Assignment 5: SEM

Abhilasha Kumar

November 1, 2017

1 Reading the Data

We first create a Wave variable that bins the months into 4-month intervals, so that we have 3 time points for each person.

```
> cell_demo = read.csv("cell_demo.csv", header = TRUE, sep = ",")
> cell = read.csv("cell_withitems_complete.csv", header = TRUE, sep = ",")
> cell = merge(cell, cell_demo, by = "ID")
> cell$ID = as.factor(as.character(cell$ID))
> for(i in 1:nrow(cell)){
+  if(cell[i,12] <= 4){
+     cell[i,20] = 1
+   }
+   else if(cell[i,12] > 4 & cell[i,12]<= 8){
+     cell[i,20] = 2
+   }
+
+     else
+     cell[i,20] = 3
+ }
> colnames(cell)[20] = "Wave"
```

2 Converting Data to Wide Format

```
> library(plyr)
> library(dplyr)
> accuracy_time = group_by(cell, ID, Wave) %>%
+     summarise_at(vars(Accuracy), mean)
> messages_time = group_by(cell, ID, Wave) %>%
+     summarise_at(vars(Messages), mean)
> memorability_time = group_by(cell, ID, Wave) %>%
+     summarise_at(vars(Memorablity), mean)
> td_time = group_by(cell, ID, Wave) %>%
+     summarise_at(vars(TimeJudgmentDistance), mean)
> vivid_time = group_by(cell, ID, Wave) %>%
+     summarise_at(vars(Vividness), mean)
> ## long to wide
>
> acc_wide = tidyr::spread(accuracy_time, Wave, Accuracy)
> colnames(acc_wide) = c("Subject", "Q1_Acc", "Q2_Acc", "Q3_Acc")
> messages_wide = tidyr::spread(messages_time, Wave, Messages)
> colnames(messages_wide) = c("Subject", "Q1_mess", "Q2_mess", "Q3_mess")
> mem_wide = tidyr::spread(memorability_time, Wave, Memorablity)
> colnames(mem_wide) = c("Subject", "Q1_mem", "Q2_mem", "Q3_mem")
> td_wide = tidyr::spread(td_time, Wave, TimeJudgmentDistance)
```

```
> colnames(td_wide) = c("Subject", "Q1_td", "Q2_td", "Q3_td")
> vivid_wide = tidyr::spread(td_time, Wave, TimeJudgmentDistance)
> colnames(vivid_wide) = c("Subject", "Q1_v", "Q2_v", "Q3_v")
> cell_wide = Reduce(function(x,y) merge(x,y, all = TRUE), list(acc_wide, messages_wide, mem_wide, t
> cell_wide[,c(5:10)] = cell_wide[,c(5:10)]/100
> cell_wide[,c(11:16)] = cell_wide[,c(11:16)]/10
>
```

3 Measurement Model with 1 Time Point

Default Marker Variable Loading

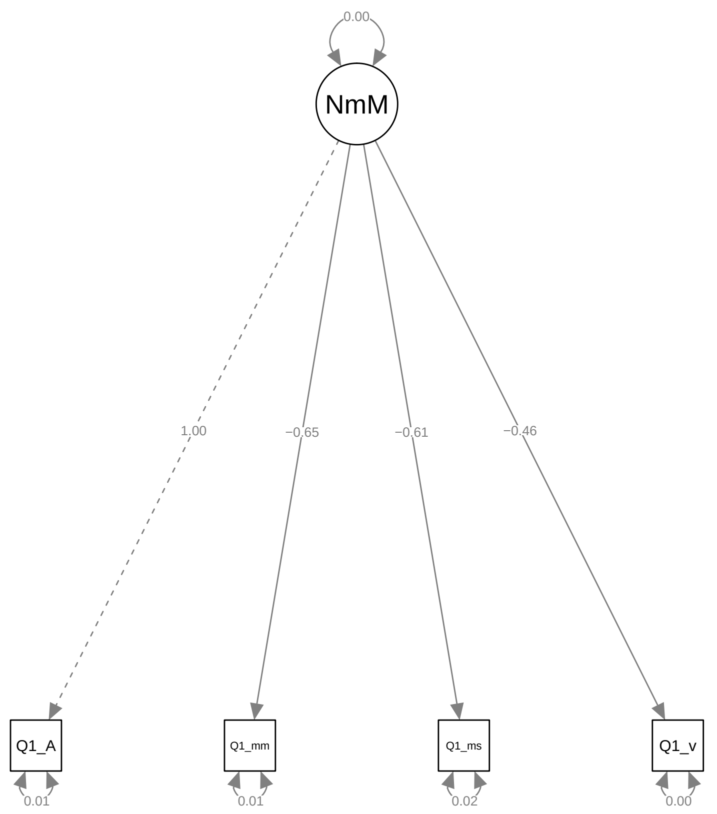
```
> cell.model.1 <- 'NameMemory =~ Q1_Acc + Q1_mem + Q1_mess + Q1_v'
> library(lavaan)
> library(semPlot)
> library(semTools)
> m_marker = lavaan::cfa(cell.model.1,  data = cell_wide)
> summary(m_marker, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  60 iterations

  Number of observations                            44

  Estimator                                         ML
  Minimum Function Test Statistic                0.639
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.726

Model test baseline model:

  Minimum Function Test Statistic                3.326
  Degrees of freedom                                 6
  P-value                                        0.767

User model versus baseline model:

  Comparative Fit Index (CFI)                    1.000
  Tucker-Lewis Index (TLI)                      -0.527

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                182.765
  Loglikelihood unrestricted model (H1)        183.084

  Number of free parameters                          8
  Akaike (AIC)                                -349.529
  Bayesian (BIC)                              -335.256
  Sample-size adjusted Bayesian (BIC)         -360.325

Root Mean Square Error of Approximation:

  RMSEA                                          0.000
  90 Percent Confidence Interval          0.000  0.213
  P-value RMSEA <= 0.05                          0.751

Standardized Root Mean Square Residual:
```

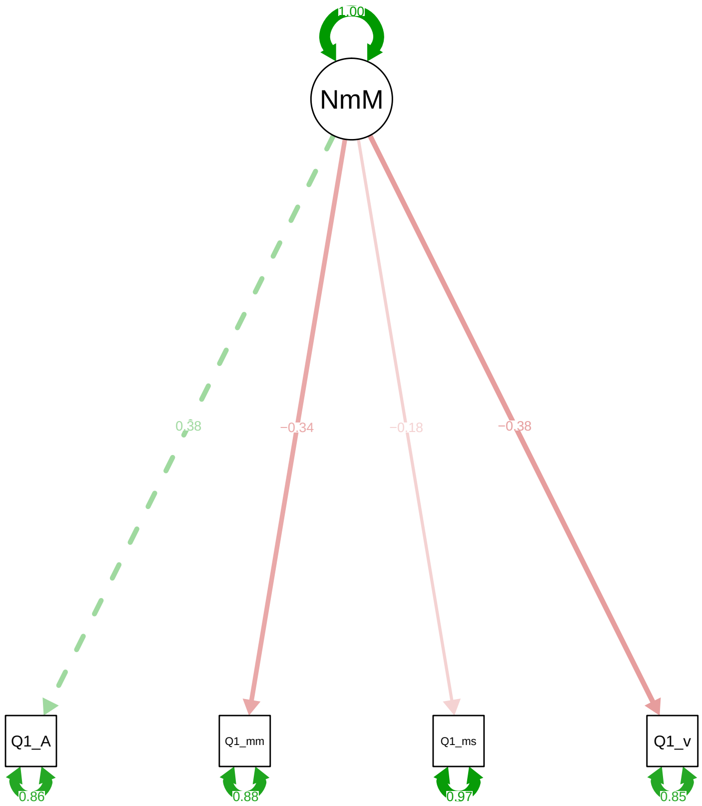
```
  SRMR                                           0.035

Parameter Estimates:

  Information                                 Expected
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  NameMemory =~
    Q1_Acc            1.000
    Q1_mem           -0.648    0.874   -0.742    0.458
    Q1_mess          -0.609    1.084   -0.562    0.574
    Q1_v             -0.461    0.637   -0.724    0.469

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.010    0.003    3.006    0.003
   .Q1_mem            0.005    0.002    3.371    0.001
   .Q1_mess           0.019    0.004    4.383    0.000
   .Q1_v              0.002    0.001    2.943    0.003
    NameMemory        0.002    0.003    0.581    0.561
```

