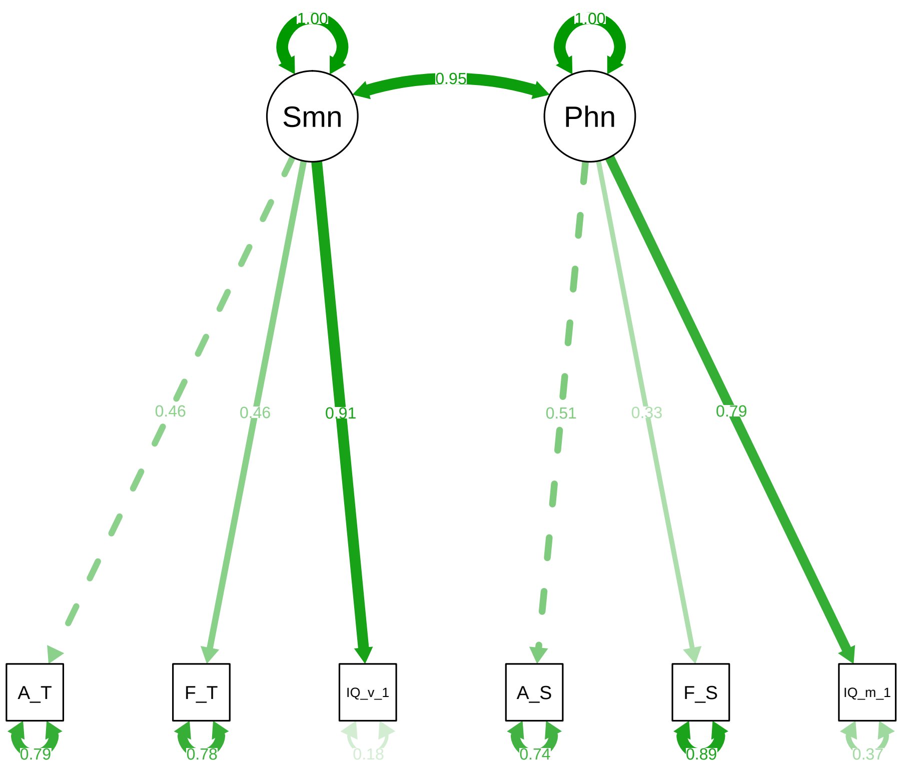
```
## lavaan (0.5-23.1097) converged normally after  53 iterations
##
##   Number of observations                            67
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               22.669
##   Degrees of freedom                                 8
##   P-value (Chi-square)                           0.004
##
## Model test baseline model:
##
##   Minimum Function Test Statistic              114.013
##   Degrees of freedom                                15
##   P-value                                        0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                    0.852
##   Tucker-Lewis Index (TLI)                       0.722
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -1161.512
##   Loglikelihood unrestricted model (H1)      -1150.178
##
##   Number of free parameters                         13
##   Akaike (AIC)                                2349.024
##   Bayesian (BIC)                              2377.685
##   Sample-size adjusted Bayesian (BIC)         2336.753
##
```

```
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.165
##   90 Percent Confidence Interval          0.087  0.248
##   P-value RMSEA <= 0.05                          0.012
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.091
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic =~
##     Anml_TW_Clst_1    1.000
##     Fd_TW_Clstrs_1    1.019    0.366    2.786    0.005
##     IQ_vocraw_1       3.862    1.109    3.483    0.000
##   Phonemic =~
##     Animl_Swtchs_1    1.000
##     Food_Switchs_1    0.755    0.338    2.231    0.026
##     IQ_mrraw_1        2.864    0.743    3.854    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic ~~
##     Phonemic          3.244    1.311    2.474    0.013
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Anml_TW_Clst_1   19.027    3.439    5.533    0.000
##    .Fd_TW_Clstrs_1   18.855    3.416    5.519    0.000
##    .IQ_vocraw_1      15.830   11.974    1.322    0.186
##    .Animl_Swtchs_1    6.776    1.261    5.373    0.000
##    .Food_Switchs_1   11.048    1.955    5.651    0.000
##    .IQ_mrraw_1       11.360    3.959    2.869    0.004
##     Semantic          4.991    2.726    1.831    0.067
##     Phonemic          2.343    1.154    2.030    0.042
```

```r
semPaths(fit.marker, layout = "tree", whatLabels = "est")
```




```r
semPaths(fit.marker, layout = "tree", what = "std")
```

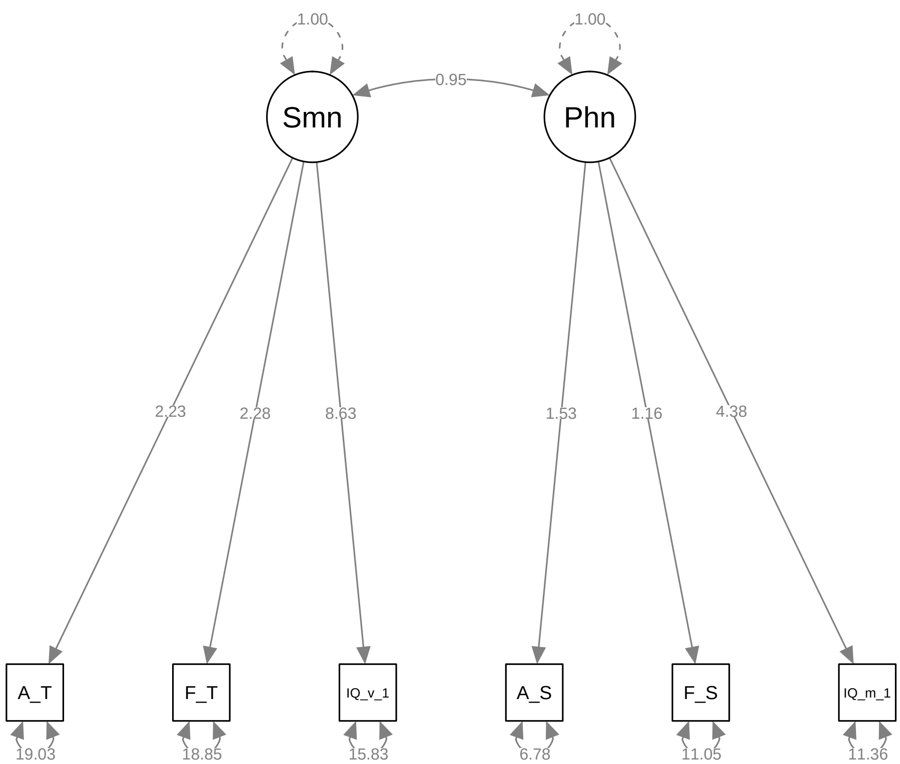



```r
#Fixed factor method
fit.fixed <- cfa(T1.mod, data=mydata_wide, std.lv = T)
summary(fit.fixed, fit.measures = TRUE)
```

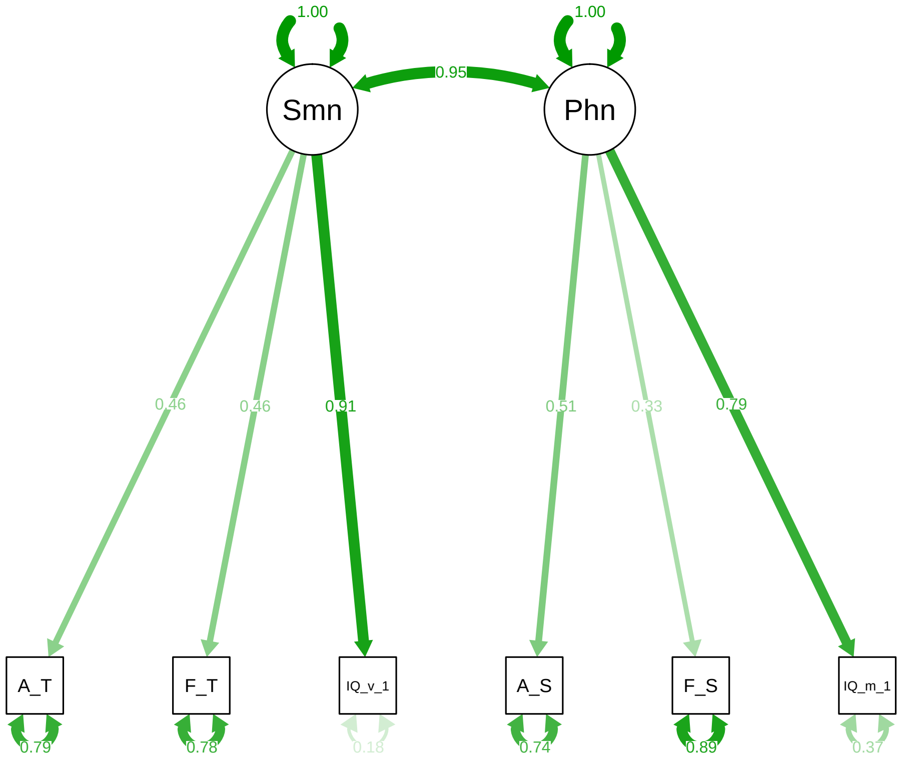
```
## lavaan (0.5-23.1097) converged normally after  48 iterations
##
##   Number of observations                            67
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               22.669
##   Degrees of freedom                                 8
##   P-value (Chi-square)                           0.004
##
## Model test baseline model:
##
##   Minimum Function Test Statistic              114.013
##   Degrees of freedom                                15
##   P-value                                        0.000
##
## User model versus baseline model:
##
##   Comparative Fit Index (CFI)                    0.852
##   Tucker-Lewis Index (TLI)                       0.722
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -1161.512
##   Loglikelihood unrestricted model (H1)      -1150.178
##
##   Number of free parameters                         13
##   Akaike (AIC)                                2349.024
##   Bayesian (BIC)                              2377.685
##   Sample-size adjusted Bayesian (BIC)         2336.753
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.165
##   90 Percent Confidence Interval          0.087  0.248
##   P-value RMSEA <= 0.05                          0.012
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.091
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic =~
##     Anml_TW_Clst_1    2.234    0.610    3.661    0.000
##     Fd_TW_Clstrs_1    2.276    0.609    3.736    0.000
##     IQ_vocraw_1       8.629    1.117    7.723    0.000
##   Phonemic =~
##     Animl_Swtchs_1    1.531    0.377    4.060    0.000
##     Food_Switchs_1    1.155    0.456    2.531    0.011
```

```
##     IQ_mrraw_1        4.383    0.683    6.416    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic ~~
##     Phonemic          0.949    0.108    8.814    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Anml_TW_Clst_1   19.027    3.439    5.533    0.000
##    .Fd_TW_Clstrs_1   18.855    3.416    5.519    0.000
##    .IQ_vocraw_1      15.830   11.974    1.322    0.186
##    .Animl_Swtchs_1    6.776    1.261    5.373    0.000
##    .Food_Switchs_1   11.048    1.955    5.651    0.000
##    .IQ_mrraw_1       11.360    3.959    2.869    0.004
##     Semantic          1.000
##     Phonemic          1.000
```

```
semPaths(fit.fixed, layout = "tree", whatLabels = "est")
```




```
semPaths(fit.fixed, layout = "tree", what = "std")
```



2. What do the fit statistics say about your latent variable? Good/bad? Is your latent variable Just identified/saturdated, under identified or over identified?

