**Structural Equation Modeling with Latent Variables or their and Plausible Values using Mplus**

These data were adapted from my dissertation work in which 152 adults age 63–87 years were measured on visual impairment (distance acuity and five degrees of contrast sensitivity), processing speed, divided visual attention, and selective visual attention (as measured by the Useful Field of View subtests for each), attentional search efficiency (DriverScan), and simulator driving impairment (as measured by six driving performance indicators).

**Hoffman, L.**, Yang, X., Bovaird, J. A., & Embretson, S. E. (2006). Measuring attention in older adults: Development and psychometric evaluation of DriverScan. *Educational and Psychological Measurement*, *66*, 984-1000.

**Hoffman, L.**, McDowd, J. M., Atchley, P., & Dubinsky R. A. (2005). The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging, 20*(4), 610-622.

This example will demonstrate how to estimate structural equation models, including models with latent variable interactions. But because simultaneous estimation of all effects of interest may not always be possible, this example will also how to generate, merge, and use plausible values instead (models with a “b” subscript). Finally, the example will also demonstrate the optimism of the results when using only a single factor score per person.

### Mplus Code to Read in Data:

```
TITLE:      SEM Example for Driverscan
DATA:       FILE = driverscanSEM.csv;     ! FILE is file to be analyzed
            FORMAT = free;                ! Free is default
            TYPE = INDIVIDUAL;            ! Individual data is default

VARIABLE:  ! Every variable in data set
      NAMES = PartID sex age75 lncs15 lncs3 lncs6 lncs12 lncs18 far lnps
              lnda lnsa Dscan lane da_task crash stop speed time;
      ! Every variable in EACH MODEL
      USEVARIABLES = (to be changed for each model);
      IDVARIABLE = PartID;                ! Will keep ID variable for merging
      MISSING ARE ALL (-9999);            ! Make sure to specify all missing values

ANALYSIS:   ESTIMATOR IS MLR; ! For continuous items whose residuals may not be normal

OUTPUT:     SAMPSTAT                ! Sample descriptives to verify data
            MODINDICES (3.84)       ! Voodoo to improve model (at p<.05)
            STDYX                   ! Requests fully standardized solution
            RESIDUAL                ! Requests standardized and normalized residuals
            SVALUES;                ! Write code with estimated parameters
            TECH4;                  ! Latent variable correlation matrix

SAVEDATA:   SAVE = FSCORES; FILE = FactorScores.dat;  ! Change .dat name by model
            MISSFLAG = 99;                            ! Missing data indicator

MODEL:      (model syntax goes here, to be changed for each model)
```

We will begin by fitting single-factor measurement models for each latent factor. This is for 2 reasons: (1) we need to ensure each factor fits *per se*, and (2) we will generate the plausible values to use later. If you are doing full SEM, you only need the “a” versions of each measurement. The “b” versions are only needed for making plausible values.

**Measurement Model for Visual Impairment (including Omega)**

***Model 1a: Estimate model using MLR (also generate SVALUES text for making plausible values)***

```
VARIABLE:  ! Every variable in THIS MODEL
            USEVARIABLES = lncs15 lncs3 lncs6 lncs12 lncs18 far;
ANALYSIS:   ESTIMATOR = MLR;
MODEL:      ! Measurement models
  Vision BY far@1
        lncs15* lncs3* lncs6* lncs12* lncs18* (L2-L6);          ! 1 marker loading
  [far* lncs15* lncs3* lncs6* lncs12* lncs18*];                  ! All intercepts
   far* lncs15* lncs3* lncs6* lncs12* lncs18* (E1-E6);          ! Residual variances
  [Vision@0]; Vision* (Fvar);                                    ! Factor M=0, Var=?

MODEL CONSTRAINT: ! TO GET OMEGA
NEW(SumLoad2 SumError SumRCov Omega);
SumLoad2 =( 1+L2+L3+L4+L5+L6)**2;
SumError = E1+E2+E3+E4+E5+E6;
SumRCov = 2*(0);
! Omega = true variance / total variance
Omega = SumLoad2*Fvar / (SumLoad2*Fvar+SumError+SumRCov);
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       18

Loglikelihood
          H0 Value                        -747.948
          H0 Scaling Correction Factor      1.1255
            for MLR
          H1 Value                        -739.282
          H1 Scaling Correction Factor      1.1171
            for MLR

Information Criteria
          Akaike (AIC)                    1531.897
          Bayesian (BIC)                  1586.327
          Sample-Size Adjusted BIC        1529.357
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit
          Value                             15.752*
          Degrees of Freedom                     9
          P-Value                           0.0722
          Scaling Correction Factor         1.1003
            for MLR

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.070
          90 Percent C.I.                    0.000  0.126
          Probability RMSEA <= .05           0.246

CFI/TLI
          CFI                                0.973
          TLI                                0.955

Chi-Square Test of Model Fit for the Baseline Model
          Value                            264.950
          Degrees of Freedom                    15
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)
          Value                              0.041
```

## Measurement Model for Vision:

MODEL RESULTS

| | Estimate | S.E. | Est./S.E. | Two-Tailed P-Value |
|---|---|---|---|---|
| VISION BY | | | | |
| FAR | 1.000 | 0.000 | 999.000 | 999.000 |
| LNCS15 | 0.497 | 0.103 | 4.815 | 0.000 |
| LNCS3 | 0.594 | 0.118 | 5.018 | 0.000 |
| LNCS6 | 0.764 | 0.136 | 5.628 | 0.000 |
| LNCS12 | 1.296 | 0.207 | 6.277 | 0.000 |
| LNCS18 | 1.504 | 0.237 | 6.353 | 0.000 |
| Means | | | | |
| VISION | 0.000 | 0.000 | 999.000 | 999.000 |
| Intercepts | | | | |
| LNCS15 | -3.698 | 0.035 | -105.136 | 0.000 |
| LNCS3 | -3.938 | 0.035 | -113.273 | 0.000 |
| LNCS6 | -3.730 | 0.043 | -87.639 | 0.000 |
| LNCS12 | -2.368 | 0.066 | -36.000 | 0.000 |
| LNCS18 | -1.406 | 0.081 | -17.389 | 0.000 |
| FAR | 3.026 | 0.067 | 45.130 | 0.000 |
| Variances | | | | |
| VISION | 0.224 | 0.067 | 3.333 | 0.001 |
| Residual Variances | | | | |
| LNCS15 | 0.133 | 0.018 | 7.435 | 0.000 |
| LNCS3 | 0.105 | 0.014 | 7.451 | 0.000 |
| LNCS6 | 0.145 | 0.028 | 5.231 | 0.000 |
| LNCS12 | 0.282 | 0.047 | 5.947 | 0.000 |
| LNCS18 | 0.488 | 0.062 | 7.933 | 0.000 |
| FAR | 0.460 | 0.055 | 8.349 | 0.000 |
| New/Additional Parameters | | | | |
| SUMLOAD2 | 31.983 | 7.564 | 4.228 | 0.000 |
| SUMERROR | 1.613 | 0.102 | 15.822 | 0.000 |
| SUMRCOV | 0.000 | 0.000 | 0.000 | 1.000 |
| OMEGA | 0.816 | 0.024 | 33.851 | 0.000 |

STANDARDIZED MODEL RESULTS
STDYX Standardization

| | Estimate | S.E. | Est./S.E. | Two-Tailed P-Value |
|---|---|---|---|---|
| VISION BY | | | | |
| FAR | 0.572 | 0.062 | 9.190 | 0.000 |
| LNCS15 | 0.541 | 0.074 | 7.305 | 0.000 |
| LNCS3 | 0.656 | 0.062 | 10.605 | 0.000 |
| LNCS6 | 0.688 | 0.057 | 12.062 | 0.000 |
| LNCS12 | 0.756 | 0.051 | 14.815 | 0.000 |
| LNCS18 | 0.713 | 0.041 | 17.293 | 0.000 |