Plotting

```
> semPaths(m_marker, whatLabels = "est")
```




```
> semPaths(m_marker, what = "std")
```



Fixed Factor Loading

```
> m_fixed = lavaan::cfa(cell.model.1, data = cell_wide, std.lv = TRUE)
> summary(m_fixed, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  51 iterations

  Number of observations                            44

  Estimator                                         ML
  Minimum Function Test Statistic                0.639
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.726

Model test baseline model:

  Minimum Function Test Statistic                3.326
  Degrees of freedom                                 6
  P-value                                        0.767

User model versus baseline model:

  Comparative Fit Index (CFI)                    1.000
  Tucker-Lewis Index (TLI)                      -0.527
```

```
Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                182.765
  Loglikelihood unrestricted model (H1)        183.084

  Number of free parameters                          8
  Akaike (AIC)                                -349.529
  Bayesian (BIC)                              -335.256
  Sample-size adjusted Bayesian (BIC)         -360.325

Root Mean Square Error of Approximation:

  RMSEA                                          0.000
  90 Percent Confidence Interval          0.000  0.213
  P-value RMSEA <= 0.05                          0.751

Standardized Root Mean Square Residual:

  SRMR                                           0.035

Parameter Estimates:

  Information                                 Expected
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  NameMemory =~
    Q1_Acc            0.040    0.035    1.162    0.245
    Q1_mem           -0.026    0.023   -1.119    0.263
    Q1_mess          -0.025    0.037   -0.662    0.508
    Q1_v             -0.019    0.016   -1.169    0.243

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.010    0.003    3.006    0.003
   .Q1_mem            0.005    0.002    3.371    0.001
   .Q1_mess           0.019    0.004    4.383    0.000
   .Q1_v              0.002    0.001    2.943    0.003
    NameMemory        1.000
```

When we keep the marker loading, the first indicator variable is weighted 1, and all the other estimates are compared to that first indicator. In fixed-factor loading we constrain the variance of the latent variable to 1. Thus, we see in the summary() output for both models that the variance estimates are different for the latent variable. Also note that the estimates for each of the indicator variables is different, which makes sense because the estimates in the are relative to Q1acc and they are not constrained in the fixed-factor model.

Note, however, that the overall fit statistics do not change i.e. both have the same values for SRMR and RMSEA.

4 Fit Statistics

The fit statistics for the marker model are as follows: RMSEA is 0, SRMR is 0.035, thus this a fairly good fitting model. CFI =1, TLI = -.527, high CFI indicates a good fit.

The degrees of freedom are 2.

5 Longitudinal CFA Model

Correlating Latent Factors

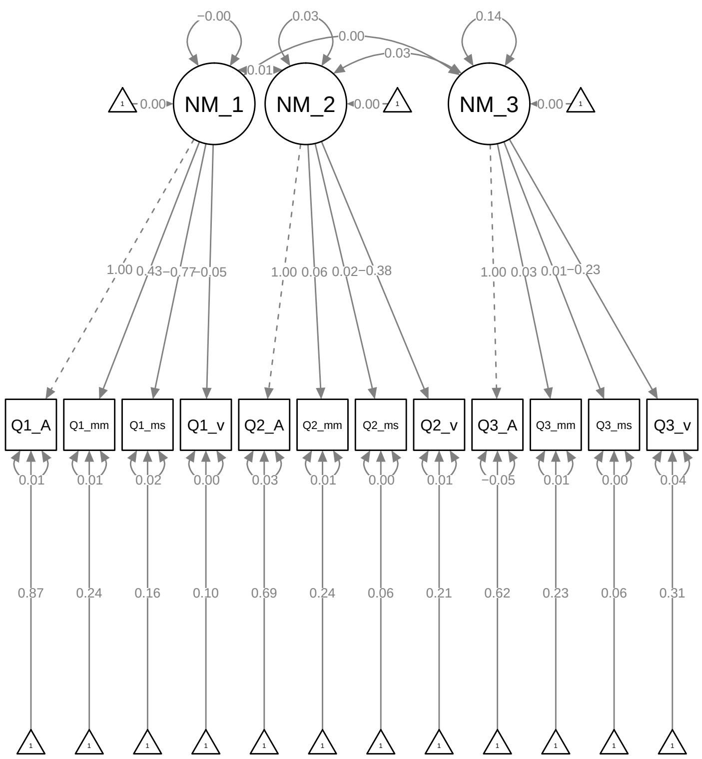
```
> long_model = 'NameMem_1 =~ Q1_Acc + Q1_mem + Q1_mess + Q1_v
+               NameMem_2 =~ Q2_Acc + Q2_mem + Q2_mess + Q2_v
+               NameMem_3 =~ Q3_Acc + Q3_mem + Q3_mess + Q3_v'
> long_model_fit = lavaan::cfa(long_model, missing = "ML", data = cell_wide)
> summary(long_model_fit, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after 175 iterations

  Number of observations                            44

  Number of missing patterns                         2

  Estimator                                         ML
  Minimum Function Test Statistic               68.879
  Degrees of freedom                                51
  P-value (Chi-square)                           0.048

Model test baseline model:

  Minimum Function Test Statistic              119.370
  Degrees of freedom                                66
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.665
  Tucker-Lewis Index (TLI)                       0.566

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                467.107
  Loglikelihood unrestricted model (H1)        501.547

  Number of free parameters                         39
  Akaike (AIC)                                -856.215
  Bayesian (BIC)                              -786.631
  Sample-size adjusted Bayesian (BIC)         -908.842

Root Mean Square Error of Approximation:

  RMSEA                                          0.089
  90 Percent Confidence Interval          0.008  0.139
  P-value RMSEA <= 0.05                          0.137

Standardized Root Mean Square Residual:

  SRMR                                           0.122

Parameter Estimates:

  Information                                 Observed
  Standard Errors                             Standard
```