- RMSEA = .165, SRMR = .091, TLI = .722, CFI = .852
- RMSEA & SRMR > .08 and TLI & CFI < .90, suggesting poor fit – that is, the latent variables are not effectively capturing commonalities among their indicator variables. This could be due to (1) high measurment error or (2) highly disparate indicators
- This model is over identified, as evidenced by the positive degrees of freedom (15).

3. Fit a longitudinal CFA model where you a) first correlate your latent factors across time and then b) a second model that predicts later times by a prevous time (ie auto regressive; t1 -> t2 -> t3). What are your conclusions? How does one differ from the other?

For the longitudinal CFA model with correlated latent fators (Long.mod), I conclude that my latent factors are strongly correlated across time. Moreover, across all three timepoints, IQ verbal loads much more strongly onto the latent factors than Animal and Food total word counts, suggesting that IQ verbal may parsimonously account for shared variance among indicators.

For the autoregressive CFA model (auto.mod), a similar picture emerges, suggesting that Semantic_1 is highly predictive of Semantic_2, which is highly predictive of Semantic_3. Of note, standardized variances for S_2 and S_3 latent factors are reduced for the autoregressive model compared to the correlated model (see path diagrams), likely due to the fact that S_2 and S_3 variances are accounted for by earlier timepoints.

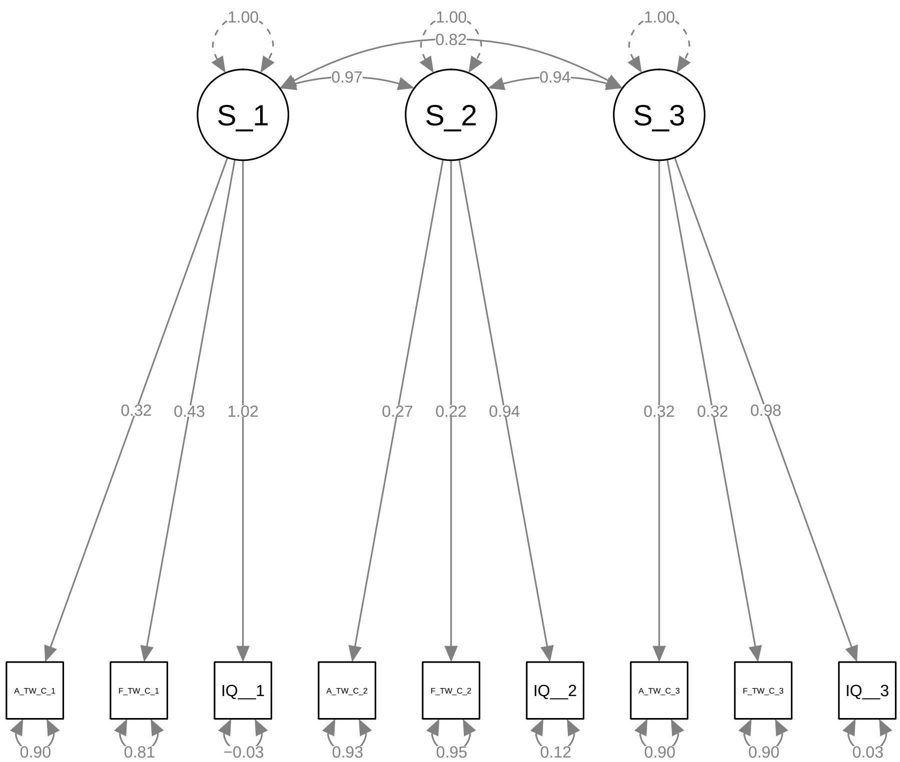
```
Long.mod <- '
            Semantic_1 =~ Animal_TW_Clusters_1 + Food_TW_Clusters_1 + IQ_vocraw_1
```

```
            Semantic_2 =~ Animal_TW_Clusters_2 + Food_TW_Clusters_2 + IQ_vocraw_2
            Semantic_3 =~ Animal_TW_Clusters_3 + Food_TW_Clusters_3 + IQ_vocraw_3
'

fit.long <- cfa(Long.mod, data=mydata_wide, std.lv = T)
summary(fit.long)
```

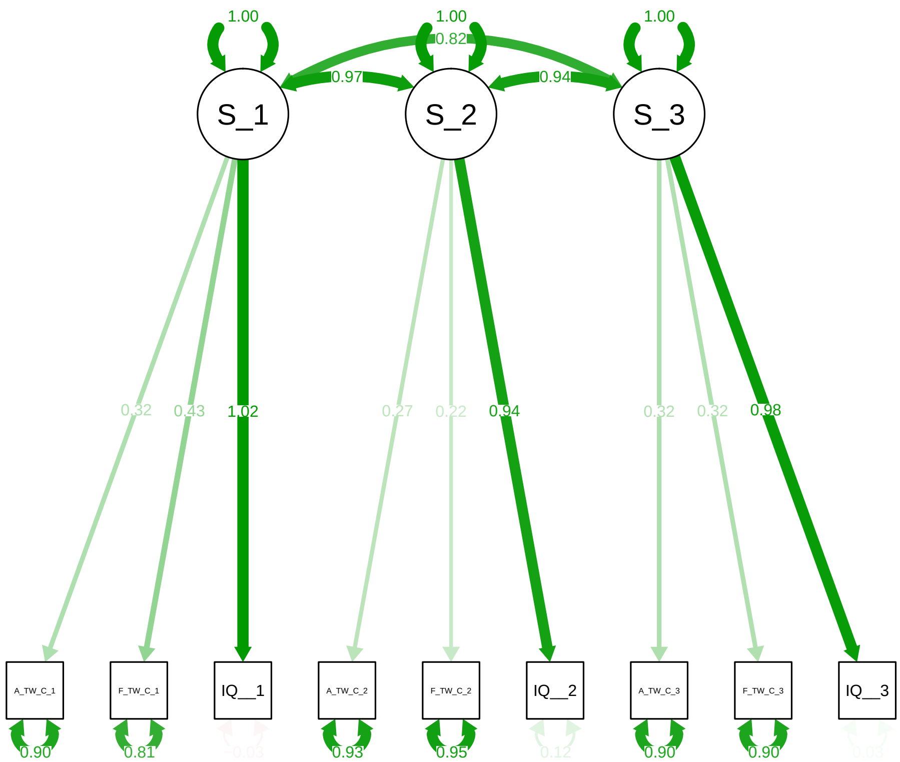
```
## lavaan (0.5-23.1097) converged normally after 103 iterations
##
##                                                   Used       Total
##   Number of observations                            64          67
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               91.983
##   Degrees of freedom                                24
##   P-value (Chi-square)                           0.000
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic_1 =~
##     Anml_TW_Clst_1    1.508    0.565    2.669    0.008
##     Fd_TW_Clstrs_1    2.148    0.593    3.622    0.000
##     IQ_vocraw_1       9.609    0.917   10.479    0.000
##   Semantic_2 =~
##     Anml_TW_Clst_2    1.209    0.553    2.187    0.029
##     Fd_TW_Clstrs_2    1.107    0.610    1.814    0.070
##     IQ_vocraw_2       8.025    0.931    8.617    0.000
##   Semantic_3 =~
##     Anml_TW_Clst_3    1.422    0.560    2.539    0.011
##     Fd_TW_Clstrs_3    1.774    0.702    2.528    0.011
##     IQ_vocraw_3       8.726    1.024    8.523    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic_1 ~~
##     Semantic_2        0.971    0.072   13.413    0.000
##     Semantic_3        0.822    0.080   10.220    0.000
##   Semantic_2 ~~
##     Semantic_3        0.942    0.092   10.283    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Anml_TW_Clst_1   19.557    3.452    5.665    0.000
##    .Fd_TW_Clstrs_1   19.929    3.534    5.639    0.000
##    .IQ_vocraw_1      -3.049    7.873   -0.387    0.699
##    .Anml_TW_Clst_2   18.646    3.299    5.652    0.000
##    .Fd_TW_Clstrs_2   23.199    4.101    5.656    0.000
##    .IQ_vocraw_2       8.726    7.815    1.117    0.264
##    .Anml_TW_Clst_3   18.073    3.214    5.623    0.000
```

```
##    .Fd_TW_Clstrs_3   28.428    5.055    5.623    0.000
##    .IQ_vocraw_3       2.694   11.201    0.241    0.810
##     Semantic_1        1.000
##     Semantic_2        1.000
##     Semantic_3        1.000
```

```r
semPaths(fit.long, whatLabels = "std")
```




```r
semPaths(fit.long, what = "std")
```