**The SVALUES output option provided this code, which we will use to generate plausible values for our factor scores for later use.**

```
MODEL COMMAND WITH FINAL
ESTIMATES USED AS STARTING
VALUES

vision BY far@1;
vision BY lncs15*0.49665 (l2);
vision BY lncs3*0.59433 (l3);
vision BY lncs6*0.76361 (l4);
vision BY lncs12*1.29636 (l5);
vision BY lncs18*1.50436 (l6);

 [ lncs15*-3.69842 ];
 [ lncs3*-3.93821 ];
 [ lncs6*-3.72997 ];
 [ lncs12*-2.36777 ];
 [ lncs18*-1.40608 ];
 [ far*3.02632 ];
 [ vision@0 ];

 lncs15*0.13297 (e2);
 lncs3*0.10479 (e3);
 lncs6*0.14501 (e4);
 lncs12*0.28191 (e5);
 lncs18*0.48808 (e6);
 far*0.46001 (e1);
 vision*0.22350;
```

**For factor score reliability**

SAMPLE STATISTICS FOR ESTIMATED FACTOR SCORES

Means

| VISION | VISION_SE |
|---|---|
| 0.000 | **0.194** |

Covariances

| | VISION |
|---|---|
| VISION | 0.186 |

$$\rho = \frac{.224}{.224 + .194^2} = .856$$

**Factor score reliability uses the factor variance, but reliability corrections will use the factor score variance instead.**

Normalized Residuals for Covariances/Correlations/Residual Correlations

| | LNCS15 | LNCS3 | LNCS6 | LNCS12 | LNCS18 |
|---|---|---|---|---|---|
| LNCS15 | 0.000 | | | | |
| LNCS3 | 1.651 | 0.000 | | | |
| LNCS6 | -0.045 | 0.261 | 0.000 | | |
| LNCS12 | -0.455 | -0.241 | 0.021 | 0.000 | |
| LNCS18 | -0.629 | -0.458 | -0.177 | 0.353 | 0.000 |
| FAR | -0.471 | -0.731 | -0.062 | 0.198 | 0.558 |

**Local fit looks good as well…**

**Now we are ready for Model 1b for the Visual Impairment factor:**
***Generate plausible values using BAYES estimation of *previous MLR model****

```
ANALYSIS:      ESTIMATOR = BAYES;
MODEL:      ! Parameters from previously estimated measurement model, but all FIXED@

Vision BY far@1;
Vision BY lncs15@0.49665(l2);
Vision BY lncs3@0.59434(l3);
Vision BY lncs6@0.76361(l4);
Vision BY lncs12@1.29636(l5);
Vision BY lncs18@1.50436(l6);

[lncs15@-3.69842];
[lncs3@-3.93821];
[lncs6@-3.72997];
[lncs12@-2.36777];
[lncs18@-1.40608];
[far@3.02632];
[vision@0];

lncs15@0.13296(e2);
lncs3@0.10479(e3);
lncs6@0.14501(e4);
lncs12@0.28191(e5);
lncs18@0.48808(e6);
far@0.46002(e1);
vision@0.22350;

DATA IMPUTATION:                   ! Creating plausible values for factor score
    NDATASETS = 100;               ! Number of separate values to create
    SAVE = PV/Vision*.dat;         ! Name of separate datasets with plausible values

SAVEDATA:   FILE = PV/VisionSummary.dat;  ! Summary about plausibles per person
            SAVE = FSCORES (100);         ! Needed to generate 100 factor scores
            FACTORS = Vision;             ! Which factors to save
            MISSFLAG = 99;                ! Missing data indicator for items

  Save file
    Vision*.dat

  Order of variables

    LNCS15
    LNCS3
    LNCS6
    LNCS12
    LNCS18
    FAR
    PARTID
    VISION
```

Now, within a subfolder of PV/, we have 100 datasets (named Vision1.dat to Vision100.dat) with these variables in this order. Thus, rather than just using the mean of a person's factor score distribution, we are *sampling* from each person's factor distribution.

It also made a text file called "Visionlist.dat" that lists these individual data files:

```
Vision1.dat
Vision2.dat
Vision3.dat
Vision4.dat
Vision5.dat
…
```

Later we will use these types of files to tell Mplus to run the same analysis on every single file, then aggregate the results (as in multiple imputation for missing data).

## Measurement Model for Driving Impairment (including Omega)

***Model 2a: Estimate model using MLR (also generate SVALUES text for making plausible values)***

```
VARIABLE:  ! Every variable in THIS MODEL
            USEVARIABLES = lane da_task crash stop speed time;
ANALYSIS:   ESTIMATOR = MLR;
MODEL:      ! Measurement models
  Driving BY crash@1
              da_task* lane* stop* speed* time* (L2-L6);   ! 1 marker loading
  [lane* da_task* crash* stop* speed* time*];              ! All intercepts
   lane* da_task* crash* stop* speed* time* (E1-E6);       ! Residual variances
  [Driving@0]; Driving* (Fvar);                            ! Factor M=0, Var=?
  speed WITH time* (ResCov);                               ! Residual covariance

MODEL CONSTRAINT: ! TO GET OMEGA
NEW(SumLoad2 SumError SumRCov Omega);
SumLoad2 = ( 1+L2+L3+L4+L5+L6)**2;
SumError =  E1+E2+E3+E4+E5+E6;
SumRCov = 2*(ResCov);
! Omega = true variance / total variance
Omega = SumLoad2*Fvar / (SumLoad2*Fvar+SumError+SumRCov);
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       19
Loglikelihood
          H0 Value                         -37.119
          H0 Scaling Correction Factor      1.1566
            for MLR
          H1 Value                         -30.710
          H1 Scaling Correction Factor      1.1108
            for MLR

Information Criteria
          Akaike (AIC)                     112.239
          Bayesian (BIC)                   167.012
          Sample-Size Adjusted BIC         106.915
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit
          Value                             12.791*
          Degrees of Freedom                     8
          P-Value                           0.1192
          Scaling Correction Factor         1.0021
            for MLR

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.067
          90 Percent C.I.                    0.000  0.133
          Probability RMSEA <= .05           0.293

CFI/TLI
          CFI                                0.922
          TLI                                0.854

Chi-Square Test of Model Fit for the Baseline Model
          Value                             76.677
          Degrees of Freedom                    15
          P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)
          Value                              0.054
```

## Measurement Model for Driving:

```
MODEL RESULTS
                                                           Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
DRIVING  BY
    CRASH              1.000      0.000    999.000    999.000
    LANE               0.150      0.057      2.608      0.009
    DA_TASK            0.173      0.074      2.348      0.019
    STOP               0.347      0.163      2.124      0.034
    SPEED              0.422      0.138      3.054      0.002
    TIME               0.048      0.043      1.104      0.270

 SPEED    WITH
    TIME              -0.023      0.004     -5.393      0.000

 Means
    DRIVING            0.000      0.000    999.000    999.000

 Intercepts
    LANE               0.815      0.015     53.293      0.000
    DA_TASK            0.256      0.013     20.102      0.000
    CRASH              0.859      0.053     16.292      0.000
    STOP               0.205      0.038      5.349      0.000
    SPEED              0.836      0.042     19.687      0.000
    TIME               3.146      0.009    349.081      0.000

 Variances
    DRIVING            0.159      0.062      2.574      0.010

 Residual Variances
    LANE               0.027      0.004      6.596      0.000
    DA_TASK            0.017      0.004      4.613      0.000
    CRASH              0.209      0.055      3.781      0.000
    STOP               0.174      0.031      5.575      0.000
    SPEED              0.210      0.028      7.391      0.000
    TIME               0.010      0.001      8.639      0.000

 New/Additional Parameters
    SUMLOAD2           4.578      1.185      3.865      0.000
    SUMERROR           0.647      0.067      9.627      0.000
    SUMRCOV           -0.046      0.009     -5.393      0.000
    OMEGA              0.548      0.076      7.166      0.000
```