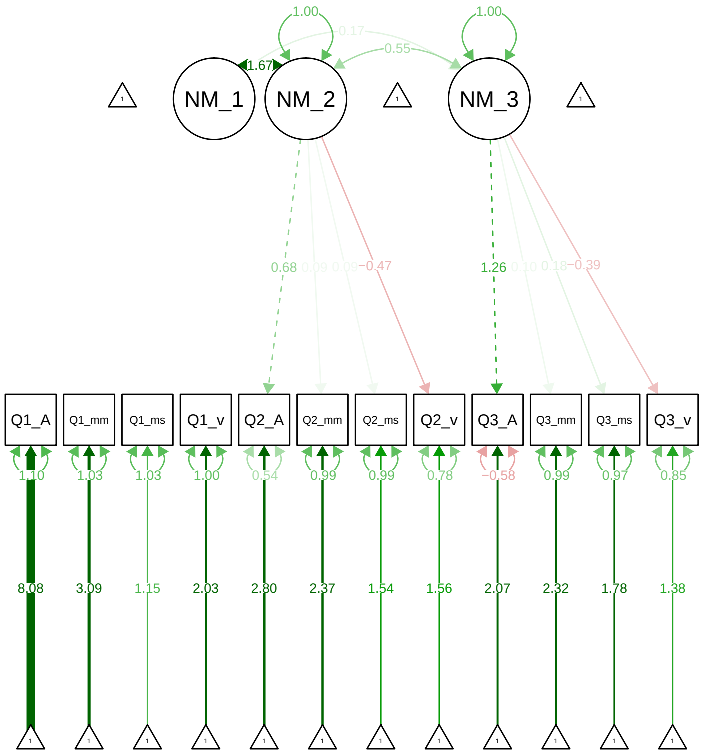
```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  NameMem_1 =~
    Q1_Acc            1.000
    Q1_mem            0.428    0.336    1.272    0.203
    Q1_mess          -0.771    0.577   -1.337    0.181
    Q1_v             -0.051    0.147   -0.346    0.729
  NameMem_2 =~
    Q2_Acc            1.000
    Q2_mem            0.056    0.119    0.472    0.637
    Q2_mess           0.022    0.047    0.456    0.648
    Q2_v             -0.380    0.137   -2.776    0.006
  NameMem_3 =~
    Q3_Acc            1.000
    Q3_mem            0.026    0.031    0.826    0.409
    Q3_mess           0.015    0.018    0.820    0.412
    Q3_v             -0.228    0.173   -1.313    0.189

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  NameMem_1 ~~
    NameMem_2         0.009    0.004    2.132    0.033
    NameMem_3         0.002    0.004    0.512    0.608
  NameMem_2 ~~
    NameMem_3         0.035    0.012    2.824    0.005

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.865    0.016   53.583    0.000
   .Q1_mem            0.237    0.012   20.510    0.000
   .Q1_mess           0.160    0.021    7.616    0.000
   .Q1_v              0.099    0.007   13.458    0.000
   .Q2_Acc            0.688    0.037   18.576    0.000
   .Q2_mem            0.242    0.015   15.699    0.000
   .Q2_mess           0.064    0.006   10.203    0.000
   .Q2_v              0.212    0.020   10.360    0.000
   .Q3_Acc            0.625    0.046   13.633    0.000
   .Q3_mem            0.229    0.015   15.184    0.000
   .Q3_mess           0.056    0.005   11.701    0.000
   .Q3_v              0.307    0.034    9.072    0.000
    NameMem_1         0.000
    NameMem_2         0.000
    NameMem_3         0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.013    0.004    3.480    0.001
   .Q1_mem            0.006    0.001    4.328    0.000
   .Q1_mess           0.020    0.005    4.384    0.000
   .Q1_v              0.002    0.001    4.683    0.000
   .Q2_Acc            0.032    0.013    2.581    0.010
   .Q2_mem            0.010    0.002    4.662    0.000
   .Q2_mess           0.002    0.000    4.665    0.000
   .Q2_v              0.014    0.003    4.368    0.000
   .Q3_Acc           -0.052    0.098   -0.537    0.591
   .Q3_mem            0.010    0.002    4.656    0.000
   .Q3_mess           0.001    0.000    4.600    0.000
   .Q3_v              0.042    0.010    4.157    0.000
```

NameMem_1	-0.001	0.002	-0.514	0.607
NameMem_2	0.028	0.015	1.851	0.064
NameMem_3	0.143	0.098	1.457	0.145

```
> semPaths(long_model_fit, whatLabels = "est")
```




```
> semPaths(long_model_fit, what = "std")
```



When we correlate the latent factors across time, we find Name Memory at Wave 2 is strongly predicted by Name Memory at Wave 1. We also find that the auto-regressive model fails to converge.

Auto-regressive

```
> long_model.2 = 'NameMem_1 =~ Q1_Acc + Q1_mem + Q1_mess + Q1_v
+                NameMem_2 =~ Q2_Acc + Q2_mem + Q2_mess + Q2_v
+                NameMem_3 =~ Q3_Acc + Q3_mem + Q3_mess + Q3_v
+
+                NameMem_2 ~ NameMem_1
+                '
>
> #auto_fit = lavaan::cfa(long_model.2, missing = "ML", data = cell_wide)
> #summary(auto_fit, fit.measures = TRUE) ## does not converge!
```

Plotting

```
> #semPaths(auto_fit, what = "est")
```

```
> #semPaths(auto_fit, what = "std")
```

6 SEM and HLM

HLM Models

```
> library(lme4)
> hlm_model_fixed = lmer(data = accuracy_time, Accuracy ~ Wave + (1|ID))
> summary(hlm_model_fixed)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Accuracy ~ Wave + (1 | ID)
   Data: accuracy_time

REML criterion at convergence: -8.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.58275 -0.56533  0.03377  0.63264  1.82296

Random effects:
 Groups   Name        Variance Std.Dev.
 ID       (Intercept) 0.01714  0.1309
 Residual             0.03855  0.1964
Number of obs: 131, groups:  ID, 44

Fixed effects:
            Estimate Std. Error t value
(Intercept)  0.96754    0.04945  19.567
Wave        -0.12115    0.02108  -5.747

Correlation of Fixed Effects:
     (Intr)
Wave -0.849
```

```
> hlm_model_random = lmer(data = accuracy_time, Accuracy ~ Wave + (Wave|ID))
> summary(hlm_model_random)
```

```
Linear mixed model fit by REML ['lmerMod']
Formula: Accuracy ~ Wave + (Wave | ID)
   Data: accuracy_time

REML criterion at convergence: -33.3

Scaled residuals:
     Min       1Q   Median       3Q      Max
-2.37678 -0.50800  0.01226  0.76407  1.61472

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 ID       (Intercept) 0.005708 0.07555
          Wave        0.012534 0.11196  -1.00
 Residual             0.025525 0.15977
Number of obs: 131, groups:  ID, 44

Fixed effects:
            Estimate Std. Error t value
```

```
(Intercept)  0.96814    0.03870  25.018
Wave        -0.12159    0.02414  -5.036

Correlation of Fixed Effects:
     (Intr)
Wave -0.838
```

SEM Models

Fixed SEM Model

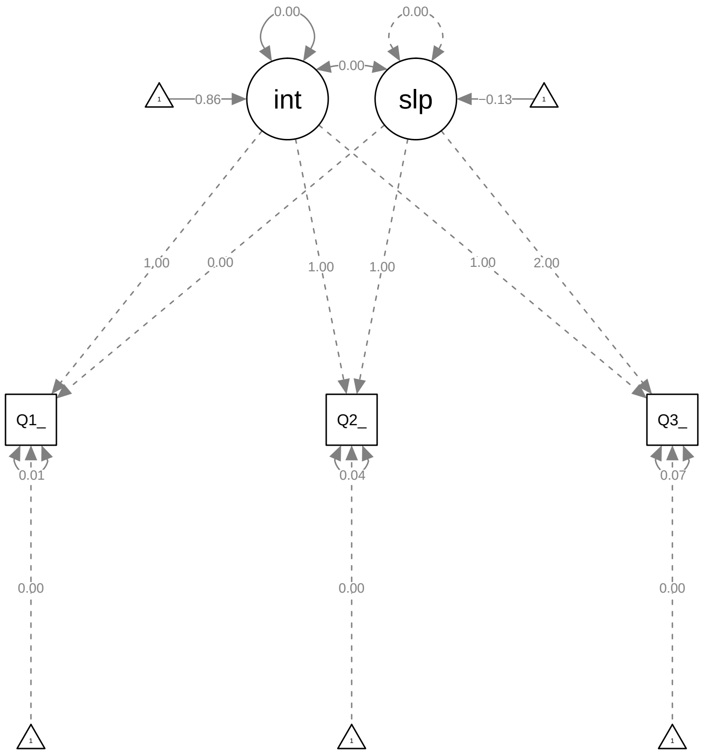
```
> ### FIXED SLOPE
>
> sem_fixed = 'intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+              slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc
+              slope ~~ 0*slope'
> sem_fixed_model = growth(sem_fixed, missing = "ML", data = cell_wide)
> summary(sem_fixed_model, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  46 iterations

  Number of observations                            44

  Number of missing patterns                         2

  Estimator                                         ML
  Minimum Function Test Statistic                9.596
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.008

Model test baseline model:

  Minimum Function Test Statistic               18.464
  Degrees of freedom                                 3
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.509
  Tucker-Lewis Index (TLI)                       0.263

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 30.097
  Loglikelihood unrestricted model (H1)         34.895

  Number of free parameters                          7
  Akaike (AIC)                                 -46.193
  Bayesian (BIC)                               -33.704
  Sample-size adjusted Bayesian (BIC)          -55.639

Root Mean Square Error of Approximation:

  RMSEA                                          0.294
  90 Percent Confidence Interval          0.127  0.491
  P-value RMSEA <= 0.05                          0.013
```

