Homework 1

2/17/2022

Question 1

The Counting and Recursive rules can be used to show the model in the figure is identified. The Counting Rule states the number of free parameters $t \leq \frac{1}{2}(s)(s+1)$ where s is the sum of endogenous, p and exogenous q variables. In this example, $p = 4$ and $q = 3$ and therefore $s = 7$ leading to the inequality $t \leq 28 = \frac{1}{2}(7)(8)$. Additionally, t is the number of variances and covariances of the exogenous variables, disturbance terms, and path coefficients or $t = 3 + 7 + 11 = 21$ respectively. Thus we have satisfied the necessary but not sufficient Counting rule because $t = 21 \leq 28$ holds.

The Recursive Rule states B must be triangular and $\Psi = Var(\zeta)$ is a diagonal matrix from the general structural equation of the model, $y = \alpha + B \cdot y + \Gamma \cdot x + \zeta$. In this example B is indeed triangular because the model is restricted in not having feedback loops present amongst the endogenous variables. We are also assuming Ψ is diagonal such that the disturbance terms are uncorrelated with the exogenous variables.

Question 2

We know the multivariate normality assumption is violated because two of the variables, GENDER and IMMIGR, are binary. This means not only are these variable non-normal, but they are also precluded from having joint normality with the other variables.

Question 3

The results from estimating the model and the subsequent interpretation are included below.

```
# Setup path model
pisa_model <- '
  # Regressions

  READING ~ A*MEMO + B*ELAB + C*CSTRAT
  MEMO ~ D*ESCS + E*GENDER
  ELAB ~ F*ESCS + G*GENDER + H*IMMIGR
  CSTRAT ~ I*ESCS + J*GENDER + K*IMMIGR

  # Mediation Analysis

  # Indirect effect of ESCS on READING through MEMO
  DA := D*A
  # Indirect effect of ESCS on READING through ELAB
  FB := F*B
  # Indirect effect of ESCS on READING through CSTRAT
  IC := I*C
  # Total indirect effect of ESCS on READING
  DA_FB_IC := DA + FB + IC

  # Indirect effect of GENDER on READING through MEMO
  EA := E*A
  # Indirect effect of GENDER on READING through ELAB
  GB := G*B
  # Indirect effect of GENDER on READING through CSTRAT
  JC := J*C
  # Total indirect effect of GENDER on READING
  EA_GB_JC := EA + GB + JC

  # Indirect effect of IMMIGR on READING through ELAB
  HB := H*B
  # Indirect effect of IMMIGR on READING through CSTRAT
  KC := K*C
  # Total indirect effect of IMMIGR on READING
  HB_KC := HB + KC
'

# Generate summary of path analysis
fitm1 <- sem(pisa_model, data = pisa_data, estimator = "ml")
```

```
## Warning in lav_data_full(data = data, group = group, cluster = cluster, : lavaan
## WARNING: some observed variances are (at least) a factor 1000 times larger than
## others; use varTable(fit) to investigate
```

```
summary(fitm1, fit.measures = TRUE, rsq = T)
```

```
## lavaan 0.6-10 ended normally after 1 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        15
##
##   Number of observations                          5053
##
## Model Test User Model:
##
##   Test statistic                              5585.354
##   Degrees of freedom                                 7
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              6845.608
##   Degrees of freedom                                18
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.183
##   Tucker-Lewis Index (TLI)                      -1.101
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -52631.946
##   Loglikelihood unrestricted model (H1)     -49839.269
##
##   Akaike (AIC)                              105293.891
##   Bayesian (BIC)                            105391.807
##   Sample-size adjusted Bayesian (BIC)       105344.142
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.397
##   90 Percent confidence interval - lower         0.388
##   90 Percent confidence interval - upper         0.406
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.213
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Regressions:
```

```
##                     Estimate  Std.Err  z-value  P(>|z|)
##   READING ~
##     MEMO       (A)  -26.367    1.113  -23.684    0.000
##     ELAB       (B)  -15.139    1.099  -13.777    0.000
##     CSTRAT     (C)   45.422    1.084   41.893    0.000
##   MEMO ~
##     ESCS       (D)    0.035    0.017    2.083    0.037
##     GENDER     (E)    0.225    0.031    7.339    0.000
##   ELAB ~
##     ESCS       (F)    0.140    0.018    7.908    0.000
##     GENDER     (G)   -0.030    0.031   -0.950    0.342
##     IMMIGR     (H)    0.166    0.041    4.018    0.000
##   CSTRAT ~
##     ESCS       (I)    0.272    0.017   15.575    0.000
##     GENDER     (J)    0.287    0.031    9.339    0.000
##     IMMIGR     (K)    0.238    0.041    5.833    0.000
##
## Variances:
##                     Estimate  Std.Err  z-value  P(>|z|)
##    .READING         7544.331  150.093   50.264    0.000
##    .MEMO               1.191    0.024   50.264    0.000
##    .ELAB               1.221    0.024   50.264    0.000
##    .CSTRAT             1.194    0.024   50.264    0.000
##
## R-Square:
##                     Estimate
##     READING            0.326
##     MEMO               0.011
##     ELAB               0.013
##     CSTRAT             0.061
##
## Defined Parameters:
##                     Estimate  Std.Err  z-value  P(>|z|)
##     DA                -0.916    0.442   -2.075    0.038
##     FB                -2.116    0.309   -6.859    0.000
##     IC                12.366    0.847   14.599    0.000
##     DA_FB_IC           9.333    1.003    9.308    0.000
##     EA                -5.945    0.848   -7.010    0.000
##     GB                 0.447    0.472    0.948    0.343
##     JC                13.045    1.431    9.115    0.000
##     EA_GB_JC           7.547    1.729    4.366    0.000
##     HB                -2.515    0.652   -3.857    0.000
##     KC                10.829    1.874    5.777    0.000
##     HB_KC              8.314    1.984    4.191    0.000
```

```
standardizedSolution(fitm1) %>% filter(op != "~~")
```

```
##          lhs op       rhs    label est.std    se       z pvalue ci.lower ci.upper
## 1   READING  ~      MEMO        A  -0.274 0.011 -24.287  0.000   -0.296   -0.251
## 2   READING  ~      ELAB        B  -0.159 0.011 -13.880  0.000   -0.182   -0.137
## 3   READING  ~    CSTRAT        C   0.484 0.010  47.073  0.000    0.464    0.504
## 4      MEMO  ~      ESCS        D   0.029 0.014   2.084  0.037    0.002    0.057
## 5      MEMO  ~    GENDER        E   0.103 0.014   7.398  0.000    0.075    0.130
## 6      ELAB  ~      ESCS        F   0.116 0.014   7.981  0.000    0.087    0.144
## 7      ELAB  ~    GENDER        G  -0.013 0.014  -0.950  0.342   -0.041    0.014
## 8      ELAB  ~    IMMIGR        H   0.059 0.015   4.028  0.000    0.030    0.087
## 9    CSTRAT  ~      ESCS        I   0.222 0.014  16.118  0.000    0.195    0.249
## 10   CSTRAT  ~    GENDER        J   0.127 0.013   9.452  0.000    0.101    0.154
## 11   CSTRAT  ~    IMMIGR        K   0.083 0.014   5.860  0.000    0.055    0.111
## 12        DA :=       D*A       DA  -0.008 0.004  -2.074  0.038   -0.016    0.000
## 13        FB :=       F*B       FB  -0.018 0.003  -6.879  0.000   -0.024   -0.013
## 14        IC :=       I*C       IC   0.108 0.007  14.984  0.000    0.093    0.122
## 15 DA_FB_IC := DA+FB+IC DA_FB_IC   0.081 0.009   9.429  0.000    0.064    0.098
## 16        EA :=       E*A       EA  -0.028 0.004  -7.034  0.000   -0.036   -0.020
## 17        GB :=       G*B       GB   0.002 0.002   0.948  0.343   -0.002    0.006
## 18        JC :=       J*C       JC   0.062 0.007   9.200  0.000    0.048    0.075
## 19 EA_GB_JC := EA+GB+JC EA_GB_JC   0.036 0.008   4.378  0.000    0.020    0.052
## 20        HB :=       H*B       HB  -0.009 0.002  -3.861  0.000   -0.014   -0.005
## 21        KC :=       K*C       KC   0.040 0.007   5.800  0.000    0.027    0.054
## 22     HB_KC :=     HB+KC    HB_KC   0.031 0.007   4.200  0.000    0.016    0.045
```