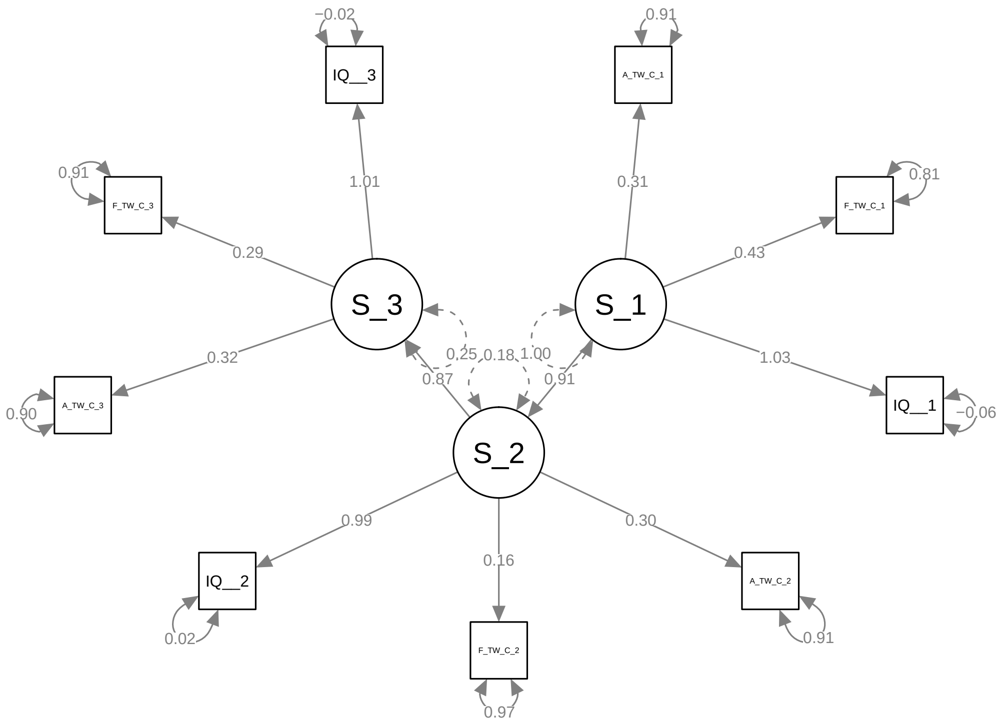



```r
Auto.mod <- '
            Semantic_1 =~ Animal_TW_Clusters_1 + Food_TW_Clusters_1 + IQ_vocraw_1
            Semantic_2 =~ Animal_TW_Clusters_2 + Food_TW_Clusters_2 + IQ_vocraw_2
            Semantic_3 =~ Animal_TW_Clusters_3 + Food_TW_Clusters_3 + IQ_vocraw_3

            Semantic_3 ~ Semantic_2
            Semantic_2 ~ Semantic_1
'

fit.auto <- cfa(Auto.mod, data=mydata_wide, std.lv = T)
summary(fit.auto)
```

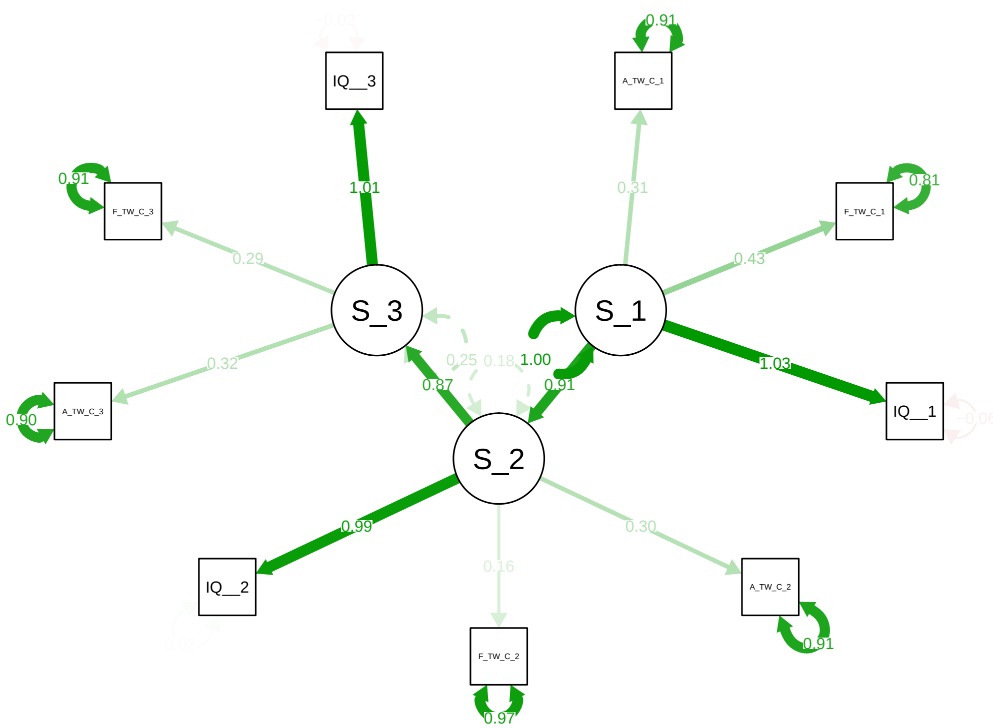
```
## lavaan (0.5-23.1097) converged normally after  81 iterations
##
##                                                   Used       Total
##   Number of observations                            64          67
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               92.425
##   Degrees of freedom                                25
##   P-value (Chi-square)                           0.000
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                      Estimate  Std.Err  z-value  P(>|z|)
##   Semantic_1 =~
```

```
##     Anml_TW_Clst_1    1.428    0.559    2.552    0.011
##     Fd_TW_Clstrs_1    2.147    0.591    3.634    0.000
##     IQ_vocraw_1       9.733    0.916   10.628    0.000
##   Semantic_2 =~
##     Anml_TW_Clst_2    0.564    0.261    2.162    0.031
##     Fd_TW_Clstrs_2    0.328    0.269    1.222    0.222
##     IQ_vocraw_2       3.548    0.849    4.181    0.000
##   Semantic_3 =~
##     Anml_TW_Clst_3    0.700    0.289    2.424    0.015
##     Fd_TW_Clstrs_3    0.815    0.360    2.264    0.024
##     IQ_vocraw_3       4.451    1.479    3.010    0.003
##
## Regressions:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   Semantic_3 ~
##     Semantic_2        0.734    0.300    2.446    0.014
##   Semantic_2 ~
##     Semantic_1        2.169    0.645    3.360    0.001
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Anml_TW_Clst_1   19.793    3.485    5.680    0.000
##    .Fd_TW_Clstrs_1   19.933    3.528    5.650    0.000
##    .IQ_vocraw_1      -5.452    8.398   -0.649    0.516
##    .Anml_TW_Clst_2   18.292    3.239    5.648    0.000
##    .Fd_TW_Clstrs_2   23.811    4.211    5.655    0.000
##    .IQ_vocraw_2       1.332    1.977    0.674    0.500
##    .Anml_TW_Clst_3   18.098    3.211    5.637    0.000
##    .Fd_TW_Clstrs_3   28.876    5.116    5.644    0.000
##    .IQ_vocraw_3      -1.773   12.578   -0.141    0.888
##     Semantic_1        1.000
##    .Semantic_2        1.000
##    .Semantic_3        1.000
```

```r
semPaths(fit.auto, layout = "circle2", whatLabels = "std")
```




```r
semPaths(fit.auto, layout = "circle2", what = "std")
```




```r
anova(fit.long, fit.auto) #simpler model (fit.long) is preferred
```

```
## Chi Square Difference Test
##
##          Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## fit.long 24         91.983
```

```
## fit.auto 25         92.425    0.44184       1     0.5062
```

4. Fit a longitdinal growth model in SEM and in HLM. Compare and contrast the differences.

Estimates of *intercept* are similar (1) between fixed slope SEM & HLM models, (2) between random slope SEM & HLM models, and (3) between fixed-slope + covariate SEM & HLM models. Estimates of *slope* are similar (1) between fixed slope SEM & HLM models and (2) between random slope SEM & HLM models. Esimates of slope differ between SEM & HLM models when a covariate is added because the HLM slope can no longer be interpreted as the straight-forward rate of increase (y over x) between timepoints. As well, logLikelihood tests designated the fixed slope + covariate model as the preferred model regardless of whether SEM or HLM was used.

```r
#HLM model
library(lme4)
library(car)
mod.HLM <- lmer(Sem_TotalCorrect ~ Timepoint + (1 | ID2), data = mydata) #fixed slope
summary(mod.HLM) #intercept = 31.06, slope = 3.85
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Sem_TotalCorrect ~ Timepoint + (1 | ID2)
##    Data: mydata
##
## REML criterion at convergence: 1405
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.20888 -0.57356 -0.05482  0.51023  2.96871
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  ID2      (Intercept) 57.40    7.576
##  Residual             38.03    6.167
## Number of obs: 200, groups:  ID2, 67
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  31.0596     1.4786  21.005
## Timepoint     3.8471     0.5355   7.184
##
## Correlation of Fixed Effects:
##           (Intr)
## Timepoint -0.722
```

```r
car::Anova(mod.HLM)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Sem_TotalCorrect
##            Chisq Df Pr(>Chisq)
## Timepoint 51.614  1  6.757e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
mod.HLM2 <- lmer(Sem_TotalCorrect ~ Timepoint + (Timepoint | ID2), data = mydata) #random slope
summary(mod.HLM2) #intercept = 31.07, slope = 3.84
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Sem_TotalCorrect ~ Timepoint + (Timepoint | ID2)
##    Data: mydata
##
## REML criterion at convergence: 1404
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.92900 -0.55230 -0.00624  0.49794  3.12399
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  ID2      (Intercept) 82.343   9.074
##           Timepoint    4.664   2.160    -0.54
##  Residual             33.485   5.787
## Number of obs: 200, groups:  ID2, 67
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  31.0709     1.5494  20.054
## Timepoint     3.8386     0.5679   6.759
##
## Correlation of Fixed Effects:
##           (Intr)
## Timepoint -0.752
```

```r
car::Anova(mod.HLM2)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Sem_TotalCorrect
##            Chisq Df Pr(>Chisq)
## Timepoint 45.685  1  1.389e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
mod.HLM3 <- lmer(Sem_TotalCorrect ~ Timepoint + Age_at_time_of_testing + (1 | ID2),
                 data = mydata) #fixed slope, with covariate
summary(mod.HLM3) #intercept = 14.83, est. timepoint = 1.83, est. age = 1.55
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: Sem_TotalCorrect ~ Timepoint + Age_at_time_of_testing + (1 |
##     ID2)
##    Data: mydata
##
## REML criterion at convergence: 1379
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.30716 -0.60338 -0.01007  0.46251  3.13253
##
## Random effects:
```

```
##  Groups   Name        Variance Std.Dev.
##  ID2      (Intercept) 35.42    5.952
##  Residual             37.79    6.147
## Number of obs: 200, groups:  ID2, 67
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)            14.8273     3.1802   4.662
## Timepoint               1.8291     0.6402   2.857
## Age_at_time_of_testing  1.5487     0.2738   5.655
##
## Correlation of Fixed Effects:
##             (Intr) Timpnt
## Timepoint    0.220
## Ag_t_tm_f_t -0.904 -0.552
```

```r
car::Anova(mod.HLM3)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Sem_TotalCorrect
##                          Chisq Df Pr(>Chisq)
## Timepoint               8.1641  1   0.004273 **
## Age_at_time_of_testing 31.9833  1  1.555e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
anova(mod.HLM, mod.HLM2, mod.HLM3) #fixed slope, covariate model is preferred
```