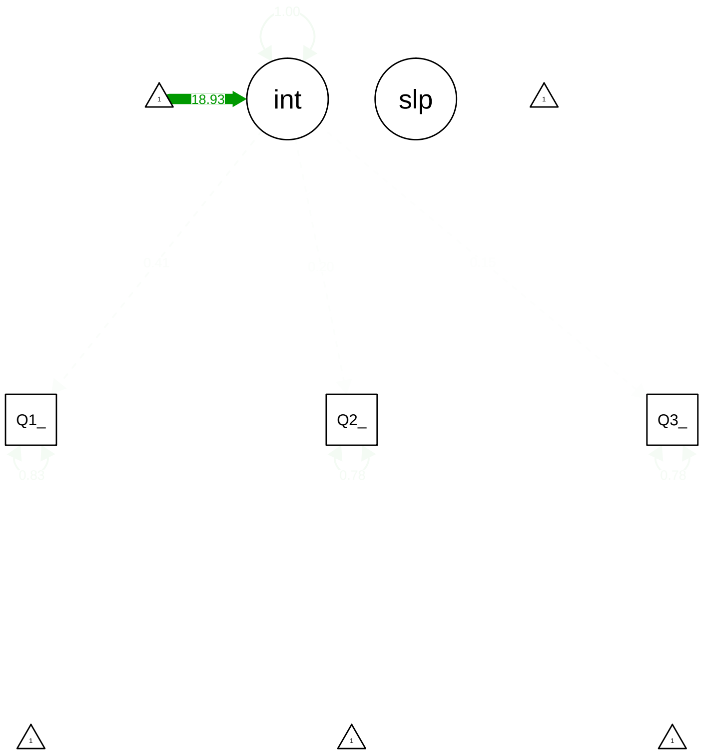
```
Standardized Root Mean Square Residual:

  SRMR                                           0.138

Parameter Estimates:

  Information                                 Observed
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope             0.005    0.004    1.149    0.251

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.000
   .Q2_Acc            0.000
   .Q3_Acc            0.000
    intercept         0.860    0.017   51.079    0.000
    slope            -0.135    0.020   -6.599    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
    slope             0.000
   .Q1_Acc            0.010    0.006    1.595    0.111
   .Q2_Acc            0.042    0.012    3.661    0.000
   .Q3_Acc            0.074    0.020    3.801    0.000
    intercept         0.002    0.006    0.349    0.727
```

Plotting

```
> semPaths(sem_fixed_model, whatLabels = "est")
```




```
> semPaths(sem_fixed_model, what = "std")
```



Random SEM Model

```
> ### RANDOM SLOPE
>
> sem_random = 'intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+                slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc'
> sem_random_model = growth(sem_random, missing = "ML", data = cell_wide)
> summary(sem_random_model, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  67 iterations

  Number of observations                            44

  Number of missing patterns                         2

  Estimator                                         ML
  Minimum Function Test Statistic                3.004
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.083

Model test baseline model:

  Minimum Function Test Statistic               18.464
  Degrees of freedom                                 3
  P-value                                        0.000
```

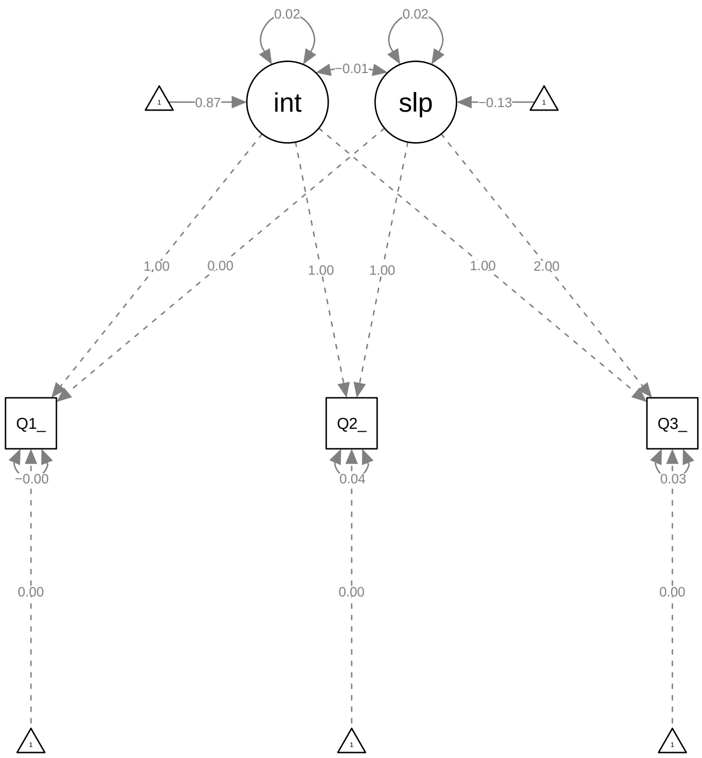
```
User model versus baseline model:

  Comparative Fit Index (CFI)                    0.870
  Tucker-Lewis Index (TLI)                       0.611

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 33.393
  Loglikelihood unrestricted model (H1)         34.895

  Number of free parameters                          8
  Akaike (AIC)                                 -50.785
  Bayesian (BIC)                               -36.512
  Sample-size adjusted Bayesian (BIC)          -61.581

Root Mean Square Error of Approximation:

  RMSEA                                          0.213
  90 Percent Confidence Interval          0.000  0.509
  P-value RMSEA <= 0.05                          0.100

Standardized Root Mean Square Residual:

  SRMR                                           0.070

Parameter Estimates:

  Information                                 Observed
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope            -0.005    0.005   -1.023    0.306

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.000
   .Q2_Acc            0.000
   .Q3_Acc            0.000
    intercept         0.868    0.017   52.132    0.000
    slope            -0.132    0.023   -5.713    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc           -0.004    0.007   -0.576    0.565
```

.Q2_Acc	0.040	0.010	3.933	0.000
.Q3_Acc	0.029	0.021	1.376	0.169
intercept	0.016	0.008	1.962	0.050
slope	0.017	0.006	2.590	0.010