In this model, all direct regression coefficients are statistically significant at the $\alpha = 0.05$ level with the exception of the GENDER on the ELAB variables. A similar pattern is present with regard to the indirect effects. All are statistically significant with the exception of the GENDER on READING through ELAB variables. All three total indirect effect parameters are statistically significant.

Question 4

The test statistics generated from estimating the model parameters suggest a poor fitting model. The RMSEA value of 0.397 does not compare well to the generally accepted standard of 0.05 or lower equating to a good fitting model. The yes/no test statistic also suggests a poor fit with a value of 5585.354 and a corresponding χ^2 p-value of approximately zero.

Question 5

The modification indices sorted by decreasing values are included below.

```
modindices(fitm1,standardized=TRUE,power=TRUE,delta=0.1,alpha=.05,high.power=.80) %>%
  filter(op != "~~") %>%
  arrange(desc(mi))
```

```
##        lhs op     rhs       mi     epc sepc.all delta          ncp power
## 1   CSTRAT  ~ READING 4029.389  -0.025   -2.317   0.1 6.612950e+04 1.000
## 2     MEMO  ~  CSTRAT 1985.352   0.624    0.641   0.1 5.096500e+01 1.000
## 3   CSTRAT  ~    MEMO 1966.556   0.624    0.608   0.1 5.043600e+01 1.000
## 4   CSTRAT  ~    ELAB 1767.623   0.585    0.577   0.1 5.168800e+01 1.000
## 5     ELAB  ~  CSTRAT 1767.623   0.598    0.606   0.1 4.939800e+01 1.000
## 6     MEMO  ~    ELAB 1116.603   0.464    0.470   0.1 5.195000e+01 1.000
## 7     ELAB  ~    MEMO 1103.579   0.473    0.467   0.1 4.930600e+01 1.000
## 8     MEMO  ~ READING  819.027   0.008    0.796   0.1 1.202311e+05 1.000
## 9  READING  ~    ESCS  616.031  33.769    0.294   0.1 5.000000e-03 0.051
## 10    ESCS  ~ READING  560.835   0.003    0.397   0.1 4.694443e+05 1.000
## 11    ELAB  ~ READING  253.713   0.004    0.405   0.1 1.401108e+05 1.000
## 12  GENDER  ~ READING   75.507   0.001    0.151   0.1 1.485835e+06 1.000
## 13 READING  ~  GENDER   68.561  20.508    0.097   0.1 2.000000e-03 0.050
## 14 READING  ~  IMMIGR   38.773 -19.343   -0.072   0.1 1.000000e-03 0.050
## 15    MEMO  ~  IMMIGR   19.374   0.180    0.064   0.1 5.997000e+00 0.688
## 16  IMMIGR  ~    MEMO   18.939   0.021    0.058   0.1 4.355650e+02 1.000
## 17    ESCS  ~    MEMO   18.936   0.165    0.197   0.1 6.919000e+00 0.749
## 18  GENDER  ~    MEMO   16.107  -2.869   -6.301   0.1 2.000000e-02 0.052
## 19  IMMIGR  ~ READING    0.295   0.000   -0.008   0.1 2.954823e+06 1.000
## 20  IMMIGR  ~    ELAB    0.000   0.000    0.000   0.1 3.000000e-03 0.050
## 21    ESCS  ~  CSTRAT    0.000   0.000    0.000   0.1 2.000000e-02 0.052
## 22  IMMIGR  ~  CSTRAT    0.000   0.000    0.000   0.1 1.720000e-01 0.070
## 23  GENDER  ~  CSTRAT    0.000   0.000    0.000   0.1 3.900000e-02 0.055
## 24    ESCS  ~  GENDER    0.000   0.000    0.000   0.1 3.000000e-02 0.053
## 25  GENDER  ~    ESCS    0.000   0.000    0.000   0.1 1.380000e-01 0.066
## 26    ESCS  ~  IMMIGR    0.000   0.000    0.000   0.1 7.000000e-03 0.051
## 27  IMMIGR  ~    ESCS    0.000   0.000    0.000   0.1 3.490000e-01 0.091
## 28  GENDER  ~  IMMIGR    0.000   0.000    0.000   0.1 3.060000e-01 0.086
## 29  IMMIGR  ~  GENDER    0.000   0.000    0.000   0.1 2.565157e+30 1.000
## 30    ESCS  ~    ELAB    0.000   0.000    0.000   0.1          NaN   NaN
## 31  GENDER  ~    ELAB    0.000   0.000    0.000   0.1          NaN   NaN
##    decision
## 1    epc:nm
## 2   *epc:m*
## 3   *epc:m*
## 4   *epc:m*
## 5   *epc:m*
## 6   *epc:m*
## 7   *epc:m*
## 8    epc:nm
## 9   **(m)**
## 10   epc:nm
## 11   epc:nm
## 12   epc:nm
## 13  **(m)**
## 14  **(m)**
## 15  **(m)**
## 16   epc:nm
## 17  **(m)**
## 18  **(m)**
## 19     (nm)
```

```
## 20      (i)
## 21      (i)
## 22      (i)
## 23      (i)
## 24      (i)
## 25      (i)
## 26      (i)
## 27      (i)
## 28      (i)
## 29     (nm)
## 30
## 31
```