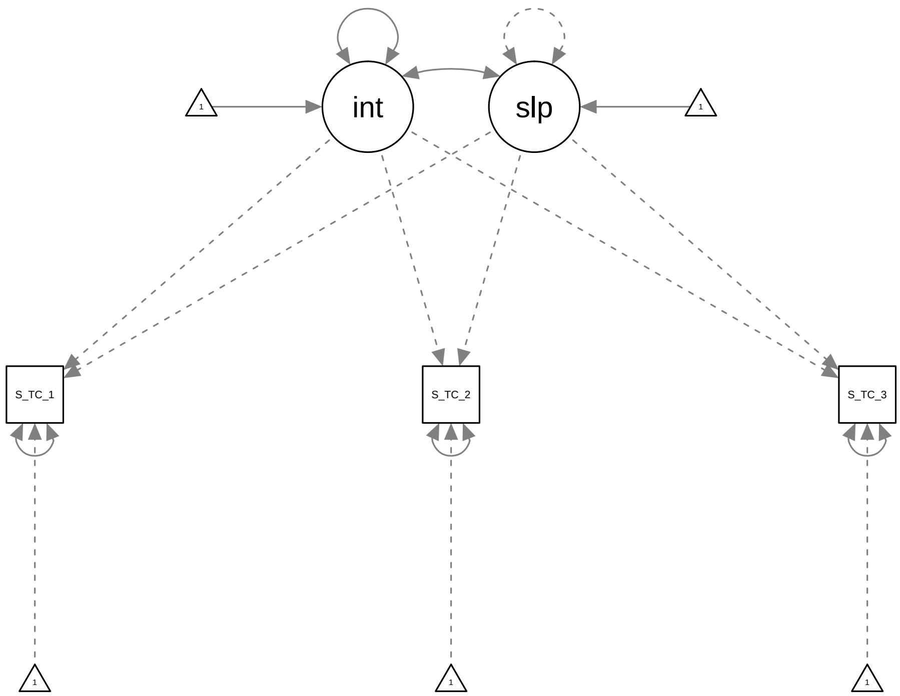
```
## Data: mydata
## Models:
## mod.HLM: Sem_TotalCorrect ~ Timepoint + (1 | ID2)
## mod.HLM3: Sem_TotalCorrect ~ Timepoint + Age_at_time_of_testing + (1 |
## mod.HLM3:     ID2)
## mod.HLM2: Sem_TotalCorrect ~ Timepoint + (Timepoint | ID2)
##          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
## mod.HLM   4 1415.5 1428.7 -703.74   1407.5
## mod.HLM3  5 1390.3 1406.8 -690.17   1380.3 27.144      1  1.888e-07 ***
## mod.HLM2  6 1418.5 1438.3 -703.27   1406.5  0.000      1          1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
#Growth model
mod.SEM <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
             slope =~ 0*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 2*Sem_TotalCorrect_3
             slope ~~ 0*slope ' #fixed slope, no variance

mod.SEM.fixed <- growth(mod.SEM, missing = "ML", data = mydata_wide)
summary(mod.SEM.fixed) #intercept = 34.93, slope = 3.85
```

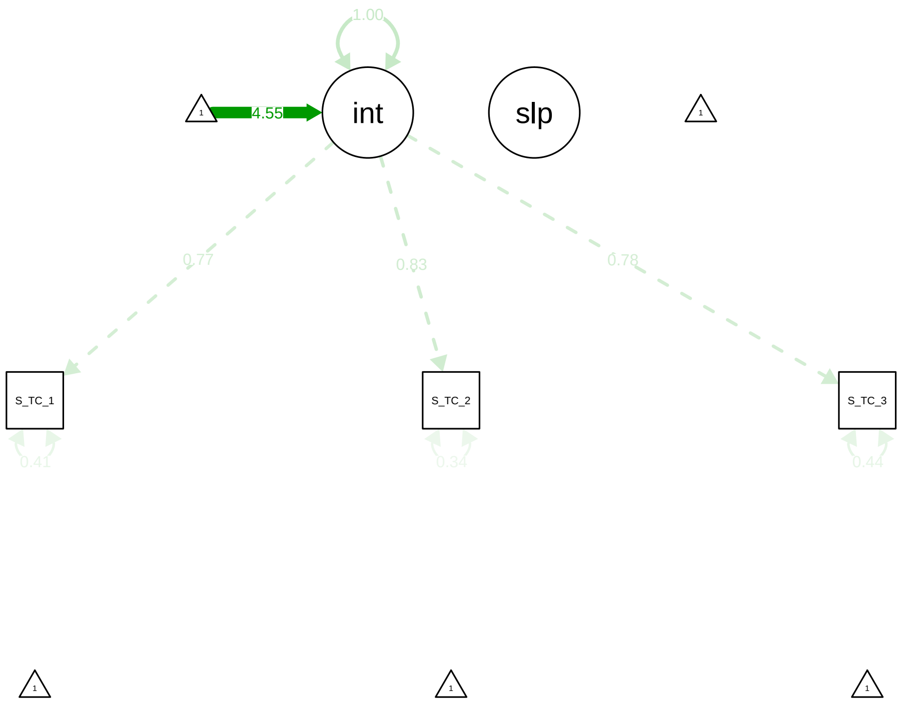
```
## lavaan (0.5-23.1097) converged normally after  68 iterations
##
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
```

```
##   Estimator                                         ML
##   Minimum Function Test Statistic                0.063
##   Degrees of freedom                                 2
##   P-value (Chi-square)                           0.969
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1    0.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    2.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope            -1.298    5.361   -0.242    0.809
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept        34.925    1.172   29.805    0.000
##     slope             3.845    0.564    6.815    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##     slope             0.000
##    .Sem_TtlCrrct_1   41.684   10.355    4.026    0.000
##    .Sem_TtlCrrct_2   28.989    7.920    3.660    0.000
##    .Sem_TtlCrrct_3   42.710   10.468    4.080    0.000
##     intercept        58.911   16.124    3.654    0.000
```

```
semPaths(mod.SEM.fixed)
```




```r
semPaths(mod.SEM.fixed, what = "std")
```

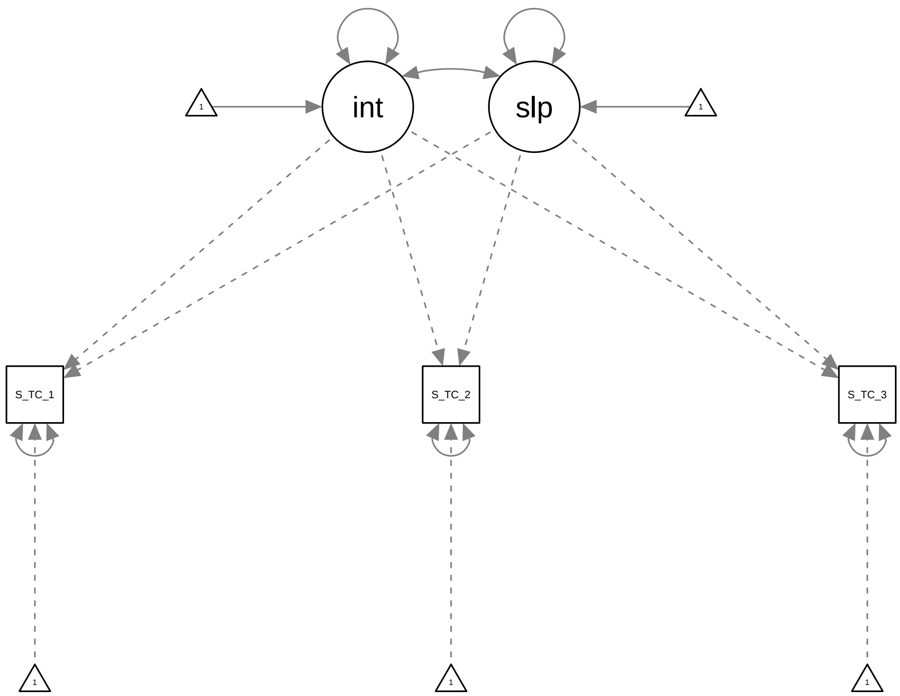



```r
mod.SEM2 <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
              slope =~ 0*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 2*Sem_TotalCorrect_3 ' #random slo

mod.SEM.random <- growth(mod.SEM2, missing = "ML", data = mydata_wide)
summary(mod.SEM.random) #intercept = 34.93, slope = 3.85
```

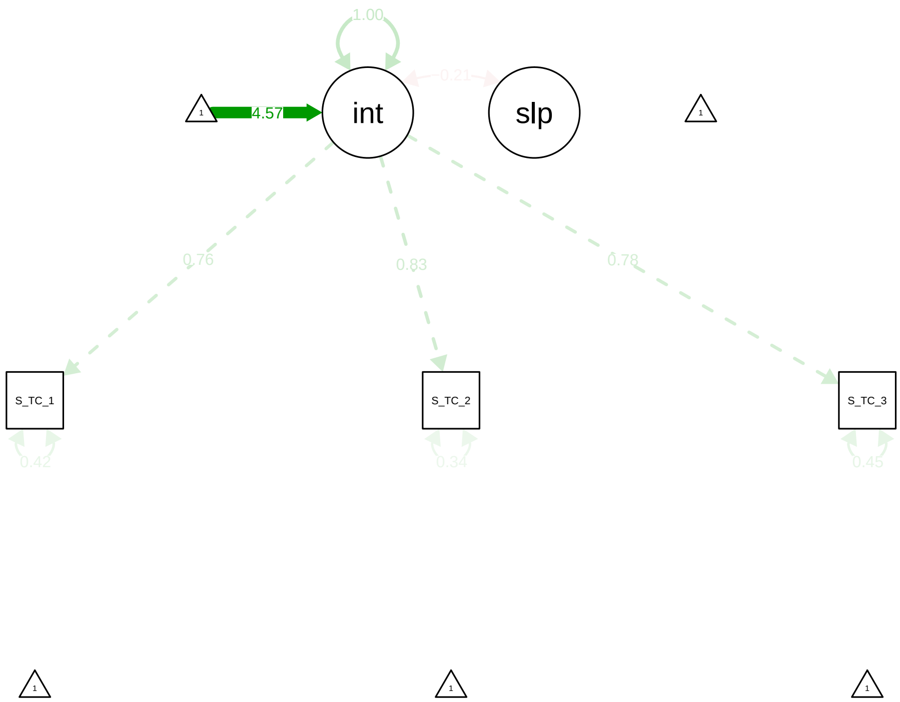
```
## lavaan (0.5-23.1097) converged normally after  78 iterations
##
```

```
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
##   Estimator                                         ML
##   Minimum Function Test Statistic                0.062
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.803
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                     Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1    0.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    2.000
##
## Covariances:
##                     Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope             -0.932   10.880   -0.086    0.932
##
## Intercepts:
##                     Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept         34.926    1.172   29.813    0.000
##     slope              3.846    0.564    6.824    0.000
##
## Variances:
##                     Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1   42.272   18.421    2.295    0.022
##    .Sem_TtlCrrct_2   28.836    8.842    3.261    0.001
##    .Sem_TtlCrrct_3   43.274   17.998    2.404    0.016
##     intercept         58.471   19.695    2.969    0.003
##     slope             -0.349    9.040   -0.039    0.969
```

```
semPaths(mod.SEM.random)
```




```
semPaths(mod.SEM.random, what = "std")
```