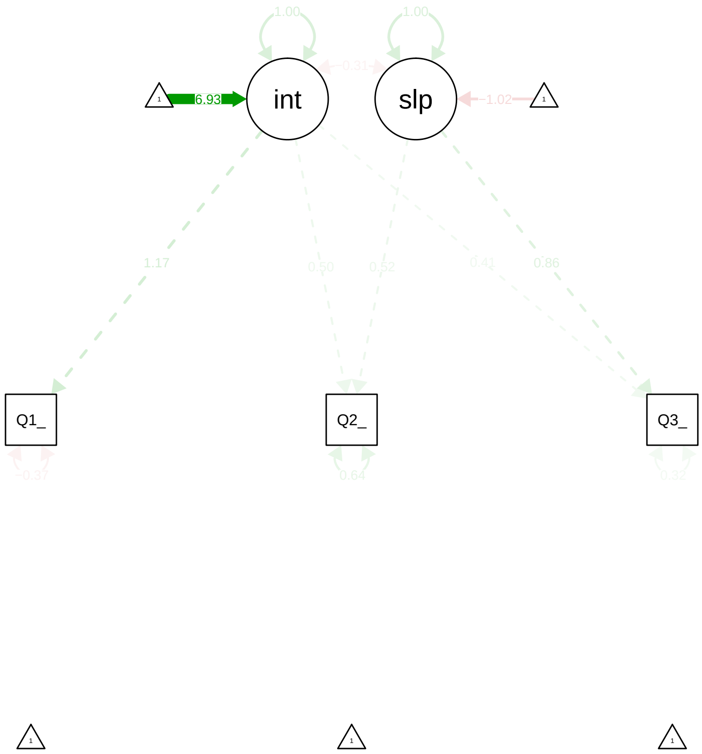
Changing from random to fixed (or the other way), changes the variance estimate of the slope. In the fixed SEM model, the variance estimate for the slope is 0 i.e. it is fixed, whereas is is 0.017 in the random SEM model. Further, also note that the estimates themselves of the intercept and slope change, although only slightly.

Plotting

```
> semPaths(sem_random_model, whatLabels = "est")
```




```
> semPaths(sem_random_model, what = "std")
```



6.1 Model Comparison

```
> anova(sem_fixed_model, sem_random_model) ## FIXED MODEL IS BETTER
```

```
Chi Square Difference Test

                 Df     AIC     BIC  Chisq Chisq diff
sem_random_model  1 -50.785 -36.512 3.0037
sem_fixed_model   2 -46.193 -33.704 9.5956     6.5919
                 Df diff Pr(>Chisq)
sem_random_model
sem_fixed_model        1    0.01024 *
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Comparing SEM and HLM models we realize that the estimates for the slope and intercept are close to each other in both models.

7 Constraining Residual Variances

```
> sem_res = 'intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+               slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc
+               Q1_Acc ~~ a*Q1_Acc
```

```
+               Q2_Acc ~~ a*Q2_Acc
+               Q3_Acc ~~ a*Q3_Acc'
> sem_res_model = growth(sem_res, missing = "ML", data = cell_wide)
> summary(sem_res_model, fit.measures =TRUE)
```

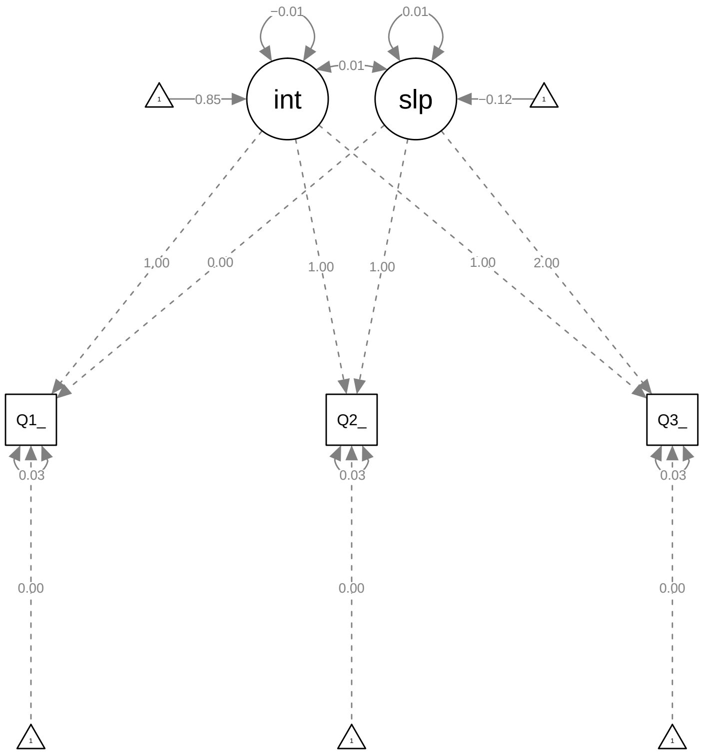
```
lavaan (0.5-23.1097) converged normally after  52 iterations

  Number of observations                            44

  Number of missing patterns                         2

  Estimator                                         ML
  Minimum Function Test Statistic               22.578
  Degrees of freedom                                 3
  P-value (Chi-square)                           0.000

Model test baseline model:

  Minimum Function Test Statistic               18.464
  Degrees of freedom                                 3
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.000
  Tucker-Lewis Index (TLI)                      -0.266

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 23.606
  Loglikelihood unrestricted model (H1)         34.895

  Number of free parameters                          6
  Akaike (AIC)                                 -35.211
  Bayesian (BIC)                               -24.506
  Sample-size adjusted Bayesian (BIC)          -43.308

Root Mean Square Error of Approximation:

  RMSEA                                          0.385
  90 Percent Confidence Interval          0.247  0.541
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.365

Parameter Estimates:

  Information                                 Observed
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
```

```
    Q3_Acc            1.000
  slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope             0.010    0.005    2.207    0.027

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.000
   .Q2_Acc            0.000
   .Q3_Acc            0.000
    intercept         0.846    0.020   42.344    0.000
    slope            -0.121    0.023   -5.294    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc     (a)    0.031    0.007    4.668    0.000
   .Q2_Acc     (a)    0.031    0.007    4.668    0.000
   .Q3_Acc     (a)    0.031    0.007    4.668    0.000
    intercept        -0.008    0.007   -1.251    0.211
    slope             0.007    0.006    1.201    0.230
```

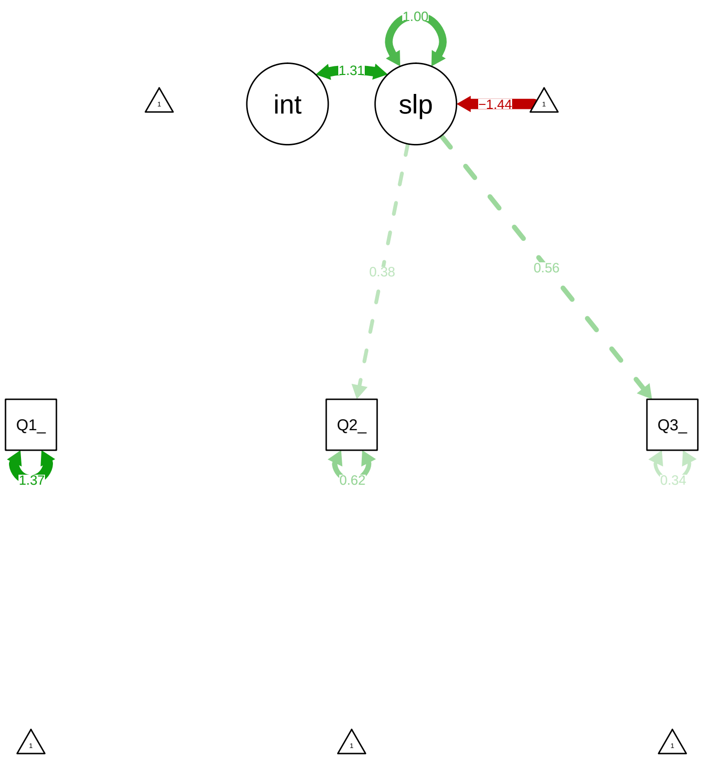
Constraining the residual variances changes the estimate of the slope and intercept, as well as the variance estimates of the slope and intercept. Note also that the variance estimate for the indicator variables is the same now, because we constrained it to be so.

Plotting

