Appendix A - Annotated Mplus Output

7/25/2020

An annotated LTA model output estimated using Mplus (version 8) is shown below. Mplus input and output syntax is presented within gray colored text boxes. Some of the output has been appended to highlight infomation most relevant to applied ressearchers. Annotation is directed towards providing assistance with interpretation and comprehension.

LTA model input:

```
TITLE: Full invariance LTA - LSAY

DATA: FILE = "Invariance_LTA_LSAY.dat";

VARIABLE: NAMES =
  CASENUM COHORT FEMALE MINORITY
  AB39M AB39T AB39U AB39W AB39X
  MATHG7 MATHG10 MATHG12 OPKNOW OPKNOW2
  SIOK SIOK2 GA33A GA33H GA33I GA33K GA33L
  CPROB11 CPROB12 CPROB13 CPROB14 CMOD1
  CPROB21 CPROB22 CPROB23 CPROB24 CMOD2;

 categorical =
  AB39M AB39T AB39U AB39W AB39X    # grade 7 items
  GA33A GA33H GA33I GA33K GA33L;   # grade 10 items

 usevar =
  AB39M AB39T AB39U AB39W AB39X
  GA33A GA33H GA33I GA33K GA33L;

 missing=all(9999);

 classes = c1(4) c2(4);
# Enumeration indicated the four class solution at each time point
# C1 is the latent class variable for 7th grade
# C2 is the latent class variable for 10th grade

ANALYSIS:
  estimator = MLR;
  type = mixture;
  starts = 500 100; # Starting values may be increased to increase
                    # the likelihood of LL being a global maximum.

MODEL:
```

```
  %overall%
# Regression of the 10th grade latent class variable,c2, on the 7th grade latent class variable, c1.
  c2 on c1;

MODEL c1: # Measurement model for c1, the latent class variable for 7th grade (time point 1)

# To estimate thresholds the 7th grade items are mentioned under each class-specific statement
  %c1#1%
  [AB39M$1-AB39X$1] (1-5);    # (1-5) is included to label the 5 class-specific item thresholds.
                             #
  %c1#2%                     # By labeling them and using the same numbers for the first
  [AB39M$1-AB39X$1] (6-10);  # class under 'MODEL C2' we are constraining the item
                             # thresholds to be the same for each class.
  %c1#3%                     #
  [AB39M$1-AB39X$1] (11-15); # This is how measurement invariance across the LCA models for each
                             # time point is specified.
  %c1#4%                     #
  [AB39M$1-AB39X$1] (16-20); #

  MODEL c2: # Measurement model for c2, the latent class variable for 10th grade (time point 2)

  %c2#1%                     #
  [GA33A$1-GA33L$1] (1-5);   # Here label numbers are repeated, all parameter labels
                             # match those above (e.g., (1-5)).
  %c2#2%                     #
  [GA33A$1-GA33L$1] (6-10);  # The symbol '$' is used to refer to binary or ordinal variable thresholds
                             # Because items in this example are binary, only a single threshold
  %c2#3%                     # is mentioned.
  [GA33A$1-GA33L$1] (11-15); #
                             #
  %c2#4%                     #
  [GA33A$1-GA33L$1] (16-20); #
```

LTA model output:

Note, due to the estimation of a series of mixture models in LTA re-ordering of the classes occurs. Although the solution is the same, the transition matrix presented in the manuscript does not match the order of the values shown in this output. This is important to check and is part of the process of conducting LTA. To avoid class switching see the Mplus manual for adding the `OPTSEED` statement.

```
SUMMARY OF ANALYSIS # The analysis sample size for this model is N = 3092


Number of groups                                         1
Number of observations                                3092

Number of dependent variables                           10
Number of independent variables                          0
Number of continuous latent variables                    0
Number of categorical latent variables                   2
```

```
Observed dependent variables

  Binary and ordered categorical (ordinal)
    AB39M        AB39T        AB39U        AB39W        AB39X        GA33A
    GA33H        GA33I        GA33K        GA33L


Categorical latent variables
   C1          C2
```

Below are the response probabilities and counts for the data. This can be used to check if your data was read in correctly. That is, category 2 reflects $P(u_j = 1)$ from equation (1) in the main text.

```
UNIVARIATE PROPORTIONS AND COUNTS FOR CATEGORICAL VARIABLES


    AB39M
      Category 1    0.387         1177.000
      Category 2    0.613         1865.000
    AB39T
      Category 1    0.596         1782.000
      Category 2    0.404         1206.000
    AB39U
      Category 1    0.508         1519.000
      Category 2    0.492         1473.000
    AB39W
      Category 1    0.596         1796.000
      Category 2    0.404         1216.000
    AB39X
      Category 1    0.539         1639.000
      Category 2    0.461         1404.000
    GA33A
      Category 1    0.425          956.000
      Category 2    0.575         1294.000
    GA33H
      Category 1    0.572         1278.000
      Category 2    0.428          957.000
    GA33I
      Category 1    0.491         1097.000
      Category 2    0.509         1137.000
    GA33K
      Category 1    0.580         1297.000
      Category 2    0.420          941.000
    GA33L
      Category 1    0.584         1314.000
      Category 2    0.416          936.000
```