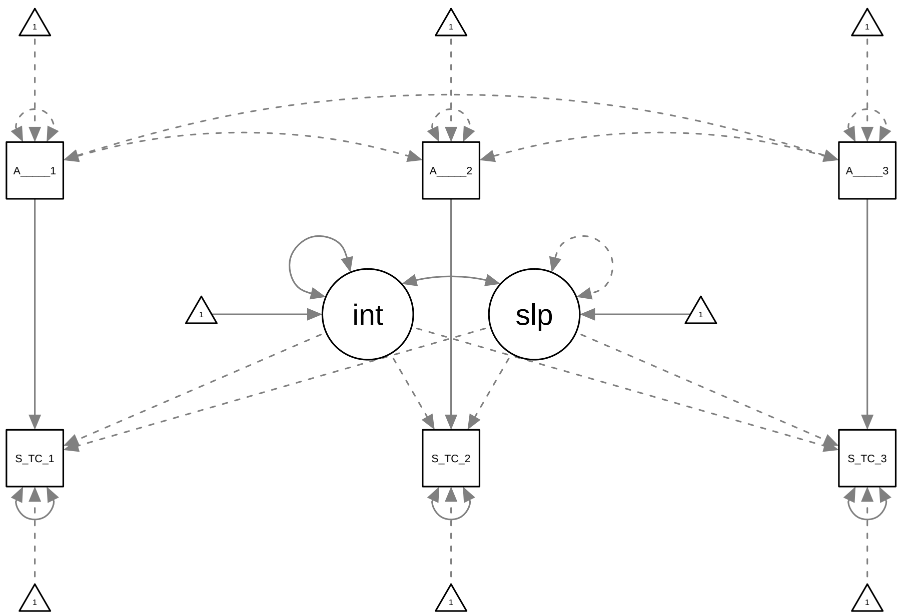

```
mod.SEM3 <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
              slope =~ 0*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 2*Sem_TotalCorrect_3
              slope ~~ 0*slope

              Sem_TotalCorrect_1 ~ Age_at_time_of_testing_1
              Sem_TotalCorrect_2 ~ Age_at_time_of_testing_2
              Sem_TotalCorrect_3 ~ Age_at_time_of_testing_3 ' #fixed slope, with covariate
```

```r
mod.SEM.cov <- growth(mod.SEM3, missing = "ML", data = mydata_wide)
summary(mod.SEM.cov) #intercept = 12.38, slope = 6.65
```

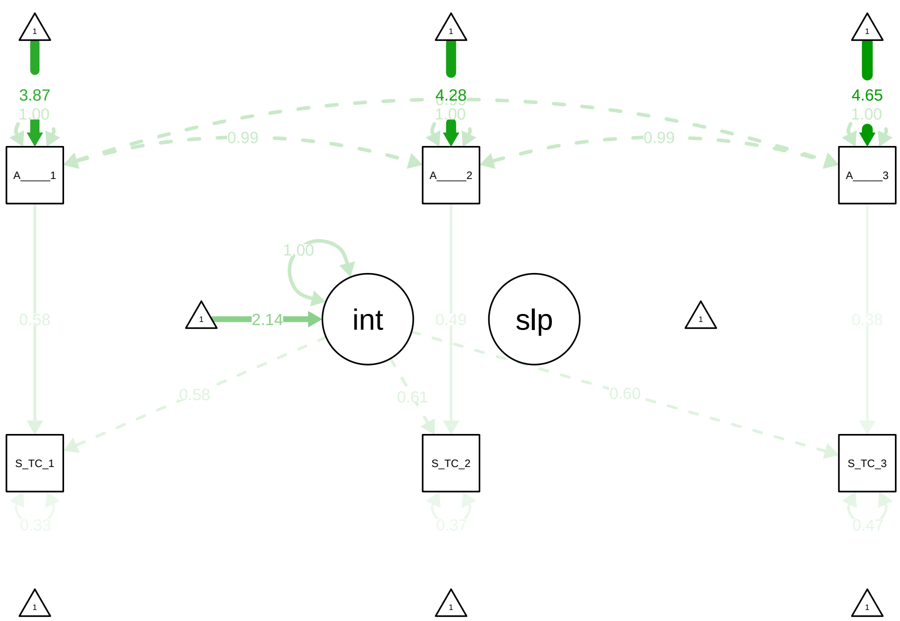
```
## lavaan (0.5-23.1097) converged normally after  67 iterations
##
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               20.725
##   Degrees of freedom                                 8
##   P-value (Chi-square)                           0.008
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1    0.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    2.000
##
## Regressions:
##                        Estimate  Std.Err  z-value  P(>|z|)
##   Sem_TotalCorrect_1 ~
##     Ag_t_tm_f_ts_1        1.904    0.316    6.027    0.000
##   Sem_TotalCorrect_2 ~
##     Ag_t_tm_f_ts_2        1.517    0.279    5.445    0.000
##   Sem_TotalCorrect_3 ~
##     Ag_t_tm_f_ts_3        1.173    0.336    3.487    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope             0.659    4.249    0.155    0.877
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept        12.377    3.863    3.204    0.001
##     slope             6.652    2.490    2.672    0.008
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
```

```
##     slope              0.000
##    .Sem_TtlCrrct_1    32.435    8.948    3.625    0.000
##    .Sem_TtlCrrct_2    32.778    8.019    4.087    0.000
##    .Sem_TtlCrrct_3    44.160   10.815    4.083    0.000
##     intercept         33.307   11.097    3.001    0.003
```

```
semPaths(mod.SEM.cov)
```




```
semPaths(mod.SEM.cov, what = "std")
```




```
anova(mod.SEM.fixed, mod.SEM.random, mod.SEM.cov) #fixed slope, covariate model is preferred
```

```
## Chi Square Difference Test
##
##                Df AIC BIC   Chisq Chisq diff Df diff Pr(>Chisq)
## mod.SEM.random  1          0.0620
```

```
## mod.SEM.fixed   2          0.0635     0.0015       1    0.96914
## mod.SEM.cov     8         20.7247    20.6613       6    0.00211 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

5. Constrain the residual variances to be equal. Does this change the fit of your model?

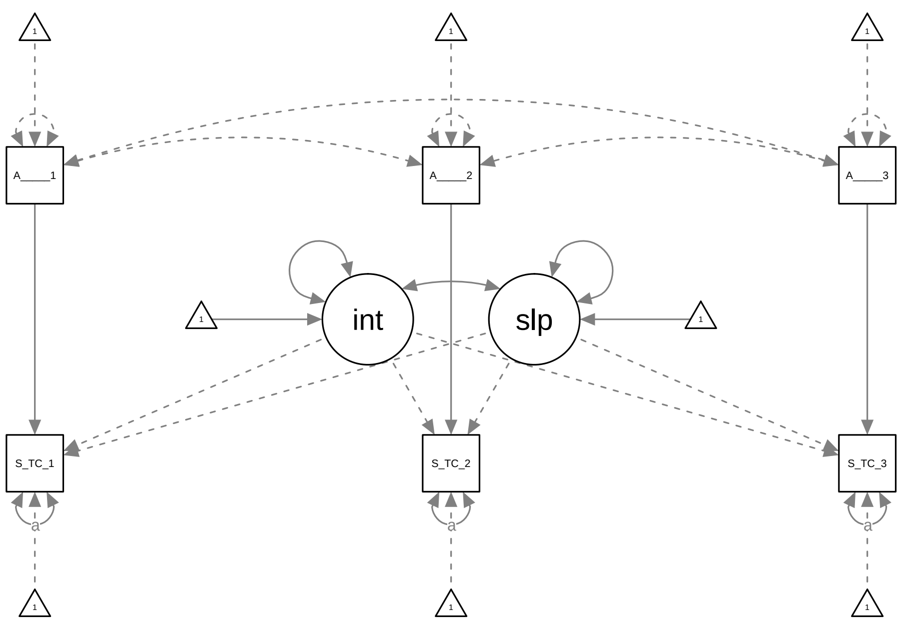
Constraining the residual variances does not significantly change model fit. LogLikelihood tests indicate that a simpler model, where residual variances are allowed to vary, is preferred to a more complex model where they are constrained to be equal.