```
> semPaths(sem_res_model, whatLabels = "est")
```




```
> semPaths(sem_res_model, what = "std")
```



7.1 Model Comparison

```
> anova(sem_random_model, sem_res_model) ## random model is better
```

```
Chi Square Difference Test

                 Df     AIC     BIC   Chisq Chisq diff
sem_random_model  1 -50.785 -36.512  3.0037
sem_res_model     3 -35.211 -24.506 22.5775     19.574
                 Df diff Pr(>Chisq)
sem_random_model
sem_res_model          2  5.618e-05 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8 Changing the Time Metric

We change centering to last wave instead of first wave. Note that we will use the random slope model as it has the lowest SRMR and RMSEA fit statistics.

```
> sem_time_model = 'intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+                slope =~ -2*Q1_Acc + -1*Q2_Acc + 0*Q3_Acc'
```

```
> sem_time_model = growth(sem_time_model, missing = "ML", data = cell_wide)
> summary(sem_time_model, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  49 iterations

  Number of observations                            44

  Number of missing patterns                         2

  Estimator                                         ML
  Minimum Function Test Statistic                3.004
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.083

Model test baseline model:

  Minimum Function Test Statistic               18.464
  Degrees of freedom                                 3
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.870
  Tucker-Lewis Index (TLI)                       0.611

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 33.393
  Loglikelihood unrestricted model (H1)         34.895

  Number of free parameters                          8
  Akaike (AIC)                                 -50.785
  Bayesian (BIC)                               -36.512
  Sample-size adjusted Bayesian (BIC)          -61.581

Root Mean Square Error of Approximation:

  RMSEA                                          0.213
  90 Percent Confidence Interval          0.000  0.509
  P-value RMSEA <= 0.05                          0.100

Standardized Root Mean Square Residual:

  SRMR                                           0.070

Parameter Estimates:

  Information                                 Observed
  Standard Errors                             Standard

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  slope =~
```

```
    Q1_Acc            -2.000
    Q2_Acc            -1.000
    Q3_Acc             0.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope              0.028    0.012    2.476    0.013

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc             0.000
   .Q2_Acc             0.000
   .Q3_Acc             0.000
    intercept          0.604    0.046   13.064    0.000
    slope             -0.132    0.023   -5.713    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            -0.004    0.007   -0.576    0.565
   .Q2_Acc             0.040    0.010    3.933    0.000
   .Q3_Acc             0.029    0.021    1.376    0.169
    intercept          0.062    0.024    2.643    0.008
    slope              0.017    0.006    2.590    0.010
```

8.1 Model Comparison

```
> anova(sem_random_model, sem_time_model) ## no diference
```

```
Chi Square Difference Test

                 Df     AIC     BIC  Chisq Chisq diff
sem_random_model  1 -50.785 -36.512 3.0037
sem_time_model    1 -50.785 -36.512 3.0037  3.908e-14
                 Df diff Pr(>Chisq)
sem_random_model
sem_time_model         0  < 2.2e-16 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

8.2 Evaluating Fits

8.2.1 Fixed SEM Model

RMSEA: .294, SRMR = .138

8.2.2 Random SEM Model

RMSEA: .213, SRMR = .07

8.2.3 Constrained residual variances SEM Model

RMSEA: .385, SRMR = .365

8.2.4 Changed Time Metric SEM Model

RMSEA: .213, SRMR = .07

Thus, as we notice, changing the time metric for the random slope model does not change the overall fit of the SEM model. However, there is a change in the intercept, since we are now looking at differences in name accuracy in the LAST wave, as opposed to the first wave. Also note that there is NO change in the estimate of the slope variable, i.e. the overall association between Name Accuracy and Time doesn't change by changing the time metric. Also, the slope estimate is negative, which means there's a negative association between naming accuracy and time, which is consistent with the MLM models run previously.

9 Different Estimators

Our final model is the random slopes model. We now use a different estimation technique for this model to see if there are any differences. Since our data is not complete, we will only use estimators that don't require complete data: MLF and MLR.

```
> sem_random = 'intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+               slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc'
> sem_random_mlf = growth(sem_random, data = cell_wide, estimator = "MLF")
> summary(sem_random_mlf, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  91 iterations

                                                  Used       Total
  Number of observations                            43          44

  Estimator                                         ML
  Minimum Function Test Statistic                2.742
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.098

Model test baseline model:

  Minimum Function Test Statistic               18.711
  Degrees of freedom                                 3
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.889
  Tucker-Lewis Index (TLI)                       0.667

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 32.672
  Loglikelihood unrestricted model (H1)         34.043

  Number of free parameters                          8
  Akaike (AIC)                                 -49.343
  Bayesian (BIC)                               -35.253
  Sample-size adjusted Bayesian (BIC)          -60.314

Root Mean Square Error of Approximation:

  RMSEA                                          0.201
  90 Percent Confidence Interval          0.000  0.503
  P-value RMSEA <= 0.05                          0.116
```

```
Standardized Root Mean Square Residual:

  SRMR                                           0.067

Parameter Estimates:

  Information                                 Observed
  Standard Errors                          First.order

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope            -0.005    0.007   -0.756    0.449

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.000
   .Q2_Acc            0.000
   .Q3_Acc            0.000
    intercept         0.865    0.023   38.076    0.000
    slope            -0.129    0.032   -4.016    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc           -0.005    0.010   -0.460    0.646
   .Q2_Acc            0.041    0.017    2.422    0.015
   .Q3_Acc            0.028    0.028    0.991    0.322
    intercept         0.016    0.012    1.324    0.186
    slope             0.017    0.009    1.797    0.072
```

```
> sem_random_mlr = growth(sem_random, data = cell_wide, estimator = "MLR")
> summary(sem_random_mlr, fit.measures = TRUE)
```

```
lavaan (0.5-23.1097) converged normally after  91 iterations

                                                  Used       Total
  Number of observations                            43          44

  Estimator                                         ML      Robust
  Minimum Function Test Statistic                2.742       3.176
  Degrees of freedom                                 1           1
  P-value (Chi-square)                           0.098       0.075
  Scaling correction factor                                  0.863
    for the Yuan-Bentler correction
```

```
Model test baseline model:

  Minimum Function Test Statistic               18.711      18.489
  Degrees of freedom                                 3           3
  P-value                                        0.000       0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.889       0.859
  Tucker-Lewis Index (TLI)                       0.667       0.578

  Robust Comparative Fit Index (CFI)                         0.880
  Robust Tucker-Lewis Index (TLI)                            0.640

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                 32.672      32.672
  Scaling correction factor                                  1.020
    for the MLR correction
  Loglikelihood unrestricted model (H1)         34.043      34.043
  Scaling correction factor                                  1.003
    for the MLR correction

  Number of free parameters                          8           8
  Akaike (AIC)                                 -49.343     -49.343
  Bayesian (BIC)                               -35.253     -35.253
  Sample-size adjusted Bayesian (BIC)          -60.314     -60.314

Root Mean Square Error of Approximation:

  RMSEA                                          0.201       0.225
  90 Percent Confidence Interval          0.000  0.503       0.000  0.000
  P-value RMSEA <= 0.05                          0.116       0.095

  Robust RMSEA                                               0.209
  90 Percent Confidence Interval                             0.000  0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.067       0.067

Parameter Estimates:

  Information                                 Observed
  Standard Errors                   Robust.huber.white

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000
```

```
Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  intercept ~~
    slope            -0.005    0.005   -1.128    0.259