**The SVALUES output option provided this code, which we will use to generate plausible values for our factor scores for later use.**

```
MODEL COMMAND WITH FINAL
ESTIMATES USED AS STARTING VALUES

 driving BY crash@1;
 driving BY lane*0.14977 (l2);
 driving BY da_task*0.17282 (l3);
 driving BY stop*0.34713 (l4);
 driving BY speed*0.42198 (l5);
 driving BY time*0.04799 (l6);

 speed WITH time*-0.02305
(rescov);

 [ lane*0.81538 ];
 [ da_task*0.25614 ];
 [ crash*0.85947 ];
 [ stop*0.20455 ];
 [ speed*0.83636 ];
 [ time*3.14598 ];
 [ driving*0 ];

 lane*0.02734 (e1);
 da_task*0.01669 (e2);
 crash*0.20856 (e3);
 stop*0.17387 (e4);
 speed*0.20994 (e5);
 time*0.01036 (e6);
 driving*0.15881;
```

```
STANDARDIZED MODEL RESULTS
STDYX Standardization
                                                           Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
DRIVING  BY
    CRASH              0.657      0.117      5.596      0.000
    LANE               0.340      0.123      2.767      0.006
    DA_TASK            0.470      0.132      3.576      0.000
    STOP               0.315      0.115      2.748      0.006
    SPEED              0.345      0.107      3.226      0.001
    TIME               0.185      0.145      1.275      0.202

SPEED    WITH
    TIME              -0.494      0.090     -5.478      0.000
```

**For factor score reliability**

```
SAMPLE STATISTICS FOR ESTIMATED
FACTOR SCORES
          Means
         DRIVING      DRIVING_SE
        ________      ________
          0.000          0.247
        Covariances
          ________      ________
          DRIVING         0.098
```

$\rho = \frac{.159}{.159 + .247^2} = .723$ Uh-oh…

**Factor score reliability uses the factor variance, but reliability corrections will use the factor score variance instead.**

```
           Normalized Residuals for Covariances/Correlations/Residual Correlations
              LANE          DA_TASK       CRASH         STOP          SPEED
              ________      ________      ________      ________      ________
 LANE           0.000
 DA_TASK       -0.487         0.000
 CRASH          0.359        -0.390         0.000
 STOP           0.769         0.503        -0.004         0.000
 SPEED          0.458        -0.836         0.471        -0.482         0.000
 TIME          -1.508         2.067        -0.346        -0.545         0.000
```

**Local fit look mostly ok…**

**Now we are ready for Model 2b for the Driving Impairment factor:**
***Generate plausible values using BAYES estimation of *previous MLR model****

```
ANALYSIS:      ESTIMATOR = BAYES;
MODEL:      ! Parameters from previously estimated measurement model, but all FIXED

 driving BY crash@1;
 driving BY lane@0.14977 (l2);
 driving BY da_task@0.17282 (l3);
 driving BY stop@0.34713 (l4);
 driving BY speed@0.42198 (l5);
 driving BY time@0.04799 (l6);

 speed WITH time@-0.02305 (rescov);

 [ lane@0.81538 ];
 [ da_task@0.25614 ];
 [ crash@0.85947 ];
 [ stop@0.20455 ];
 [ speed@0.83636 ];
 [ time@3.14598 ];
 [ driving@0 ];

 lane@0.02734 (e1);
 da_task@0.01669 (e2);
 crash@0.20856 (e3);
 stop@0.17387 (e4);
 speed@0.20994 (e5);
 time@0.01036 (e6);
 driving@0.15881;


DATA IMPUTATION:                   ! Creating plausible values for factor score
    NDATASETS = 100;               ! Number of separate values to create
    SAVE = PV/Driving*.dat;        ! Name of separate datasets with plausible values

SAVEDATA:   FILE = PV/DrivingSummary.dat;  ! Summary about plausibles per person
            SAVE = FSCORES (100);          ! Needed to generate 100 factor scores
            FACTORS = Driving;             ! Which factors to save
            MISSFLAG = 99;                 ! Missing data indicator for items
```

```
  Save file
    Driving*.dat

  Order of variables

    LANE
    DA_TASK
    CRASH
    STOP
    SPEED
    TIME
    PARTID
    DRIVING
```

Now, within a subfolder of PV/, we have 100 datasets (named Driving1.dat to Driving100.dat) with these variables in this order. Thus, rather than just using the mean, we are *sampling* from each person's factor score distribution. It also made a text file called "Drivinglist.dat" that lists these individual data files:

```
Driving1.dat
Driving2.dat
Driving3.dat
Driving4.dat
Driving5.dat
…..
```

Later we will use these types of files to tell Mplus to run the same analysis on every single file, then aggregate the results (as in multiple imputation).

## Measurement Model for Attentional Impairment (including Omega)

***Model 3a: Estimate model using MLR (also generate SVALUES text for making plausible values)***

```
VARIABLE:  ! Every variable in THIS MODEL
            USEVARIABLES = lnda lnsa Dscan;
ANALYSIS:   ESTIMATOR = MLR;

MODEL:       ! Measurement models
  Attn BY lnda@1
          lnsa* dscan* (L2-L3); ! 1 marker loading
  [lnda* lnsa* dscan*];         ! All intercepts
   lnda* lnsa* dscan* (E1-E3);  ! Residual variances
  [Attn@0]; Attn* (Fvar);       ! Factor M=0, Var=?

MODEL CONSTRAINT:  ! TO GET OMEGA
NEW(SumLoad2 SumError SumRCov Omega);
SumLoad2 = ( 1+L2+L3)**2;
SumError =  E1+E2+E3;
SumRCov = 2*(0);
! Omega = true variance / total variance
Omega = SumLoad2*Fvar / (SumLoad2*Fvar+SumError+SumRCov);
```

**The SVALUES output option provided this code, which we will use to generate plausible values for our factor scores for later use.**

```
MODEL COMMAND WITH FINAL
ESTIMATES USED AS STARTING
VALUES

 attn BY lnda@1;
 attn BY lnsa*0.51567 (l2);
 attn BY dscan*1.10655 (l3);

 [ lnda*4.35396 ];
 [ lnsa*5.58076 ];
 [ dscan*-0.01244 ];
 [ attn*0 ];

 lnda*0.51556 (e1);
 lnsa*0.08112 (e2);
 dscan*0.44855 (e3);
 attn*0.44310;
```

***Can you guess why I didn't include the model fit?***

### Measurement Model for Attention:

MODEL RESULTS

| | Estimate | S.E. | Est./S.E. | Two-Tailed P-Value |
|---|---|---|---|---|
| ATTN BY | | | | |
| LNDA | 1.000 | 0.000 | 999.000 | 999.000 |
| LNSA | 0.516 | 0.071 | 7.275 | 0.000 |
| DSCAN | 1.107 | 0.139 | 7.933 | 0.000 |
| Means | | | | |
| ATTN | 0.000 | 0.000 | 999.000 | 999.000 |
| Intercepts | | | | |
| LNDA | 4.354 | 0.079 | 54.825 | 0.000 |
| LNSA | 5.581 | 0.036 | 154.256 | 0.000 |
| DSCAN | -0.012 | 0.081 | -0.154 | 0.878 |
| Variances | | | | |
| ATTN | 0.443 | 0.088 | 5.008 | 0.000 |
| Residual Variances | | | | |
| LNDA | 0.516 | 0.068 | 7.597 | 0.000 |
| LNSA | 0.081 | 0.017 | 4.674 | 0.000 |
| DSCAN | 0.449 | 0.086 | 5.243 | 0.000 |
| New/Additional Parameters | | | | |
| SUMLOAD2 | 6.876 | 0.960 | 7.165 | 0.000 |
| SUMERROR | 1.045 | 0.102 | 10.212 | 0.000 |
| SUMRCOV | 0.000 | 0.000 | 0.000 | 1.000 |
| OMEGA | 0.745 | 0.038 | 19.728 | 0.000 |

STANDARDIZED MODEL RESULTS
STDYX Standardization

| | Estimate | S.E. | Est./S.E. | Two-Tailed P-Value |
|---|---|---|---|---|
| ATTN BY | | | | |
| LNDA | 0.680 | 0.055 | 12.275 | 0.000 |
| LNSA | 0.770 | 0.055 | 14.087 | 0.000 |
| DSCAN | 0.740 | 0.056 | 13.153 | 0.000 |

**For factor score reliability**

SAMPLE STATISTICS FOR ESTIMATED FACTOR SCORES

Means

| ATTN | ATTN_SE |
|---|---|
| 0.000 | 0.313 |

Covariances

| | ATTN |
|---|---|
| ATTN | 0.345 |

$$\rho = \frac{.443}{.443 + .313^2} = .819$$

**Factor score reliability uses the factor variance, but reliability corrections will use the factor score variance instead.**