Based on these results, it appears the model should be altered to relax the restriction of zero effect between ESCS and READING to allow for a direct effect between them. This appears to be the case due to the large values of the modification index, MI, as well as the expected change parameter, EPC. A new model will be estimated with the updated path.

```
# Add READING ~ ESCS path into model based on MI/EPC values
pisa_model2 <- '
  # Regressions

  READING ~ A*MEMO + B*ELAB + C*CSTRAT + L*ESCS
  MEMO ~ D*ESCS + E*GENDER
  ELAB ~ F*ESCS + G*GENDER + H*IMMIGR
  CSTRAT ~ I*ESCS + J*GENDER + K*IMMIGR

  # Mediation Analysis

  # Indirect effect of ESCS on READING through MEMO
  DA := D*A
  # Indirect effect of ESCS on READING through ELAB
  FB := F*B
  # Indirect effect of ESCS on READING through CSTRAT
  IC := I*C
  # Total indirect effect of ESCS on READING
  DA_FB_IC := DA + FB + IC
  # Total effect of ESCS on READING
  DA_FB_IC_L := DA + FB + IC + L

  # Indirect effect of GENDER on READING through MEMO
  EA := E*A
  # Indirect effect of GENDER on READING through ELAB
  GB := G*B
  # Indirect effect of GENDER on READING through CSTRAT
  JC := J*C
  # Total indirect effect of GENDER on READING
  EA_GB_JC := EA + GB + JC

  # Indirect effect of IMMIGR on READING through ELAB
  HB := H*B
  # Indirect effect of IMMIGR on READING through CSTRAT
  KC := K*C
  # Total indirect effect of IMMIGR on READING
  HB_KC := HB + KC
'
fitm2 <- sem(pisa_model2, data = pisa_data, estimator = "ml")
```

```
## Warning in lav_data_full(data = data, group = group, cluster = cluster, : lavaan
## WARNING: some observed variances are (at least) a factor 1000 times larger than
## others; use varTable(fit) to investigate
```

```
summary(fitm2, fit.measures = TRUE, rsq = T)
```

```
## lavaan 0.6-10 ended normally after 1 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        16
##
##   Number of observations                          5053
##
## Model Test User Model:
##
##   Test statistic                              4924.514
##   Degrees of freedom                                 6
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              6845.608
##   Degrees of freedom                                18
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.280
##   Tucker-Lewis Index (TLI)                      -1.161
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -52301.526
##   Loglikelihood unrestricted model (H1)     -49839.269
##
##   Akaike (AIC)                              104635.051
##   Bayesian (BIC)                            104739.495
##   Sample-size adjusted Bayesian (BIC)       104688.652
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.403
##   90 Percent confidence interval - lower         0.393
##   90 Percent confidence interval - upper         0.412
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.196
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Regressions:
```

```
##                      Estimate  Std.Err  z-value  P(>|z|)
##   READING ~
##     MEMO       (A)   -21.963    1.043  -21.055    0.000
##     ELAB       (B)   -15.048    1.034  -14.552    0.000
##     CSTRAT     (C)    37.269    1.035   35.999    0.000
##     ESCS       (L)    33.957    1.274   26.644    0.000
##   MEMO ~
##     ESCS       (D)     0.035    0.017    2.083    0.037
##     GENDER     (E)     0.225    0.031    7.339    0.000
##   ELAB ~
##     ESCS       (F)     0.140    0.018    7.908    0.000
##     GENDER     (G)    -0.030    0.031   -0.950    0.342
##     IMMIGR     (H)     0.166    0.041    4.018    0.000
##   CSTRAT ~
##     ESCS       (I)     0.272    0.017   15.575    0.000
##     GENDER     (J)     0.287    0.031    9.339    0.000
##     IMMIGR     (K)     0.238    0.041    5.833    0.000
##
## Variances:
##                      Estimate  Std.Err  z-value  P(>|z|)
##    .READING          6619.465  131.693   50.264    0.000
##    .MEMO                1.191    0.024   50.264    0.000
##    .ELAB                1.221    0.024   50.264    0.000
##    .CSTRAT              1.194    0.024   50.264    0.000
##
## R-Square:
##                      Estimate
##     READING             0.371
##     MEMO                0.011
##     ELAB                0.013
##     CSTRAT              0.061
##
## Defined Parameters:
##                      Estimate  Std.Err  z-value  P(>|z|)
##     DA                 -0.763    0.368   -2.073    0.038
##     FB                 -2.104    0.303   -6.948    0.000
##     IC                 10.146    0.710   14.294    0.000
##     DA_FB_IC            7.279    0.855    8.516    0.000
##     DA_FB_IC_L         41.236    1.476   27.943    0.000
##     EA                 -4.952    0.715   -6.930    0.000
##     GB                  0.445    0.469    0.948    0.343
##     JC                 10.703    1.184    9.039    0.000
##     EA_GB_JC            6.196    1.460    4.244    0.000
##     HB                 -2.499    0.645   -3.873    0.000
##     KC                  8.885    1.543    5.758    0.000
##     HB_KC               6.386    1.673    3.818    0.000
```