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc            0.000
   .Q2_Acc            0.000
   .Q3_Acc            0.000
    intercept         0.865    0.017   51.187    0.000
    slope            -0.129    0.023   -5.636    0.000

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc           -0.005    0.006   -0.762    0.446
   .Q2_Acc            0.041    0.007    5.741    0.000
   .Q3_Acc            0.028    0.017    1.605    0.108
    intercept         0.016    0.006    2.611    0.009
    slope             0.017    0.006    2.737    0.006
```

Note that the estimates of slope and intercept slightly change if we change the estimator, but the overall fit remains the same.

10 Adding New Latent Variable

```
> sem_two_lv = 'name_intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+               name_slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc
+               time_intercept =~ 1*Q1_td + 1*Q2_td + 1*Q3_td
+               time_slope =~ 0*Q1_td + 1*Q2_td + 2*Q3_td'
> sem_twolv_model = growth(sem_two_lv, data = cell_wide, estimator = "MLF")
> summary(sem_twolv_model, fit.measures = TRUE)
```

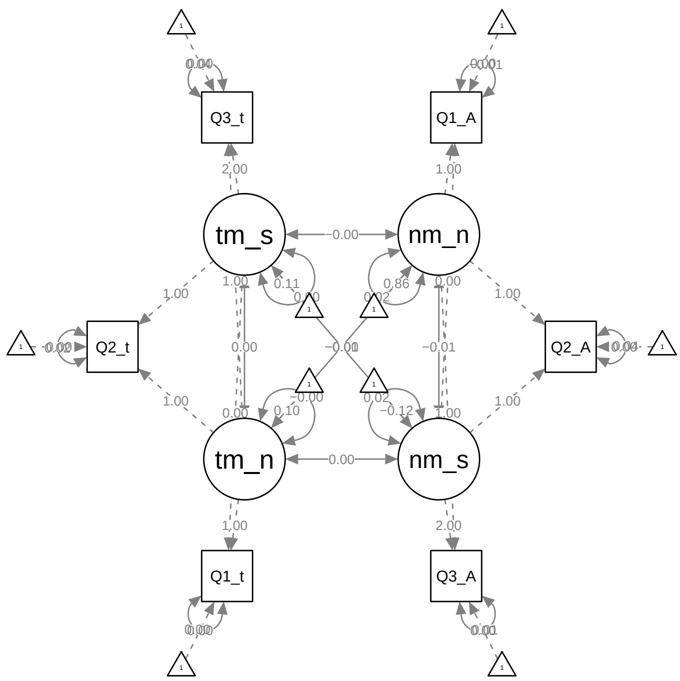
```
lavaan (0.5-23.1097) converged normally after 160 iterations

                                                  Used       Total
  Number of observations                            43          44

  Estimator                                         ML
  Minimum Function Test Statistic                8.663
  Degrees of freedom                                 7
  P-value (Chi-square)                           0.278

Model test baseline model:

  Minimum Function Test Statistic               43.590
  Degrees of freedom                                15
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.942
  Tucker-Lewis Index (TLI)                       0.875

Loglikelihood and Information Criteria:
```

```
  Loglikelihood user model (H0)                141.135
  Loglikelihood unrestricted model (H1)        145.466

  Number of free parameters                         20
  Akaike (AIC)                                -242.270
  Bayesian (BIC)                              -207.046
  Sample-size adjusted Bayesian (BIC)         -269.697

Root Mean Square Error of Approximation:

  RMSEA                                          0.074
  90 Percent Confidence Interval          0.000  0.211
  P-value RMSEA <= 0.05                          0.349

Standardized Root Mean Square Residual:

  SRMR                                           0.080

Parameter Estimates:

  Information                                 Observed
  Standard Errors                          First.order

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  name_intercept =~
    Q1_Acc            1.000
    Q2_Acc            1.000
    Q3_Acc            1.000
  name_slope =~
    Q1_Acc            0.000
    Q2_Acc            1.000
    Q3_Acc            2.000
  time_intercept =~
    Q1_td             1.000
    Q2_td             1.000
    Q3_td             1.000
  time_slope =~
    Q1_td             0.000
    Q2_td             1.000
    Q3_td             2.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  name_intercept ~~
    name_slope       -0.005    0.008   -0.631    0.528
    time_intercept   -0.001    0.002   -0.463    0.644
    time_slope       -0.002    0.003   -0.804    0.422
  name_slope ~~
    time_intercept    0.001    0.002    0.477    0.633
    time_slope       -0.007    0.005   -1.245    0.213
  time_intercept ~~
    time_slope        0.001    0.005    0.298    0.766

Intercepts:
                  Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc           0.000
   .Q2_Acc           0.000
```

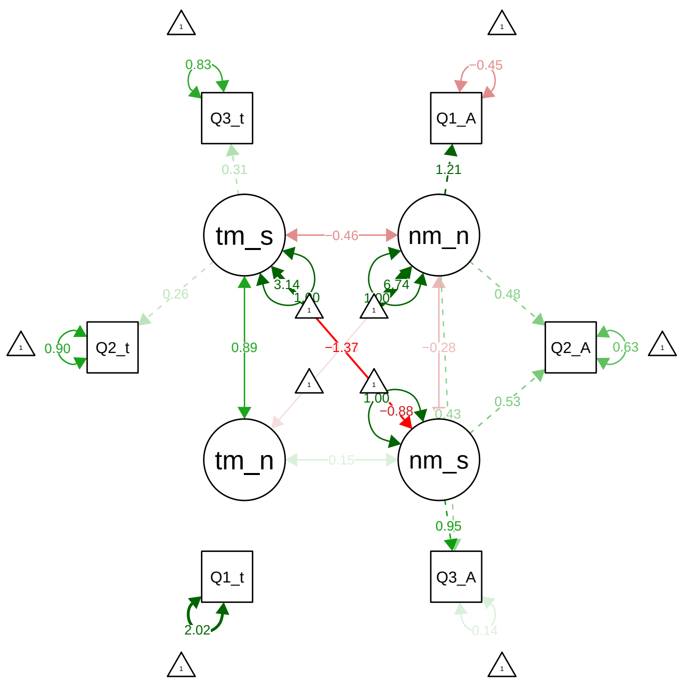
```
   .Q3_Acc           0.000
   .Q1_td            0.000
   .Q2_td            0.000
   .Q3_td            0.000
    name_intercept   0.865    0.026   33.263    0.000
    name_slope      -0.125    0.047   -2.665    0.008
    time_intercept   0.101    0.014    7.389    0.000
    time_slope       0.108    0.028    3.800    0.000

Variances:
                  Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc          -0.005    0.014   -0.371    0.710
   .Q2_Acc           0.045    0.024    1.901    0.057
   .Q3_Acc           0.013    0.038    0.336    0.737
   .Q1_td            0.005    0.006    0.791    0.429
   .Q2_td            0.016    0.006    2.548    0.011
   .Q3_td            0.042    0.015    2.758    0.006
    name_intercept   0.016    0.015    1.134    0.257
    name_slope       0.020    0.014    1.430    0.153
    time_intercept  -0.002    0.006   -0.407    0.684
    time_slope       0.001    0.007    0.166    0.868
```

Plotting

```
> semPaths(sem_twolv_model, layout = "circle2", whatLabels = "est")
```



```
> semPaths(sem_twolv_model, layout = "circle2",  what = "std")
```



11 Adding Another Predictor