**<u>Now we are ready for Model 3b for the Attentional Impairment factor:</u>**
***Generate plausible values using BAYES estimation of *previous MLR model****

```
ANALYSIS:      ESTIMATOR = BAYES;

MODEL:      ! Parameters from previously estimated measurement model, but all FIXED

 attn BY lnda@1;
 attn BY lnsa@0.51567 (l2);
 attn BY dscan@1.10655 (l3);

 [ lnda@4.35396 ];
 [ lnsa@5.58076 ];
 [ dscan@-0.01244 ];
 [ attn@0 ];

 lnda@0.51556 (e1);
 lnsa@0.08112 (e2);
 dscan@0.44855 (e3);
 attn@0.44310;

DATA IMPUTATION:                ! Creating plausible values for factor score
    NDATASETS = 100;            ! Number of separate values to create
    SAVE = PV/Attn*.dat;        ! Name of separate datasets with plausible values

SAVEDATA:   FILE = PV/AttnSummary.dat;   ! Summary about plausibles per person
            SAVE = FSCORES (100);        ! Needed to generate 100 factor scores
            FACTORS = Attn;              ! Which factors to save
            MISSFLAG = 99;               ! Missing data indicator for items
```

```
  Save file
    Attn*.dat

  Order of variables

    LNDA
    LNSA
    DSCAN
    PARTID
    ATTN
```

Now, within a subfolder of PV/, we have 100 datasets (named Attn1.dat to Attn100.dat) with these variables in this order. Thus, rather than just using the mean of a person's factor score distribution, we are *sampling* from it.

It also made a text file called "Attnlist.dat" that lists these individual data files:

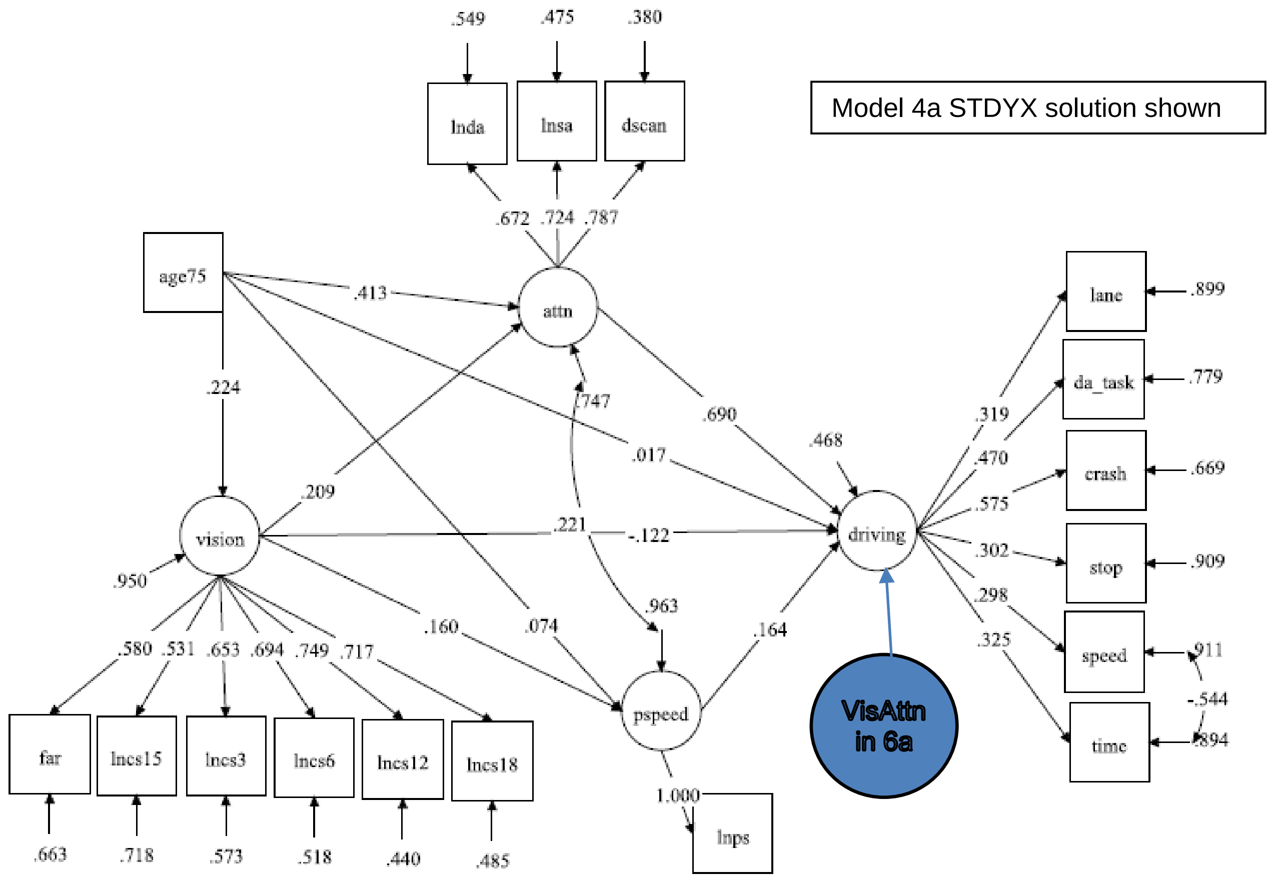
```
Attn1.dat
Attn2.dat
Attn3.dat
Attn4.dat
Attn5.dat
…..
```

Later we will use these types of files to tell Mplus to run the same analysis on every single file, then aggregate the results (as in multiple imputation).