The new model's BIC value decrease by ≈ 652.3 suggesting the new model is preferred over the original one. This new model can be justified when we consider how the ESCS variable is constructed as it includes the number of books in the home. The presence of increased number of books would seem to be a reasonable indicator of a family that places higher importance on reading skills which would lead to a higher measures of that skill regardless of any approaches to teaching.

With this new model in mind, we will repeat the process of identifying potential misspecifications.

```
modindices(fitm2,standardized=TRUE,power=TRUE,delta=0.1,alpha=.05,high.power=.80) %>%
  filter(op != "~~") %>%
  arrange(desc(mi))
```

```
##        lhs op     rhs       mi    epc sepc.all delta         ncp power decision
## 1   CSTRAT  ~ READING 3496.620 -0.029   -2.607   0.1   42593.614 1.000   epc:nm
## 2     MEMO  ~  CSTRAT 1985.352  0.624    0.641   0.1      50.965 1.000 *epc:m*
## 3   CSTRAT  ~    MEMO 1966.556  0.624    0.608   0.1      50.436 1.000 *epc:m*
## 4     ELAB  ~  CSTRAT 1767.623  0.598    0.606   0.1      49.398 1.000 *epc:m*
## 5   CSTRAT  ~    ELAB 1767.623  0.585    0.577   0.1      51.688 1.000 *epc:m*
## 6     MEMO  ~    ELAB 1116.603  0.464    0.470   0.1      51.950 1.000 *epc:m*
## 7     ELAB  ~    MEMO 1103.579  0.473    0.467   0.1      49.306 1.000 *epc:m*
## 8     MEMO  ~ READING  695.513  0.009    0.844   0.1   85194.275 1.000   epc:nm
## 9     ELAB  ~ READING  375.105  0.006    0.585   0.1   93308.358 1.000   epc:nm
## 10 READING  ~  GENDER   97.611 22.942    0.112   0.1       0.002 0.050 **(m)**
## 11  GENDER  ~ READING   70.032  0.001    0.127   0.1 1815175.659 1.000   epc:nm
## 12    MEMO  ~  IMMIGR   19.374  0.180    0.064   0.1       5.997 0.688 **(m)**
## 13  IMMIGR  ~    MEMO   18.937  0.021    0.058   0.1     435.613 1.000   epc:nm
## 14    ESCS  ~    MEMO   18.611  0.163    0.194   0.1       7.040 0.756 **(m)**
## 15  GENDER  ~    MEMO    6.642 -1.183   -2.598   0.1       0.047 0.055 **(m)**
## 16 READING  ~  IMMIGR    1.990  4.315    0.017   0.1       0.001 0.050      (i)
## 17    ESCS  ~ READING    0.182  0.000    0.019   0.1   60925.298 1.000     (nm)
## 18  IMMIGR  ~ READING    0.000  0.000    0.000   0.1 3429120.208 1.000     (nm)
## 19    ESCS  ~    ELAB    0.000  0.000    0.000   0.1       0.094 0.061      (i)
## 20    ESCS  ~  CSTRAT    0.000  0.000    0.000   0.1       0.505 0.110      (i)
## 21  GENDER  ~  CSTRAT    0.000  0.000    0.000   0.1       1.532 0.236      (i)
## 22  GENDER  ~    ESCS    0.000  0.000    0.000   0.1      26.126 0.999     (nm)
## 23  IMMIGR  ~  GENDER    0.000  0.000    0.000   0.1       0.959 0.165      (i)
## 24    ESCS  ~  GENDER    0.000  0.000    0.000   0.1       2.286 0.327      (i)
## 25  GENDER  ~  IMMIGR    0.000  0.000    0.000   0.1       0.719 0.136      (i)
## 26    ESCS  ~  IMMIGR    0.000  0.000    0.000   0.1       1.006 0.171      (i)
## 27  IMMIGR  ~    ESCS    0.000  0.000    0.000   0.1      45.446 1.000     (nm)
## 28  GENDER  ~    ELAB    0.000  0.000    0.000   0.1         NaN   NaN
## 29  IMMIGR  ~    ELAB    0.000  0.000    0.000   0.1         NaN   NaN
## 30  IMMIGR  ~  CSTRAT    0.000  0.000    0.000   0.1         NaN   NaN
```