```
> sem_newpred = 'name_intercept =~ 1*Q1_Acc + 1*Q2_Acc + 1*Q3_Acc
+               name_slope =~ 0*Q1_Acc + 1*Q2_Acc + 2*Q3_Acc
+               time_intercept =~ 1*Q1_td + 1*Q2_td + 1*Q3_td
+               time_slope =~ 0*Q1_td + 1*Q2_td + 2*Q3_td
+
+                 Q1_Acc ~ Q1_v
+                 Q2_Acc ~ Q2_v
+                 Q3_Acc ~ Q3_v'
> sem_newpred = growth(sem_newpred, data = cell_wide, estimator = "MLF")
> summary(sem_newpred, fit.measures = TRUE)
```

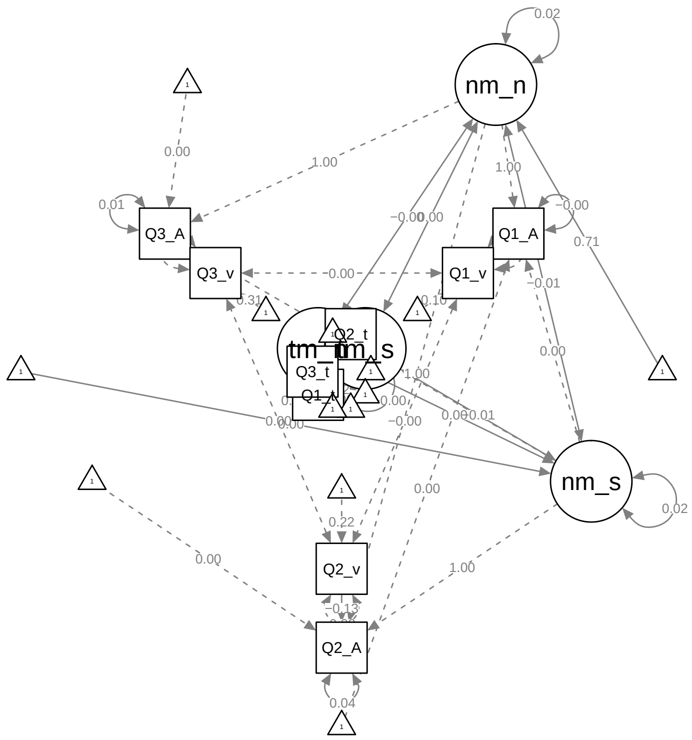
```
lavaan (0.5-23.1097) converged normally after 157 iterations

                                                  Used       Total
  Number of observations                            43          44

  Estimator                                         ML
  Minimum Function Test Statistic              920.951
  Degrees of freedom                                22
  P-value (Chi-square)                           0.000
```

```
Model test baseline model:

  Minimum Function Test Statistic              962.853
  Degrees of freedom                                33
  P-value                                        0.000

User model versus baseline model:

  Comparative Fit Index (CFI)                    0.033
  Tucker-Lewis Index (TLI)                      -0.450

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)                244.531
  Loglikelihood unrestricted model (H1)        705.007

  Number of free parameters                         23
  Akaike (AIC)                                -443.063
  Bayesian (BIC)                              -402.555
  Sample-size adjusted Bayesian (BIC)         -474.605

Root Mean Square Error of Approximation:

  RMSEA                                          0.975
  90 Percent Confidence Interval          0.922  1.029
  P-value RMSEA <= 0.05                          0.000

Standardized Root Mean Square Residual:

  SRMR                                           0.325

Parameter Estimates:

  Information                                 Observed
  Standard Errors                           First.order

Latent Variables:
                    Estimate  Std.Err  z-value  P(>|z|)
  name_intercept =~
    Q1_Acc             1.000
    Q2_Acc             1.000
    Q3_Acc             1.000
  name_slope =~
    Q1_Acc             0.000
    Q2_Acc             1.000
    Q3_Acc             2.000
  time_intercept =~
    Q1_td              1.000
    Q2_td              1.000
    Q3_td              1.000
  time_slope =~
    Q1_td              0.000
    Q2_td              1.000
    Q3_td              2.000

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)
```

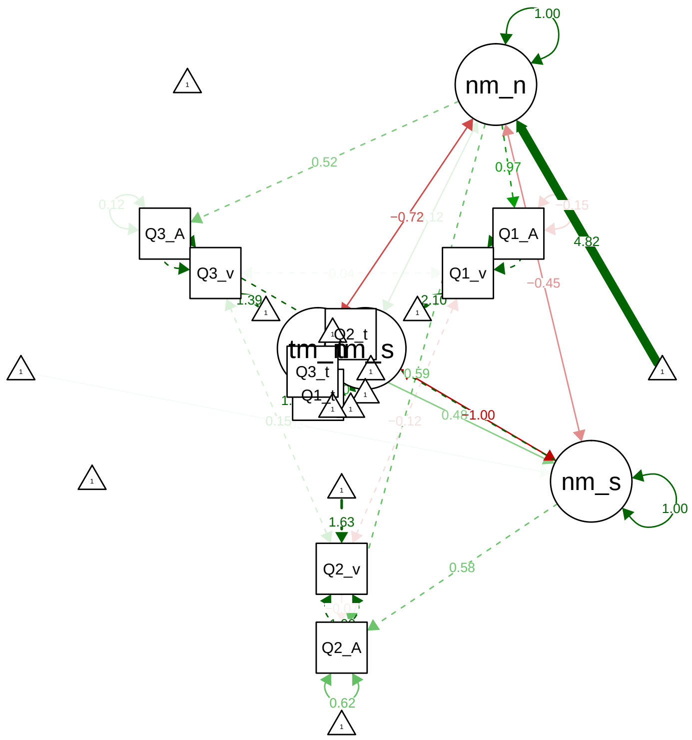
```
  Q1_Acc ~
    Q1_v              1.483    1.406    1.055    0.291
  Q2_Acc ~
    Q2_v             -0.127    0.286   -0.446    0.655
  Q3_Acc ~
    Q3_v             -0.322    0.367   -0.878    0.380

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)
  name_intercept ~~
    name_slope       -0.010    0.009   -1.047    0.295
    time_intercept   -0.004    0.004   -1.083    0.279
    time_slope        0.001    0.004    0.148    0.883
  name_slope ~~
    time_intercept    0.003    0.002    1.366    0.172
    time_slope       -0.006    0.006   -0.871    0.384
  time_intercept ~~
    time_slope        0.001    0.005    0.181    0.856

Intercepts:
                  Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc           0.000
   .Q2_Acc           0.000
   .Q3_Acc           0.000
   .Q1_td            0.000
   .Q2_td            0.000
   .Q3_td            0.000
    name_intercept   0.713    0.154    4.629    0.000
    name_slope       0.004    0.121    0.035    0.972
    time_intercept   0.101    0.015    6.972    0.000
    time_slope       0.108    0.031    3.499    0.000

Variances:
                  Estimate  Std.Err  z-value  P(>|z|)
   .Q1_Acc          -0.003    0.012   -0.280    0.780
   .Q2_Acc           0.038    0.023    1.630    0.103
   .Q3_Acc           0.010    0.037    0.267    0.789
   .Q1_td            0.004    0.006    0.656    0.512
   .Q2_td            0.015    0.007    2.281    0.023
   .Q3_td            0.044    0.022    1.994    0.046
    name_intercept   0.022    0.018    1.224    0.221
    name_slope       0.021    0.012    1.648    0.099
    time_intercept  -0.001    0.006   -0.260    0.795
    time_slope       0.002    0.007    0.229    0.819
```

```
>
```

```
> semPaths(sem_newpred, layout = "circle2", whatLabels = "est")
```




```
> semPaths(sem_newpred, layout = "circle2",  what = "std")
```


nm_n
nm_s
tm_n
tm_s
Q1_A
Q1_v
Q1_t
Q2_A
Q2_v
Q2_t
Q3_A
Q3_v
Q3_t
1.00
0.52
0.97
−0.72
4.82
−0.45
1.39
2.10
0.59
−1.00
1.63
0.58
0.62
1.00