---

Now we are ready to test the model of interest, **Model 4a** as shown below (drawn by Mplus, made prettier by me). We'll begin with a saturated structural model that has main effects of the latent variables only.

```
VARIABLE:  ! Every variable in THIS MODEL
            USEVARIABLES = lncs15 lncs3 lncs6 lncs12 lncs18 far
                           lane da_task crash stop speed time
                           lnda lnsa Dscan age75 lnps;
ANALYSIS:    ESTIMATOR = MLR;
```




```
MODEL:       ! Measurement models
  Vision BY far@1 lncs15* lncs3* lncs6* lncs12* lncs18*;   ! 1 marker loading
  [far* lncs15* lncs3* lncs6* lncs12* lncs18*];            ! All intercepts
   far* lncs15* lncs3* lncs6* lncs12* lncs18*;             ! Residual variances
  [Vision@0]; Vision*;                                     ! Factor M=0, Var=?

  Driving BY crash@1 da_task* lane* stop* speed* time*;    ! 1 marker loading
  [lane* da_task* crash* stop* speed* time*];              ! All intercepts
   lane* da_task* crash* stop* speed* time*;               ! Residual variances
  [Driving@0]; Driving*;                                   ! Factor M=0, Var=?
  speed WITH time* (ResCov);                               ! Residual covariance

 Attn BY lnda@1 lnsa* dscan*;                              ! 1 marker loading
  [lnda* lnsa* dscan*];                                    ! All intercepts
   lnda* lnsa* dscan*;                                     ! Residual variances
  [Attn@0]; Attn*;                                         ! Factor M=0, Var=?

 Pspeed BY lnps@1; lnps@0;          ! Bring proc speed into likelihood
  [lnps* Pspeed@0]; Pspeed*;        ! Move its variance to a factor, factor mean=0

! Structural model with all possible main effects
  Vision Attn Pspeed Driving ON Age75*  (Age1-Age4);   ! Age --> outcomes
         Attn Pspeed Driving ON Vision* (Vis1-Vis3);   ! Vision --> outcomes
  Attn WITH Pspeed*;                                   ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn* (Speed1 Attn1);             ! Pspeed, Attn --> Driving

MODEL CONSTRAINT:
NEW(AgeVis AgeSpeed AgeAttn);
AgeVis   = Age1*Vis3;    ! Indirect effect of age to vision to driving
AgeSpeed = Age3*Speed1; ! Indirect effect of age to proc speed to driving
AgeAttn  = Age2*Attn1;  ! Indirect effect of age to attention to driving
```

```
MODEL FIT INFORMATION
Number of Free Parameters                        58
Loglikelihood
          H0 Value                        -1310.811
          H0 Scaling Correction Factor       1.1063
            for MLR
          H1 Value                        -1238.221
          H1 Scaling Correction Factor       1.0405
            for MLR

Information Criteria
          Akaike (AIC)                     2737.622
          Bayesian (BIC)                   2913.007
          Sample-Size Adjusted BIC         2729.438
            (n* = (n + 2) / 24)

Chi-Square Test of Model Fit
          Value                             144.331*
          Degrees of Freedom                    110
          P-Value                            0.0156
          Scaling Correction Factor          1.0059
            for MLR

RMSEA (Root Mean Square Error Of Approximation)
          Estimate                            0.045
          90 Percent C.I.                     0.021  0.064
          Probability RMSEA <= .05            0.635

CFI/TLI
          CFI                                 0.936
          TLI                                 0.921

SRMR (Standardized Root Mean Square Residual)
          Value                               0.063

UNSTANDARDIZED MODEL RESULTS (TRUNCATED FOR SPACE)
                                                      Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
 VISION   BY
    FAR                1.000      0.000    999.000    999.000
    LNCS15             0.481      0.099      4.837      0.000
    LNCS3              0.584      0.115      5.076      0.000
    LNCS6              0.759      0.136      5.583      0.000
    LNCS12             1.265      0.203      6.248      0.000
    LNCS18             1.491      0.232      6.416      0.000

DRIVING  BY
    CRASH              1.000      0.000    999.000    999.000
    LANE               0.161      0.066      2.444      0.015
    DA_TASK            0.197      0.065      3.022      0.003
    STOP               0.381      0.164      2.330      0.020
    SPEED              0.418      0.164      2.540      0.011
    TIME               0.097      0.053      1.819      0.069

ATTN     BY
    LNDA               1.000      0.000    999.000    999.000
    LNSA               0.491      0.061      8.000      0.000
    DSCAN              1.192      0.170      7.022      0.000

PSPEED   BY
    LNPS               1.000      0.000    999.000    999.000

ATTN     ON
    VISION             0.287      0.137      2.095      0.036

PSPEED   ON
    VISION             0.167      0.100      1.658      0.097

DRIVING  ON
    VISION            -0.089      0.109     -0.814      0.415
    PSPEED             0.114      0.083      1.387      0.165
    ATTN               0.365      0.127      2.884      0.004

 VISION   ON
    AGE75              0.024      0.011      2.187      0.029
```

```
 ATTN     ON
    AGE75              0.059      0.014      4.393      0.000

 PSPEED   ON
    AGE75              0.008      0.008      0.988      0.323

 DRIVING  ON
    AGE75              0.001      0.011      0.119      0.905

ATTN     WITH
    PSPEED             0.061      0.027      2.292      0.022

SPEED    WITH
    TIME              -0.025      0.004     -5.512      0.000

 New/Additional Parameters
    AGEVIS            -0.002      0.003     -0.830      0.406
    AGESPEED           0.001      0.001      0.764      0.445
    AGEATTN            0.022      0.009      2.507      0.012

STANDARDIZED MODEL RESULTS (TRUNCATED FOR SPACE)
STDYX Standardization
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
 VISION   BY
    FAR                0.580      0.062      9.424      0.000
    LNCS15             0.531      0.076      6.999      0.000
    LNCS3              0.653      0.061     10.646      0.000
    LNCS6              0.694      0.059     11.851      0.000
    LNCS12             0.749      0.051     14.647      0.000
    LNCS18             0.717      0.042     17.024      0.000
DRIVING  BY
    CRASH              0.575      0.107      5.378      0.000
    LANE               0.319      0.130      2.446      0.014
    DA_TASK            0.470      0.100      4.694      0.000
    STOP               0.302      0.115      2.630      0.009
    SPEED              0.298      0.102      2.911      0.004
    TIME               0.325      0.132      2.470      0.014
 ATTN     BY
    LNDA               0.672      0.058     11.501      0.000
    LNSA               0.724      0.053     13.543      0.000
    DSCAN              0.787      0.045     17.608      0.000
 PSPEED   BY
    LNPS               1.000      0.000    999.000    999.000

DRIVING  ON
    VISION            -0.122      0.148     -0.826      0.409
    PSPEED             0.164      0.120      1.368      0.171
    ATTN               0.690      0.149      4.617      0.000

PSPEED   ON
    VISION             0.160      0.094      1.715      0.086
 ATTN     ON
    VISION             0.209      0.096      2.191      0.028
 DRIVING  ON
    AGE75              0.017      0.148      0.118      0.906
 VISION   ON
    AGE75              0.224      0.087      2.582      0.010
 ATTN     ON
    AGE75              0.413      0.081      5.085      0.000
 PSPEED   ON
    AGE75              0.074      0.075      0.986      0.324

ATTN     WITH
    PSPEED             0.221      0.088      2.523      0.012
 SPEED    WITH
    TIME              -0.544      0.090     -6.061      0.000

R-SQUARE
     Latent                                         Two-Tailed
    Variable        Estimate       S.E.  Est./S.E.    P-Value
    VISION             0.050      0.039      1.291      0.197
    DRIVING            0.532      0.151      3.526      0.000
    ATTN               0.253      0.077      3.264      0.001
    PSPEED             0.037      0.032      1.129      0.259
```

```
! Reduced structural model 5a (no age or vision --> driving)
  Vision Attn Pspeed ON Age75*  (Age2-Age4)      ! Age --> outcomes, not driving
         Attn Pspeed ON Vision* (Vis2-Vis3);     ! Vision --> outcomes, not driving
  Attn WITH Pspeed*;                             ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn*  (Speed1 Attn1);      ! Pspeed, Attn --> Driving
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       56
Loglikelihood
          H0 Value                       -1311.286
          H0 Scaling Correction Factor      1.0933
            for MLR
          H1 Value                       -1238.221
          H1 Scaling Correction Factor      1.0405
            for MLR
Information Criteria
          Akaike (AIC)                    2734.572
          Bayesian (BIC)                  2903.909
          Sample-Size Adjusted BIC        2726.670
            (n* = (n + 2) / 24)
Chi-Square Test of Model Fit
          Value                            144.090*
          Degrees of Freedom                   112
          P-Value                           0.0221
          Scaling Correction Factor         1.0142
            for MLR
RMSEA (Root Mean Square Error Of Approximation)
          Estimate                           0.043
          90 Percent C.I.                    0.018  0.063
          Probability RMSEA <= .05           0.691
CFI/TLI
          CFI                                0.940
          TLI                                0.927
SRMR (Standardized Root Mean Square Residual)
          Value                              0.063
```

Did constraining these two paths to 0 make the model worse?
Rescaled $-2\Delta LL(2) = 0.646$, $p = .72$, so no

This is the appropriate way to test a structural model, whose job is to reproduce the covariance among the latent factors and any observed predictors (but not among any observed predictors themselves).

Relying on good global model fit (which will mostly reflect the measurement models) is not sufficient to say a structural model fits.

We will continue with a full structural model instead so we can be sure that model mis-fit is not a reason behind any discrepancies.