```
mod.SEM4 <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
              slope =~ 0*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 2*Sem_TotalCorrect_3

              Sem_TotalCorrect_1 ~ Age_at_time_of_testing_1
              Sem_TotalCorrect_2 ~ Age_at_time_of_testing_2
              Sem_TotalCorrect_3 ~ Age_at_time_of_testing_3

              Sem_TotalCorrect_1 ~~ a*Sem_TotalCorrect_1
              Sem_TotalCorrect_2 ~~ a*Sem_TotalCorrect_2
              Sem_TotalCorrect_3 ~~ a*Sem_TotalCorrect_3 ' #random slope, with covariate, residual vari

mod.SEM.cov2 <- growth(mod.SEM4, missing = "ML", data = mydata_wide)
summary(mod.SEM.cov2) #intercept = 12.53, slope = 6.23
```

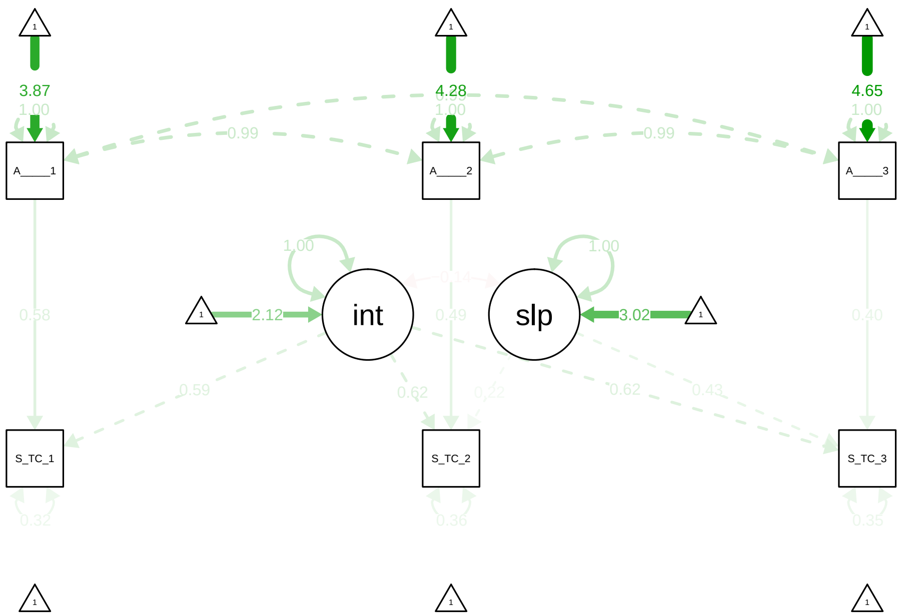
```
## lavaan (0.5-23.1097) converged normally after  86 iterations
##
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
##   Estimator                                         ML
##   Minimum Function Test Statistic               20.579
##   Degrees of freedom                                 9
##   P-value (Chi-square)                           0.015
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1    0.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    2.000
##
## Regressions:
```

```
##                          Estimate  Std.Err  z-value  P(>|z|)
##   Sem_TotalCorrect_1 ~
##     Ag_t_tm_f_ts_1        1.892    0.316    5.991    0.000
##   Sem_TotalCorrect_2 ~
##     Ag_t_tm_f_ts_2        1.535    0.277    5.544    0.000
##   Sem_TotalCorrect_3 ~
##     Ag_t_tm_f_ts_3        1.218    0.335    3.631    0.000
##
## Covariances:
##                      Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope              -1.747    5.666   -0.308    0.758
##
## Intercepts:
##                      Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept        12.534    3.858    3.249    0.001
##     slope             6.234    2.513    2.481    0.013
##
## Variances:
##                      Estimate  Std.Err  z-value  P(>|z|)
##    .Sm_TtlCr_1 (a)   32.259    5.580    5.781    0.000
##    .Sm_TtlCr_2 (a)   32.259    5.580    5.781    0.000
##    .Sm_TtlCr_3 (a)   32.259    5.580    5.781    0.000
##     intercept        34.943   11.672    2.994    0.003
##     slope             4.275    4.506    0.949    0.343
```

```r
semPaths(mod.SEM.cov2)
```




```r
semPaths(mod.SEM.cov2, what = "std")
```




```r
anova(mod.SEM.cov, mod.SEM.cov2) #simpler model (mod.SEM.cov) is preferred
```

```
## Chi Square Difference Test
##
##              Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## mod.SEM.cov   8         20.725
## mod.SEM.cov2  9         20.579   -0.14545       1          1
```

6. Contrain your slope to be fixed, not random. How does this change your model?

Constraining slopes to be fixed does not sigificantly change my model (p = .97).

```r
anova(mod.SEM.fixed, mod.SEM.random) #see problem #4 for model specifics
```

```
## Chi Square Difference Test
##
##                Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## mod.SEM.random  1         0.0620
## mod.SEM.fixed   2         0.0635  0.0014966       1     0.9691
```

7. Change the time metric in your SEM growth model. How does that change your estimates? Does it change your fit statistics?

I changed my time metric such that the intercept was centered at TP 3 rather than TP1. This increased the intercept (which makes sense, given age-related change) but did not change model fit.

I also changed my time metric such that "duration" between TP3 & TP2 > "duration" between TP2 & TP1. This had little impact on the intercept, but decreased the slope and improved model fit.

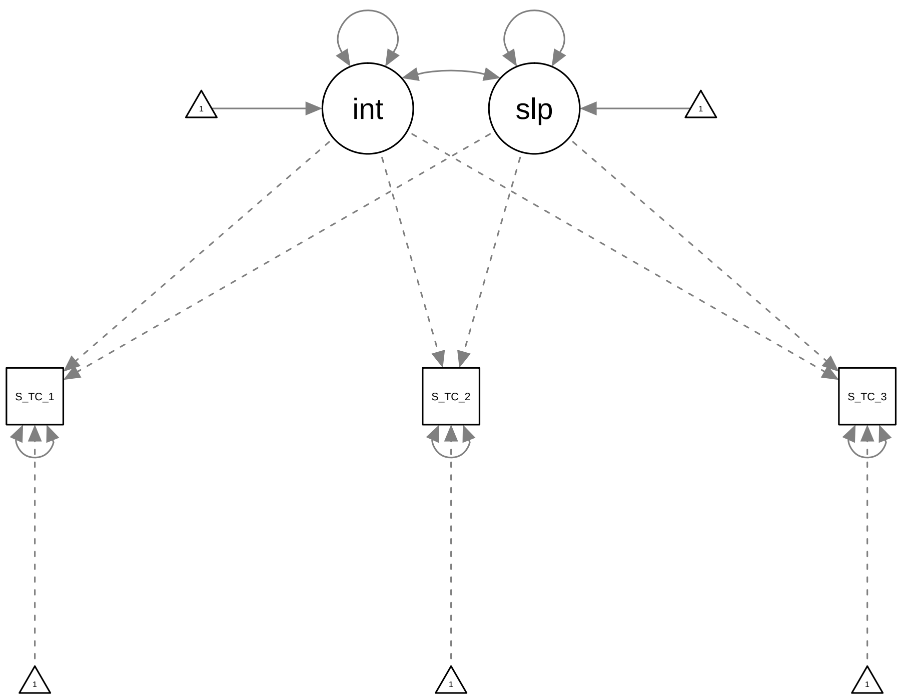
```r
# mydata_wide$time_1 <- 0
# mydata_wide$time_2 <- as.numeric(mydata_wide$Age_at_time_of_testing_2) -
#   as.numeric(mydata_wide$Age_at_time_of_testing_1)
# mydata_wide$time_3 <- as.numeric(mydata_wide$Age_at_time_of_testing_3) -
```

```r
#   as.numeric(mydata_wide$Age_at_time_of_testing_1)

mod.SEM5 <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
              slope =~ -2*Sem_TotalCorrect_1 + -1*Sem_TotalCorrect_2 + 0*Sem_TotalCorrect_3 '

mod.SEM.time <- growth(mod.SEM5, missing = "ML", data = mydata_wide)
summary(mod.SEM.time)
```

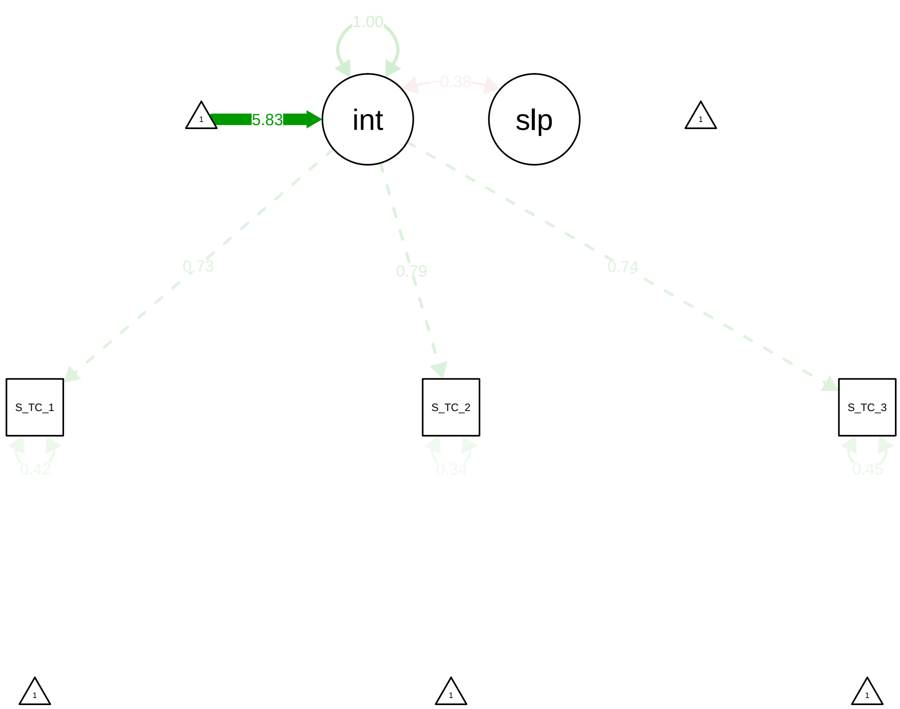
```
## lavaan (0.5-23.1097) converged normally after  93 iterations
##
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
##   Estimator                                         ML
##   Minimum Function Test Statistic                0.062
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.803
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1   -2.000
##     Sem_TtlCrrct_2   -1.000
##     Sem_TtlCrrct_3    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope            -1.630   10.117   -0.161    0.872
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept        42.618    1.146   37.181    0.000
##     slope             3.846    0.564    6.824    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1   42.272   18.421    2.295    0.022
##    .Sem_TtlCrrct_2   28.836    8.842    3.261    0.001
##    .Sem_TtlCrrct_3   43.274   17.998    2.404    0.016
##     intercept        53.347   18.762    2.843    0.004
```