When considering the modification indices of the second model, it appears another modification should be made in allowing a path from GENDER to READING again based on the large MI and EPC values. This third model will now be estimated.

```
# Add READING ~ GENDER path into model based on MI/EPC values
pisa_model3 <- '
  # Regressions

  READING ~ A*MEMO + B*ELAB + C*CSTRAT + L*ESCS + M*GENDER
  MEMO ~ D*ESCS + E*GENDER
  ELAB ~ F*ESCS + G*GENDER + H*IMMIGR
  CSTRAT ~ I*ESCS + J*GENDER + K*IMMIGR

  # Mediation Analysis

  # Indirect effect of ESCS on READING through MEMO
  DA := D*A
  # Indirect effect of ESCS on READING through ELAB
  FB := F*B
  # Indirect effect of ESCS on READING through CSTRAT
  IC := I*C
  # Total indirect effect of ESCS on READING
  DA_FB_IC := DA + FB + IC
  # Total effect of ESCS on READING
  DA_FB_IC_L := DA + FB + IC + L

  # Indirect effect of GENDER on READING through MEMO
  EA := E*A
  # Indirect effect of GENDER on READING through ELAB
  GB := G*B
  # Indirect effect of GENDER on READING through CSTRAT
  JC := J*C
  # Total indirect effect of GENDER on READING
  EA_GB_JC := EA + GB + JC
  # Total direct effect of GENDER on READING
  EA_GB_JC_M := EA + GB + JC + M

  # Indirect effect of IMMIGR on READING through ELAB
  HB := H*B
  # Indirect effect of IMMIGR on READING through CSTRAT
  KC := K*C
  # Total indirect effect of IMMIGR on READING
  HB_KC := HB + KC
'
fitm3 <- sem(pisa_model3, data = pisa_data, estimator = "ml")
```

```
## Warning in lav_data_full(data = data, group = group, cluster = cluster, : lavaan
## WARNING: some observed variances are (at least) a factor 1000 times larger than
## others; use varTable(fit) to investigate
```

```
summary(fitm3, fit.measures = TRUE, rsq = T)
```

```
## lavaan 0.6-10 ended normally after 1 iterations
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                        17
##
##   Number of observations                          5053
##
## Model Test User Model:
##
##   Test statistic                              4825.567
##   Degrees of freedom                                 5
##   P-value (Chi-square)                           0.000
##
## Model Test Baseline Model:
##
##   Test statistic                              6845.608
##   Degrees of freedom                                18
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    0.294
##   Tucker-Lewis Index (TLI)                      -1.542
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)             -52252.052
##   Loglikelihood unrestricted model (H1)     -49839.269
##
##   Akaike (AIC)                              104538.105
##   Bayesian (BIC)                            104649.076
##   Sample-size adjusted Bayesian (BIC)       104595.056
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.437
##   90 Percent confidence interval - lower         0.427
##   90 Percent confidence interval - upper         0.447
##   P-value RMSEA <= 0.05                          0.000
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.194