**<u>What if we wanted to test a latent variable interaction? Model 6a (full structural model shown only)</u>**

*<u>Note that latent variable interactions can only be model predictors (and they cannot have covariances)</u>*

```
ANALYSIS:   ESTIMATOR = MLR;
            TYPE = RANDOM; ALGORITHM = INTEGRATION;   ! New estimation options needed
! Full structural model
  Vision Attn Pspeed Driving ON Age75*  (Age1-Age4);  ! Age --> outcomes
         Attn Pspeed Driving ON Vision* (Vis1-Vis3);  ! Vision --> outcomes
  Attn WITH Pspeed*;                                  ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn*  (Speed1 Attn1);           ! Pspeed, Attn --> Driving

! Interaction between two latent variables (would be same if one variable was observed)
  VisAttn | Vision XWITH Attn;      ! VisAttn = new latent variable interaction
  Driving ON VisAttn* (VxA);        ! Latent variable interaction --> Driving

MODEL CONSTRAINT: ! latent factor variance of attn = .443, of vision = .224
NEW (V4low V4high A4low A4high);
  V4low =  Vis3  - VxA*SQRT(.443);  ! Vision slope for -1SD attn
  V4high = Vis3  + VxA*SQRT(.443);  ! Vision slope for +1SD attn
  A4low =  Attn1 - VxA*SQRT(.224);  ! Attn slope for -1SD vision
  A4high = Attn1 + VxA*SQRT(.224);  ! Attn slope for -+1SD vision
```

```
MODEL FIT INFORMATION
Number of Free Parameters                       59
Loglikelihood
          H0 Value                       -1310.261
          H0 Scaling Correction Factor      1.1066
            for MLR
Information Criteria
          Akaike (AIC)                    2738.522
          Bayesian (BIC)                  2916.931
          Sample-Size Adjusted BIC        2730.197
            (n* = (n + 2) / 24)
```

Model fit has disappeared once we've used numeric integration (no H1 saturated covariance matrix to come back to anymore). STDYX disappears for the same reason.

**New structural model output only—note that the VisAttn interaction is related only to driving:**

```
UNSTANDARDIZED MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
ATTN       ON
    VISION             0.305      0.142      2.140      0.032

 PSPEED     ON
    VISION             0.168      0.101      1.662      0.096

DRIVING    ON
    VISION            -0.106      0.114     -0.924      0.355  simple vision slope at attn=0
    PSPEED             0.118      0.083      1.423      0.155
    ATTN               0.363      0.130      2.785      0.005  simple attn slope at vision=0
    VISATTN            0.139      0.142      0.978      0.328  n.s. interaction

VISION     ON
    AGE75              0.024      0.011      2.188      0.029

ATTN       ON
    AGE75              0.059      0.014      4.399      0.000

PSPEED     ON
    AGE75              0.008      0.008      0.982      0.326

DRIVING    ON
    AGE75              0.002      0.011      0.135      0.892

ATTN     WITH
    PSPEED             0.060      0.027      2.222      0.026

New/Additional Parameters
    V4LOW             -0.198      0.167     -1.181      0.237 simple vision slope at attn=-1SD
    V4HIGH             0.013      0.126     -0.105      0.916 simple vision slope at attn=+1SD
    A4LOW              0.297      0.139      2.134      0.033 simple attn slope at vision=-1SD
    A4HIGH             0.428      0.153      2.793      0.005 simple attn slope at vision=+1SD

STDYX Standardization
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 ATTN       ON
    VISION             0.220      0.099      2.233      0.026

 PSPEED     ON
    VISION             0.160      0.093      1.720      0.085

 DRIVING    ON
    VISION            -0.145      0.155     -0.939      0.348
    PSPEED             0.170      0.120      1.417      0.157
    ATTN               0.692      0.152      4.564      0.000
    VISATTN            0.125      0.126      0.999      0.318

 VISION     ON
    AGE75              0.227      0.088      2.594      0.009

 ATTN       ON
    AGE75              0.413      0.081      5.071      0.000

 PSPEED     ON
    AGE75              0.074      0.075      0.981      0.327

 DRIVING    ON
    AGE75              0.020      0.151      0.133      0.894

 ATTN     WITH
    PSPEED             0.217      0.088      2.448      0.014
```

What if we wanted to test *multiple* interactions between latent variables, or the model wouldn't converge (or there are too many latent variables to estimate all at once)? Or what if we had non-normal outcomes but we wanted to use maximum likelihood for our IRT/IFA model factor relations?

**Plausible values (PVs) to the rescue!** So far we've only done the measurement models "a" version (run model in MLR) and the "b" version (use Bayes estimation to generate plausible values for each latent factor). Now we need steps "c" and "d".

**Step "c"**: Merge the plausible values of different factors together, so that you have 100 complete datasets. Here is a SAS macro to automate that process:

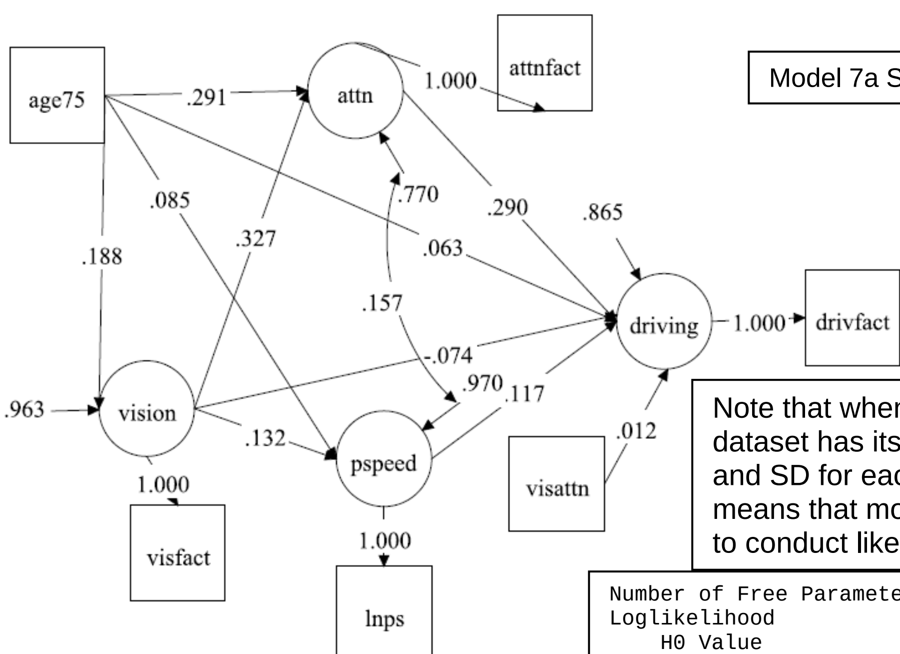
```
*********************************************************************************;
*****               INFO NEEDED TO ENTER TO MERGE PLAUSIBLE VALUES          ******;
*********************************************************************************;
* Folder for plausible files;     %LET filesave = C:\Dropbox\18_CLDP948\DriverScan\SEM\PV;
* SAS original data file name;    %LET datafile = MyOriginalData;
* Name of person ID variable;     %LET IDvar = PartID;
* Suffix # of FIRST file;         %LET startP = 1;
* Suffix # of LAST file;          %LET endP = 100;
* Total # of sets of files;       %LET total = 3;

%MACRO LabelThem;
* index = count of how many sets of files,
* prefix = name of file prefix;
* # items needed to drop from front of file;
%IF &index. = 1 %THEN %DO; %LET prefix = Vision;  %LET ndrop = 6; %END;
%IF &index. = 2 %THEN %DO; %LET prefix = Driving; %LET ndrop = 6; %END;
%IF &index. = 3 %THEN %DO; %LET prefix = Attn;    %LET ndrop = 3; %END;
**** REPEAT THE ABOVE FOR ALL YOUR SETS OF FILES TO BE MERGED ***********;
%MEND LabelThem;

*********************************************************************************;
*****               NOTHING NEEDS TO BE CHANGED FROM HERE, JUST RUN IT      ******;
*********************************************************************************;
* Sort original data by ID; PROC SORT DATA=&datafile.; BY &IDvar.; RUN;
%GLOBAL index prefix ndrop; %MACRO Import;
%DO num=&startp. %TO &endp.; DATA Merge&num.; SET &datafile.; RUN;
       %DO index=1 %TO &total.; %LabelThem; * Import plausible file;
              DATA &prefix.&num.; INFILE  "&filesave.\&prefix.&num..dat" DLM=TAB LRECL=1000;
              INPUT var1-var&ndrop. &IDvar. &prefix.; KEEP &IDvar. &prefix.; RUN;
              PROC SORT DATA=&prefix.&num.; BY &IDvar.; RUN;
              * Merge with original data, replace missing values;
              DATA Merge&num.; MERGE Merge&num. &prefix.&num.; BY &IDvar.;
              IF &prefix.=. THEN &prefix.=-9999; IF &IDvar.=. THEN &IDvar.=-9999; RUN;
              * Remove SAS datasets;
              PROC DATASETS LIB=WORK NOLIST; DELETE &prefix.&num.; RUN; QUIT;
       %END; * Export to .csv for use in Mplus;
              PROC EXPORT DATA=Merge&num. OUTFILE= "&filesave.\PV&num..csv"
              DBMS=CSV REPLACE; PUTNAMES=NO; RUN;
       * Remove SAS datasets;
       PROC DATASETS LIB=WORK NOLIST; DELETE Merge&num.; RUN; QUIT;
%END; %MEND Import;
* Run macro; %Import;
* Build list of plausible values files;
%MACRO Makelist;
DATA _NULL_; * Name of file to print to;
       FILE "&filesave.\PVFilesList.dat" NOPAD NOTITLES;
       * Print all dataset names;
       %DO i=&startP. %TO &endP.; PUT @1 "PV&i..csv"; %END;
RUN; %MEND Makelist;
* Run macro; %Makelist;
*********************************************************************************;
```

**When the SAS program is done running,** you will have 100 .csv files called PV*.csv with all plausible values for the latent variables merged together, as well as a file called "PVFileList.dat" that lists all these files. Now we are ready to analyze! (Btw, why 100? Because more should be better, right?)