```
##     slope            -0.349    9.040    -0.039    0.969
```

```r
semPaths(mod.SEM.time)
```




```r
semPaths(mod.SEM.time, what = "std")
```




```r
anova(mod.SEM.random, mod.SEM.time)
```

```
## Chi Square Difference Test
##
##                  Df AIC BIC Chisq  Chisq diff Df diff Pr(>Chisq)
```

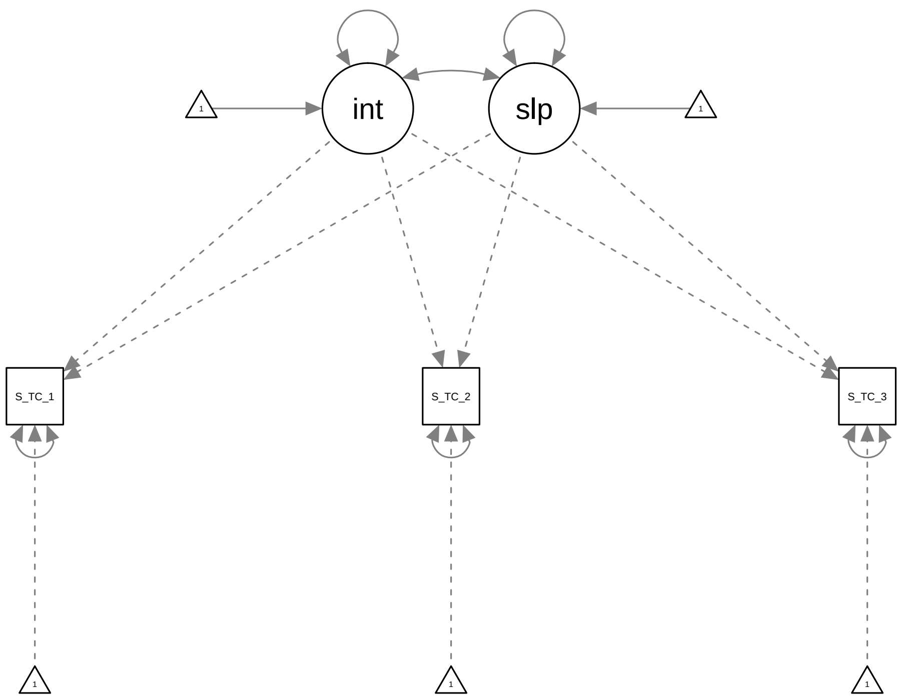
```
## mod.SEM.random  1         0.062
## mod.SEM.time    1         0.062 -2.3803e-13       0          1
```

```
mod.SEM6 <- ' intercept =~ 1*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 1*Sem_TotalCorrect_3
              slope =~ 0*Sem_TotalCorrect_1 + 1*Sem_TotalCorrect_2 + 6*Sem_TotalCorrect_3 '

mod.SEM.time2 <- growth(mod.SEM6, missing = "ML", data = mydata_wide)
summary(mod.SEM.time2)
```

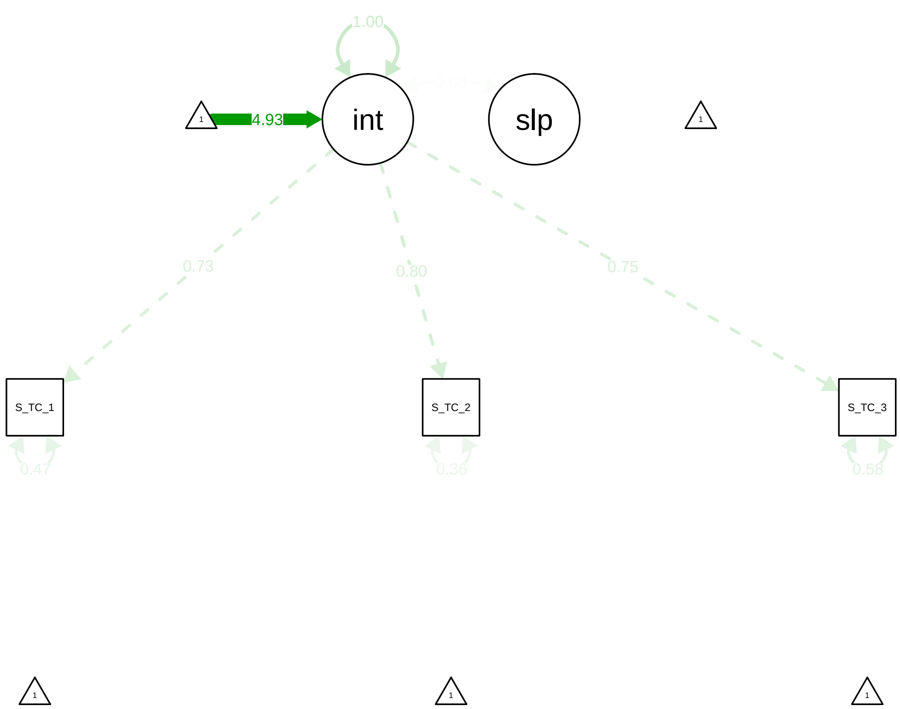
```
## lavaan (0.5-23.1097) converged normally after  89 iterations
##
##   Number of observations                            67
##
##   Number of missing patterns                         2
##
##   Estimator                                         ML
##   Minimum Function Test Statistic                8.210
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.004
##
## Parameter Estimates:
##
##   Information                                 Observed
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Sem_TtlCrrct_1    1.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    1.000
##   slope =~
##     Sem_TtlCrrct_1    0.000
##     Sem_TtlCrrct_2    1.000
##     Sem_TtlCrrct_3    6.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope             0.411    2.280    0.180    0.857
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1    0.000
##    .Sem_TtlCrrct_2    0.000
##    .Sem_TtlCrrct_3    0.000
##     intercept        36.518    1.155   31.627    0.000
##     slope             1.067    0.184    5.794    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Sem_TtlCrrct_1   48.629   14.123    3.443    0.001
##    .Sem_TtlCrrct_2   31.574   10.548    2.994    0.003
##    .Sem_TtlCrrct_3   56.364   63.287    0.891    0.373
##     intercept        54.935   14.286    3.845    0.000
```

```
##     slope             -0.535    2.126   -0.252    0.801
```

```r
semPaths(mod.SEM.time2)
```




```r
semPaths(mod.SEM.time2, what = "std")
```




```r
anova(mod.SEM.random, mod.SEM.time2)
```

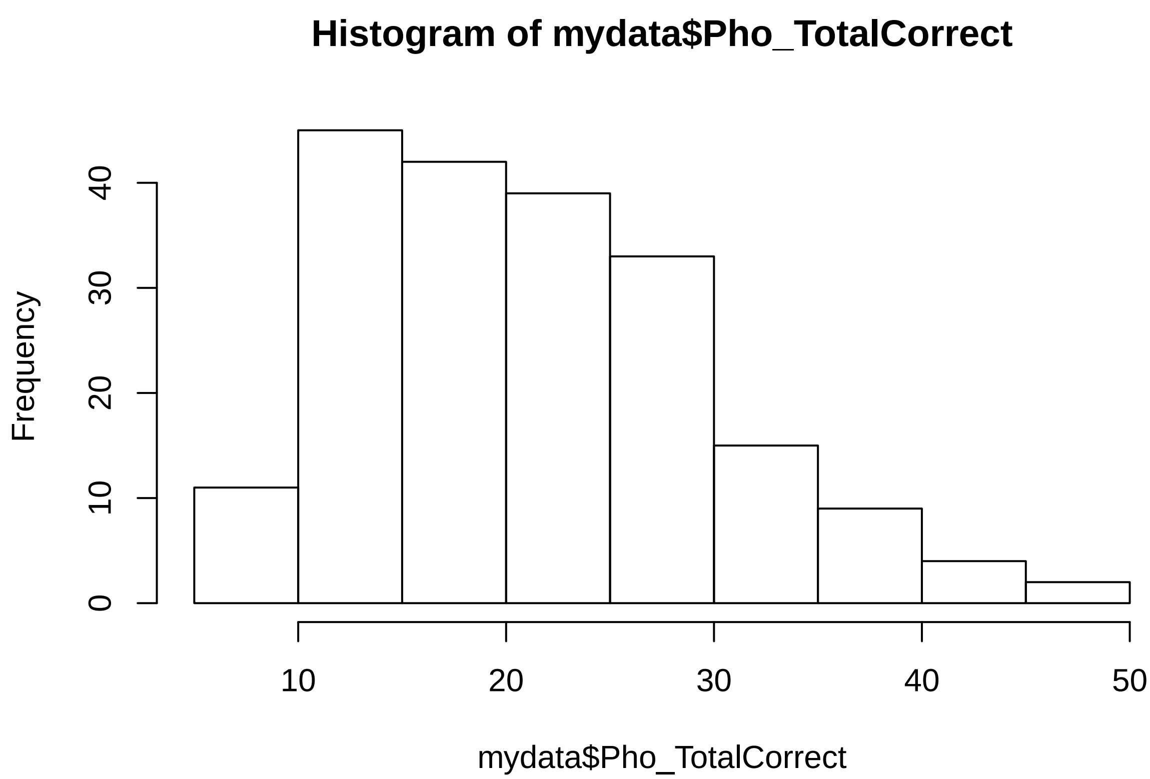
```
## Chi Square Difference Test
##
##                  Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
```

```
## mod.SEM.random  1         0.0620
## mod.SEM.time2   1         8.2098     8.1478       0  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8. Try a different type of estimation (see lavaan tutorial for details). How does that change your model?

Despite the fact that Pho_TotalCorrect is non-normal, changing the esimator has no effect on model fit, intercept, or slope.

```
#What if we want to try to model non-normal data?
hist(mydata$Pho_TotalCorrect)
```

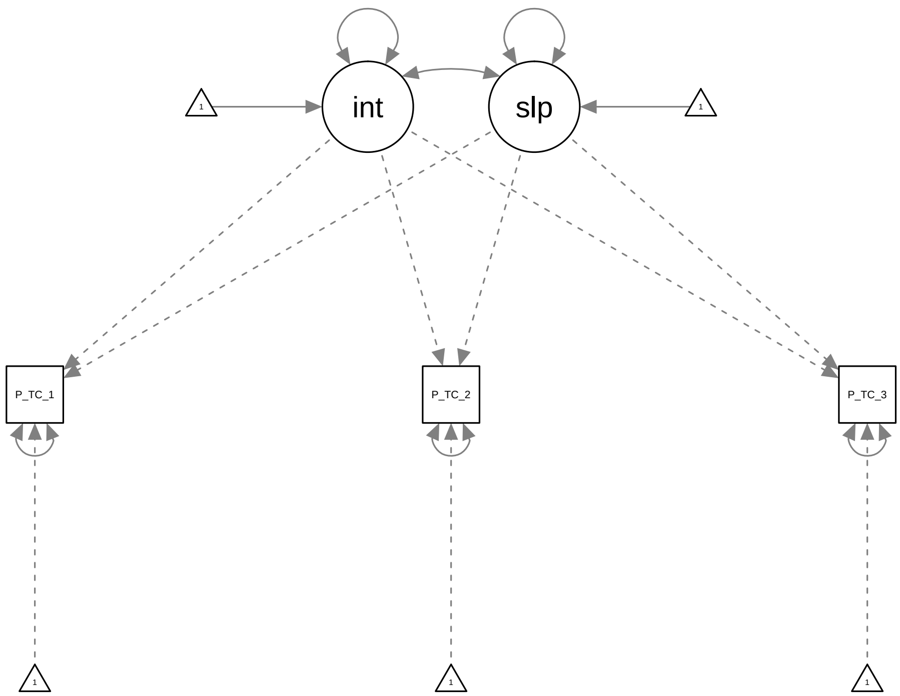



```
mod.SEM7 <- ' intercept =~ 1*Pho_TotalCorrect_1 + 1*Pho_TotalCorrect_2 + 1*Pho_TotalCorrect_3
              slope =~ -2*Pho_TotalCorrect_1 + -1*Pho_TotalCorrect_2 + 0*Pho_TotalCorrect_3 '

mod.SEM.MLM <- growth(mod.SEM7, estimator = "MLM", data = mydata_wide)
mod.SEM.ML <- growth(mod.SEM7, estimator = "ML", data = mydata_wide)

summary(mod.SEM.MLM)
```