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Regressions:
```

```
##                      Estimate  Std.Err  z-value  P(>|z|)
##   READING ~
##     MEMO       (A)  -22.558    1.038  -21.725    0.000
##     ELAB       (B)  -13.545    1.024  -13.226    0.000
##     CSTRAT     (C)   35.402    1.034   34.240    0.000
##     ESCS       (L)   34.449    1.263   27.271    0.000
##     GENDER     (M)   23.030    2.299   10.015    0.000
##   MEMO ~
##     ESCS       (D)    0.035    0.017    2.083    0.037
##     GENDER     (E)    0.225    0.031    7.339    0.000
##   ELAB ~
##     ESCS       (F)    0.140    0.018    7.908    0.000
##     GENDER     (G)   -0.030    0.031   -0.950    0.342
##     IMMIGR     (H)    0.166    0.041    4.018    0.000
##   CSTRAT ~
##     ESCS       (I)    0.272    0.017   15.575    0.000
##     GENDER     (J)    0.287    0.031    9.339    0.000
##     IMMIGR     (K)    0.238    0.041    5.833    0.000
##
## Variances:
##                      Estimate  Std.Err  z-value  P(>|z|)
##    .READING          6491.106  129.139   50.264    0.000
##    .MEMO                1.191    0.024   50.264    0.000
##    .ELAB                1.221    0.024   50.264    0.000
##    .CSTRAT              1.194    0.024   50.264    0.000
##
## R-Square:
##                      Estimate
##     READING             0.376
##     MEMO                0.011
##     ELAB                0.013
##     CSTRAT              0.061
##
## Defined Parameters:
##                      Estimate  Std.Err  z-value  P(>|z|)
##     DA                 -0.784    0.378   -2.073    0.038
##     FB                 -1.894    0.279   -6.787    0.000
##     IC                  9.638    0.680   14.177    0.000
##     DA_FB_IC            6.960    0.826    8.424    0.000
##     DA_FB_IC_L         41.409    1.449   28.571    0.000
##     EA                 -5.087    0.732   -6.953    0.000
##     GB                  0.400    0.422    0.947    0.343
##     JC                 10.167    1.128    9.010    0.000
##     EA_GB_JC            5.481    1.410    3.888    0.000
##     EA_GB_JC_M         28.510    2.643   10.787    0.000
##     HB                 -2.250    0.585   -3.845    0.000
##     KC                  8.440    1.468    5.750    0.000
##     HB_KC               6.190    1.580    3.918    0.000
```

The third model's BIC value decrease by ≈ 90.4 suggesting this is preferred over the second one. This third model can be justified when considering how GENDER and READING are related. Possible explanations could include both societal reasons, such as boys being placed in sporting activities at the expense of learning activities, and biological reasons such as vision differences in sex due to crucial genes being on the Y-chromosome. Importantly, these potential justifications would directly affect the reading measure.