**Step “d”**: Estimate the same model, but using the plausible values instead of the latent factors to build an observed interaction term. This tells Mplus to do so for all 100 files and then combine the results.

## Model 7a: Using Plausible Values to Estimate an SEM with a latent variable interaction

```
TITLE:  SEM Example for Driverscan using Plausible Values
DATA:
  FILE = PV/PVFilesList.dat;     ! FILE that lists all the data file names
  TYPE = IMPUTATION;             ! Analyze and combine results across files
VARIABLE:
! List of ALL variables in data file
  NAMES = PartID sex age75 lncs15 lncs3 lncs6 lncs12 lncs18 far lnps
               lnda lnsa Dscan lane da_task crash stop speed time
          Vision Driving Attn;  ! New factor scores
! Variables to be analyzed in this model
  USEVARIABLE = age75 lnps Vision Driving Attn VisAttn;
! Missing data identifier
  MISSING ARE ALL (-9999);
! ID variable;
  IDVARIABLE = PartID;
DEFINE:     VisAttn = Vision * Attn; ! Now interaction is observed variable
                                     ! but it will only predict driving for comparability
ANALYSIS:   ESTIMATOR = MLR;
OUTPUT:     STDYX RESIDUAL;      ! Standardized model, local fit
            SAMPSTAT;            ! Get descriptive stats for variables
MODEL:
! Measurement models for "factors" (factor mean=0 used for centering)
! Now assuming perfect reliability because of PVs
  Vision  BY VisFact@1;  Vision*  VisFact@0;  [Vision@0  VisFact*];
  Attn    BY AttnFact@1; Attn*    AttnFact@0; [Attn@0    AttnFact*];
  Pspeed  BY lnps@1;     Pspeed*  lnps@0;     [Pspeed@0  lnps*];
  Driving BY DrivFact@1; Driving* DrivFact@0; [Driving@0 DrivFact*];

! Structural model among "factors"
  Vision Attn Pspeed Driving ON Age75*  (Age1-Age4);   ! Age --> outcomes
         Attn Pspeed Driving ON Vision* (Vis1-Vis3);   ! Vision --> outcomes
  Attn WITH Pspeed*;                                   ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn* (Speed1 Attn1);             ! Pspeed, Attn --> Driving
  Driving ON VisAttn* (VxA);                           ! Interaction --> Driving

MODEL CONSTRAINT: ! Plausible score variance of attn = .441, of vision = .221
NEW (V4low V4high A4low A4high);
V4low =  Vis3  - VxA*SQRT(.441);  ! Vision slope for -1SD attn
V4high = Vis3  + VxA*SQRT(.441);  ! Vision slope for +1SD attn
A4low =  Attn1 - VxA*SQRT(.221);  ! Attn slope for -1SD vision
A4high = Attn1 + VxA*SQRT(.221);  ! Attn slope for -+1SD vision
```




Model 7a STDYX solution shown

Note that when using plausible values, each dataset has its own set of fit statistics, the mean and SD for each are given (see left example). This means that modified procedures must be followed to conduct likelihood ratio tests across datasets.

```
Number of Free Parameters                        19
Loglikelihood
    H0 Value
        Mean                                -392.820
        Std Dev                                9.076
        Number of successful computations        100
```

**What would have happened if we used the mean of each person's factor score distribution from the single-factor models as a single observed variable instead? Let's examine two ways of doing this.**

```
TITLE:  SEM Example for Driverscan using Uncorrected Single Factor Scores;
DATA:
  FILE = SEMfactorscores.dat;    ! Mean factor score merged back into data
  TYPE = INDIVIDUAL;             ! Now just a regular analysis
VARIABLE:
! List of ALL variables in data file
  NAMES = PartID sex age75 lncs15 lncs3 lncs6 lncs12 lncs18 far lnps
          lnda lnsa Dscan lane da_task crash stop speed time
          VisFact DrivFact AttnFact; ! New factor scores
! Variables to be analyzed in this model
  USEVARIABLE = age75 lnps VisFact DrivFact AttnFact VisAttn;
! Missing data identifier
  MISSING ARE ALL (-9999);
! ID variable;
  IDVARIABLE = PartID;
DEFINE:     VisAttn = VisFact * AttnFact; ! Interaction is observed variable
            CENTER VisAttn (GrandMean);   ! Mean-center for comparability
ANALYSIS:   ESTIMATOR = MLR;
OUTPUT:     STDYX RESIDUAL;     ! Standardized model, local fit
            SAMPSTAT;           ! Get descriptive stats for variables
```

## Model 8a: Using Reliability-Corrected Single Factor Scores (Model 10a does same using PVs)

```
MODEL:
! Measurement models for "factors" (factor mean=0 used for centering)
! Incorporates factor score unreliability
  Vision  BY VisFact@1;  Vision*;  VisFact*(ResVis);   [Vision@0  VisFact*];
  Attn    BY AttnFact@1; Attn*;    AttnFact*(ResAttn); [Attn@0    AttnFact*];
  Pspeed  BY lnps@1;     Pspeed*;  lnps@0;             [Pspeed@0  lnps*];
  Driving BY DrivFact@1; Driving*; DrivFact*(ResDriv); [Driving@0 DrivFact*];

! Structural model among "factors"
  Vision Attn Pspeed Driving ON Age75*  (Age1-Age4);    ! Age --> outcomes
         Attn Pspeed Driving ON Vision* (Vis1-Vis3);    ! Vision --> outcomes
  Attn WITH Pspeed*;                                    ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn* (Speed1 Attn1);              ! Pspeed, Attn --> Driving
  Driving ON VisAttn* (VxA);                            ! Interaction --> Driving

MODEL CONSTRAINT: ! New factor score variance of attn = .345, of vision = .186
NEW (V4low V4high A4low A4high);
V4low =  Vis3  - VxA*SQRT(.345);  ! Vision slope for -1SD attn
V4high = Vis3  + VxA*SQRT(.345);  ! Vision slope for +1SD attn
A4low =  Attn1 - VxA*SQRT(.186);  ! Attn slope for -1SD vision
A4high = Attn1 + VxA*SQRT(.186);  ! Attn slope for -+1SD vision

ResVis =(1-.856)*0.186;  ! Fix residual variances to "unreliable" part of factor score
ResAttn=(1-.819)*0.345;  ! Each comes from "a" version of measurement model
ResDriv=(1-.723)*0.097;
```