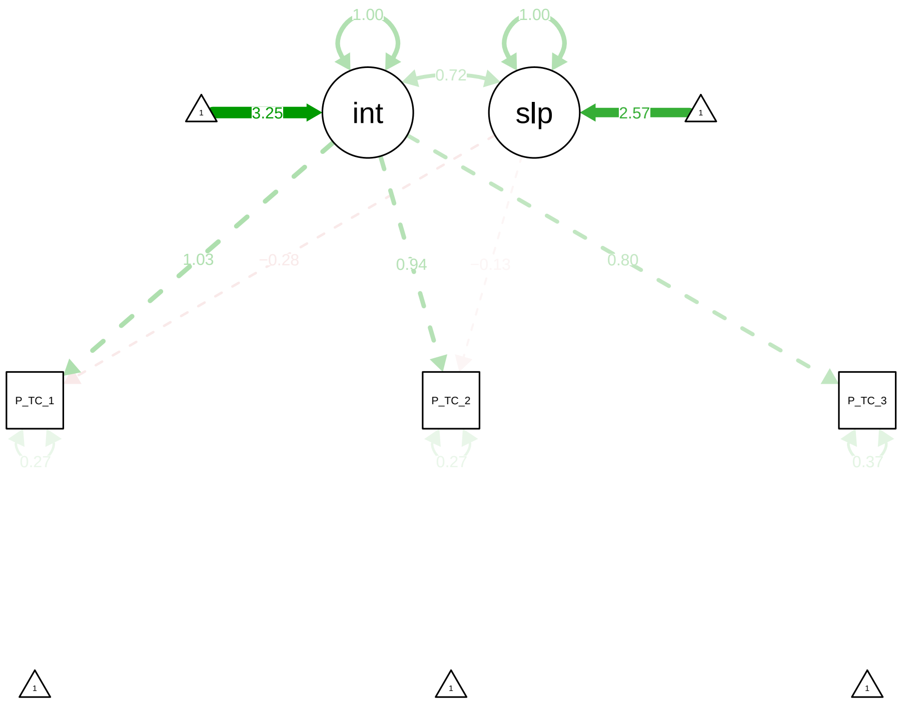
```
## lavaan (0.5-23.1097) converged normally after  80 iterations
##
##                                                   Used       Total
##   Number of observations                            66          67
##
##   Estimator                                         ML      Robust
##   Minimum Function Test Statistic                4.151       4.420
##   Degrees of freedom                                 1           1
```

```
##   P-value (Chi-square)                           0.042       0.036
##   Scaling correction factor                                 0.939
##     for the Satorra-Bentler correction
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                           Robust.sem
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Pho_TtlCrrct_1    1.000
##     Pho_TtlCrrct_2    1.000
##     Pho_TtlCrrct_3    1.000
##   slope =~
##     Pho_TtlCrrct_1   -2.000
##     Pho_TtlCrrct_2   -1.000
##     Pho_TtlCrrct_3    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope             5.513    6.633    0.831    0.406
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Pho_TtlCrrct_1    0.000
##    .Pho_TtlCrrct_2    0.000
##    .Pho_TtlCrrct_3    0.000
##     intercept        24.238    1.101   22.023    0.000
##     slope             2.622    0.429    6.113    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Pho_TtlCrrct_1   14.290    9.002    1.588    0.112
##    .Pho_TtlCrrct_2   17.235    5.583    3.087    0.002
##    .Pho_TtlCrrct_3   32.117   12.723    2.524    0.012
##     intercept        55.557   15.868    3.501    0.000
##     slope             1.043    5.083    0.205    0.837
```

```
semPaths(mod.SEM.MLM)
```




```r
semPaths(mod.SEM.MLM, what = "std")
```

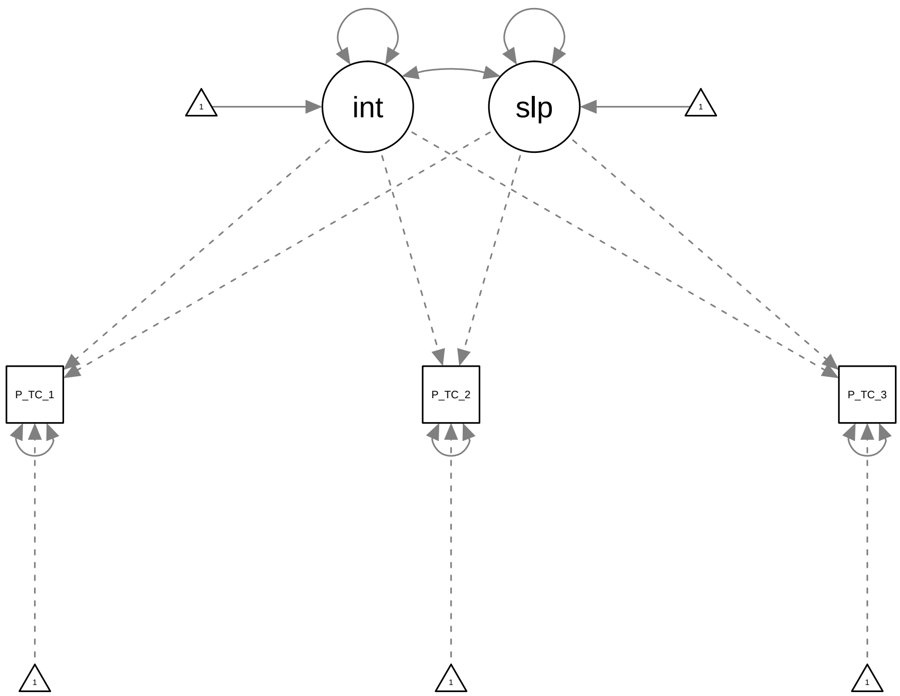



```r
summary(mod.SEM.ML)
```

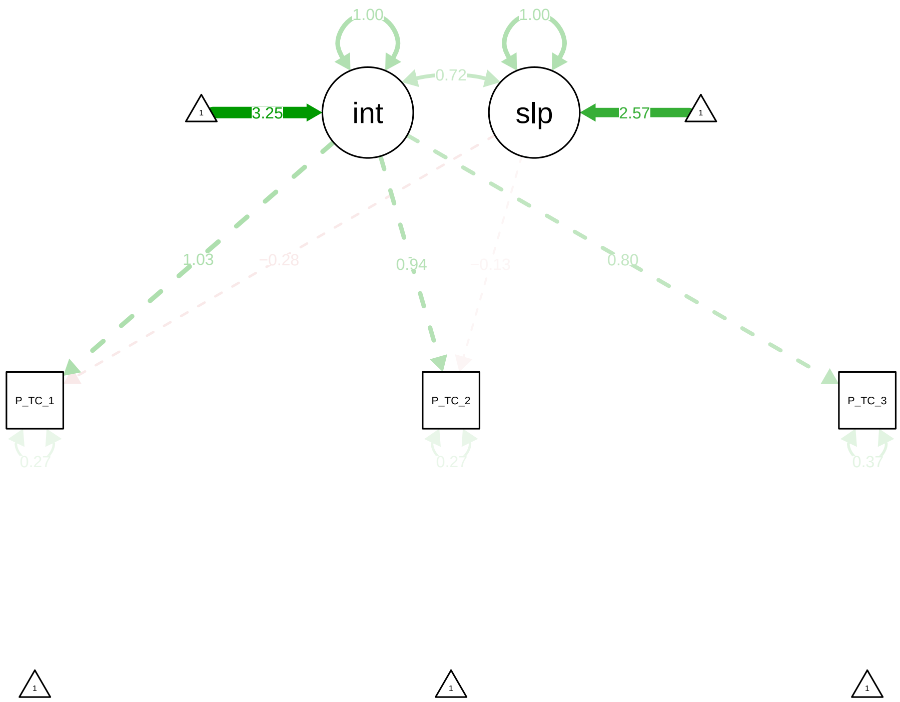
```
## lavaan (0.5-23.1097) converged normally after  80 iterations
##
##                                                   Used       Total
##   Number of observations                            66          67
##
##   Estimator                                         ML
```

```
##   Minimum Function Test Statistic                4.151
##   Degrees of freedom                                 1
##   P-value (Chi-square)                           0.042
##
## Parameter Estimates:
##
##   Information                                 Expected
##   Standard Errors                             Standard
##
## Latent Variables:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept =~
##     Pho_TtlCrrct_1    1.000
##     Pho_TtlCrrct_2    1.000
##     Pho_TtlCrrct_3    1.000
##   slope =~
##     Pho_TtlCrrct_1   -2.000
##     Pho_TtlCrrct_2   -1.000
##     Pho_TtlCrrct_3    0.000
##
## Covariances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##   intercept ~~
##     slope             5.513    6.612    0.834    0.404
##
## Intercepts:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Pho_TtlCrrct_1    0.000
##    .Pho_TtlCrrct_2    0.000
##    .Pho_TtlCrrct_3    0.000
##     intercept        24.238    1.092   22.192    0.000
##     slope             2.622    0.426    6.159    0.000
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)
##    .Pho_TtlCrrct_1   14.290    9.691    1.475    0.140
##    .Pho_TtlCrrct_2   17.235    5.652    3.050    0.002
##    .Pho_TtlCrrct_3   32.117   12.994    2.472    0.013
##     intercept        55.557   15.203    3.654    0.000
##     slope             1.043    5.427    0.192    0.848
```

```r
semPaths(mod.SEM.ML)
```




```r
semPaths(mod.SEM.ML, what = "std")
```




```r
anova(mod.SEM.MLM, mod.SEM.ML)
```

```
## Chi Square Difference Test
##
##              Df AIC BIC  Chisq Chisq diff Df diff Pr(>Chisq)
## mod.SEM.MLM  1          4.1505
## mod.SEM.ML   1          4.1505          0       0          1
```

9. Provide semplots for each of the models

Incorporated throughout code (above)