With this third model in mind, we will again repeat the process of identifying potential misspecifications.

```
modindices(fitm3,standardized=TRUE,power=TRUE,delta=0.1,alpha=.05,high.power=.80) %>%
  filter(op != "~~") %>%
  arrange(desc(mi))
```

```
##        lhs op     rhs       mi    epc sepc.all delta         ncp power decision
## 1   CSTRAT  ~ READING 3428.751 -0.030   -2.755   0.1   36980.818 1.000   epc:nm
## 2     MEMO  ~  CSTRAT 1985.352  0.624    0.641   0.1      50.965 1.000 *epc:m*
## 3   CSTRAT  ~    MEMO 1966.556  0.624    0.608   0.1      50.436 1.000 *epc:m*
## 4   CSTRAT  ~    ELAB 1767.623  0.585    0.577   0.1      51.688 1.000 *epc:m*
## 5     ELAB  ~  CSTRAT 1767.623  0.598    0.606   0.1      49.398 1.000 *epc:m*
## 6     MEMO  ~    ELAB 1116.603  0.464    0.470   0.1      51.950 1.000 *epc:m*
## 7     ELAB  ~    MEMO 1103.579  0.473    0.467   0.1      49.306 1.000 *epc:m*
## 8     MEMO  ~ READING  874.335  0.011    1.016   0.1   73175.570 1.000   epc:nm
## 9     ELAB  ~ READING  302.697  0.006    0.538   0.1   87849.420 1.000   epc:nm
## 10    MEMO  ~  IMMIGR   19.374  0.180    0.064   0.1       5.997 0.688 **(m)**
## 11  IMMIGR  ~    MEMO   18.927  0.021    0.058   0.1     435.845 1.000   epc:nm
## 12    ESCS  ~    MEMO   18.919  0.165    0.197   0.1       6.925 0.749 **(m)**
## 13  GENDER  ~    MEMO   16.091 -2.866   -6.295   0.1       0.020 0.052 **(m)**
## 14 READING  ~  IMMIGR    2.086  4.375    0.017   0.1       0.001 0.050      (i)
## 15  IMMIGR  ~ READING    0.006  0.000    0.001   0.1 3525544.289 1.000     (nm)
## 16    ESCS  ~ READING    0.005  0.000    0.003   0.1   62325.720 1.000     (nm)
## 17  GENDER  ~ READING    0.002  0.000   -0.040   0.1     440.217 1.000     (nm)
## 18  GENDER  ~    ELAB    0.000  0.000    0.000   0.1       0.029 0.053      (i)
## 19  GENDER  ~  CSTRAT    0.000  0.000    0.000   0.1       0.079 0.059      (i)
## 20    ESCS  ~    ELAB    0.000  0.000    0.000   0.1       0.094 0.061      (i)
## 21    ESCS  ~  CSTRAT    0.000  0.000    0.000   0.1       0.342 0.090      (i)
## 22  IMMIGR  ~    ELAB    0.000  0.000    0.000   0.1       0.606 0.122      (i)
## 23  GENDER  ~    ESCS    0.000  0.000    0.000   0.1       0.140 0.066      (i)
## 24  IMMIGR  ~  CSTRAT    0.000  0.000    0.000   0.1       4.295 0.545      (i)
## 25  GENDER  ~  IMMIGR    0.000  0.000    0.000   0.1       0.886 0.156      (i)
## 26  IMMIGR  ~  GENDER    0.000  0.000    0.000   0.1       2.635 0.368      (i)
## 27    ESCS  ~  GENDER    0.000  0.000    0.000   0.1       0.066 0.058      (i)
## 28    ESCS  ~  IMMIGR    0.000  0.000    0.000   0.1       1.058 0.177      (i)
## 29  IMMIGR  ~    ESCS    0.000  0.000    0.000   0.1      47.120 1.000     (nm)
```

These results suggest there are no more major model misspecifications.

Question 6

The final model chosen based on BIC values is shown below.