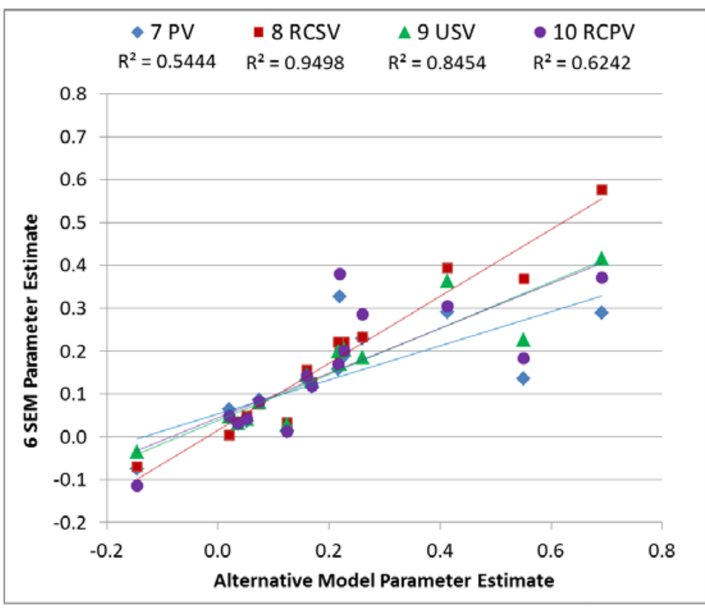
## Model 9a: Using Uncorrected Single Factor Scores

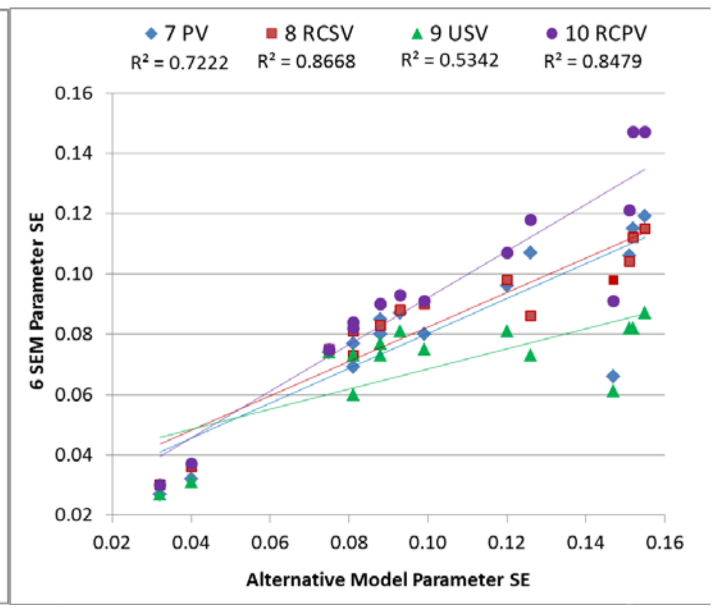
```
MODEL:
! Measurement models for "factors" (factor mean=0 used for centering)
! Now assuming perfect reliability
  Vision  BY VisFact@1;  Vision*  VisFact@0;  [Vision@0  VisFact*];
  Attn    BY AttnFact@1; Attn*    AttnFact@0; [Attn@0    AttnFact*];
  Pspeed  BY lnps@1;     Pspeed*  lnps@0;     [Pspeed@0  lnps*];
  Driving BY DrivFact@1; Driving* DrivFact@0; [Driving@0 DrivFact*];

! Structural model among "factors"
  Vision Attn Pspeed Driving ON Age75*  (Age1-Age4);    ! Age --> outcomes
         Attn Pspeed Driving ON Vision* (Vis1-Vis3);    ! Vision --> outcomes
  Attn WITH Pspeed*;                                    ! Res cov for Attn and Pspeed
  Driving ON Pspeed* Attn* (Speed1 Attn1);              ! Pspeed, Attn --> Driving
  Driving ON VisAttn* (VxA);                            ! Interaction --> Driving

MODEL CONSTRAINT: ! New factor score variance of attn = .345, of vision = .186
NEW (V4low V4high A4low A4high);
V4low =  Vis1  - VxA*SQRT(.345);  ! Vision slope for -1SD attn
V4high = Vis1  + VxA*SQRT(.345);  ! Vision slope for +1SD attn
A4low =  Attn1 - VxA*SQRT(.186);  ! Attn slope for -1SD vision
A4high = Attn1 + VxA*SQRT(.186);  ! Attn slope for -+1SD vision
```

**Model fit is equivalent between Models 8a and 9a, which are based on the same information and the same input data:** $\chi^2(3) = 0.912$, $p = .82$, RMSEA = 0 (0-0.82, $p = .89$), CFI = 1, SRMR = .02.

What about the results? Let's compare the standardized solution across our five options:

| | Estimates | | | | | Standard Errors | | | | | P-Values | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **MODEL** | **6 SEM** | **7 PV** | **8 RCSV** | **9 USV** | **10 RCPV** | **6 SEM** | **7 PV** | **8 RCSV** | **9 USV** | **10 RCPV** | **6 SEM** | **7 PV** | **8 RCSV** | **9 USV** | **10 RCPV** |
| **Age -->** | | | | | | | | | | | | | | | |
| VISION | .227 | .188 | .220 | .203 | .201 | .088 | .085 | .083 | .077 | .090 | **.009** | **.027** | **.008** | **.008** | **.026** |
| ATTN | .413 | **.291** | .393 | .362 | **.305** | .081 | .077 | .081 | .073 | .084 | **.000** | **.000** | **.000** | **.000** | **.000** |
| PSPEED | .074 | .085 | .076 | .081 | .082 | .075 | .074 | .075 | .074 | .075 | .327 | .251 | .313 | .275 | .276 |
| DRIVING | .020 | .063 | .004 | .046 | .048 | .151 | .106 | .104 | .082 | .121 | .894 | .554 | .968 | .576 | .691 |
| **Vision -->** | | | | | | | | | | | | | | | |
| PSPEED | .160 | .132 | .156 | .144 | .142 | .093 | .087 | .088 | .081 | .093 | .085 | .127 | .076 | .077 | .127 |
| ATTN | .220 | **.327** | .205 | .170 | **.379** | .099 | .080 | .090 | .075 | .091 | **.026** | **.000** | **.022** | **.022** | **.000** |
| **ATTN<-->PSPEED** | .217 | .157 | .220 | .198 | .169 | .088 | .080 | .083 | .073 | .090 | **.014** | .051 | **.008** | **.007** | .061 |
| **DRIVING <--** | | | | | | | | | | | | | | | |
| PSPEED | .170 | .117 | .126 | .129 | .119 | .120 | .096 | .098 | .081 | .107 | .157 | .227 | .198 | .110 | .269 |
| VISION | -.145 | -.074 | -.069 | -.035 | -.114 | .155 | .119 | .115 | .087 | .147 | .348 | .534 | .548 | .686 | .436 |
| ATTN | .692 | **.290** | .576 | .415 | .372 | .152 | .115 | .112 | .082 | .147 | **.000** | **.011** | **.000** | **.000** | **.011** |
| VISATTN | **.125** | .012 | .033 | .028 | .013 | .126 | .107 | .086 | .073 | .118 | .318 | .910 | .705 | .705 | .910 |
| **R2 Latent Variable** | | | | | | | | | | | | | | | |
| VISION | .052 | .037 | .048 | .041 | .042 | .040 | .032 | .036 | .031 | .037 | .195 | .257 | .185 | .186 | .256 |
| ATTN | .260 | .230 | .232 | .185 | .286 | .081 | .069 | .073 | .060 | .082 | **.001** | **.001** | **.002** | **.002** | .001 |
| PSPEED | .037 | .030 | .035 | .032 | .032 | .032 | .027 | .030 | .027 | .030 | .258 | .274 | .241 | .237 | .280 |
| DRIVING | .551 | **.135** | .369 | .226 | **.184** | .147 | .066 | .098 | .061 | .091 | **.000** | **.043** | **.000** | **.000** | **.043** |




From our informal comparison of methods, it looks like reliability-corrected versions of the models (8a and 10a) do the best job of reproducing parameter estimates (left figure) and standard errors (right figure) relative to the original SEM. Further, it appears the single-factor-score model did better than the plausible-values model, although we'd want to see this result replicate via simulation before making any conclusions. Results should differ more given greater unreliability of the factor scores (such as for driving here).

Note that a single estimate of reliability cannot be used when factors are created using IRT/IFA, in which reliability is trait-specific instead.

Also potentially problematic in creating the plausible values is that the model parameter estimates were fixed (treated as known) using the ML solution, rather than allowed to vary during the Bayesian estimation. So the factor scores are less variable than they would have been using full Bayesian estimation.