This Mplus model includes a total of 16 estimated response patterns, that is all combinations of the 4 classes from C1 by the 4 classes from C2, as shown in the output below. Class counts and proportions are shown in the first and second columns respectively. This section provides the joint distribution of C1 and C2. Note that the proportions in the last column sum to one.

```
FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS
BASED ON THE ESTIMATED MODEL
```

```
Latent Class
  Pattern

   1  1         501.60729              0.16223
   1  2          81.70191              0.02642
   1  3         139.59137              0.04515
   1  4         174.43899              0.05642
   2  1         244.35669              0.07903
   2  2         245.36893              0.07936
   2  3         136.04557              0.04400
   2  4         293.87883              0.09504
   3  1         156.28764              0.05055
   3  2          35.18978              0.01138
   3  3         132.52501              0.04286
   3  4         116.81557              0.03778
   4  1         179.19485              0.05795
   4  2         121.23235              0.03921
   4  3          96.73600              0.03129
   4  4         437.02922              0.14134
```

In this section the first column lists the latent variables `C1` and `C2`. The second Column lists the latent classes k that compose each latent variable. The third column is the model estimated class counts which are distinct from modal assignment counts as indicated by the non-integer estimates. Column four are the model estimated class proportions, note that each latent variable's respective proportions sum to one (i.e., For `C1` the first 4 rows).

```
FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE
BASED ON THE ESTIMATED MODEL

  Latent Class
    Variable     Class

    C1             1         897.33954              0.29021
                   2         919.65002              0.29743
                   3         440.81802              0.14257
                   4         834.19238              0.26979
    C2             1        1081.44653              0.34976
                   2         483.49295              0.15637
                   3         504.89795              0.16329
                   4        1022.16260              0.33058
```

The transition probability matrix describes stability (diagonal of matrix) or change (off-diagonal of matrix) for each class across the two latent class variables. Typically, these parameters are of central interest in an LTA model and are used to determine transition of observations between discrete classes across time-points. The values in the transition matrix are derived from the multinomial logistic regression of `C2` on `C1` and the latent class intercepts or the class size logit estimates (e.g., `[C1#1]`). The value in the upper-left of the matrix is interpreted as: 56% of those in Class 1 in 7th grade remained in Class 1 in 10th grade. The value in the top row column 2 is interpreted as: 9.1% of those in Class 1 in 7th grade transitioned to Class 2 in 10th grade. Note that the multinomial logistic regression estimates are presented farther down in the output.

```
LATENT TRANSITION PROBABILITIES BASED ON THE ESTIMATED MODEL

  C1 Classes (Rows) by C2 Classes (Columns)
```

	1	2	3	4
1	0.559	0.091	0.156	0.194
2	0.266	0.267	0.148	0.320
3	0.355	0.080	0.301	0.265
4	0.215	0.145	0.116	0.524

This section presents the joint probability patterns by latent class. Given there are 4 classes in 7th grade and 4 classes in 10th grade, there are 16 possible patterns. This set of marginals presents the prevalence of each of these 16 patterns. These patterns can be used to see the prevalence of specific transition patterns. For example, the pattern of being in class 1 at time 1 and class 1 at time 2 has 16% of the sample, whereas the pattern of being in class 3 in 7th grade and class 2 in 10th grade only has 1.1% of the sample.

```
FINAL CLASS COUNTS AND PROPORTIONS FOR THE LATENT CLASS PATTERNS
BASED ON ESTIMATED POSTERIOR PROBABILITIES

  Latent Class
    Pattern

    1  1         501.60733          0.16223
    1  2          81.70194          0.02642
    1  3         139.59130          0.04515
    1  4         174.43903          0.05642
    2  1         244.35675          0.07903
    2  2         245.36899          0.07936
    2  3         136.04550          0.04400
    2  4         293.87886          0.09504
    3  1         156.28756          0.05055
    3  2          35.18964          0.01138
    3  3         132.52514          0.04286
    3  4         116.81549          0.03778
    4  1         179.19486          0.05795
    4  2         121.23236          0.03921
    4  3          96.73600          0.03129
    4  4         437.02925          0.14134
```

Marginal probabilities based on `estimated posterior probabilities`. This estimate of the relative size of the classes are the most commonly reported. Note that it accounts for the non-perfect assignment to classes, thus shown see non-integer values for the class counts (column 3). This estimate is based on summing the columns for each posterior probability.

```
FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE
BASED ON ESTIMATED POSTERIOR PROBABILITIES

  Latent Class
    Variable     Class

    C1              1         897.33960          0.29021
                    2         919.65009          0.29743
                    3         440.81781          0.14257
                    4         834.19250          0.26979
    C2              1        1081.44641          0.34976
                    2         483.49292          0.15637
```

3	504.89792	0.16329
4	1022.16260	0.33058

Modal assignment: Below are the class counts and proportions if you assign individuals into classes based on their modal class assignment (e.g., the class with the highest posterior probability of membership). These are not commonly used for describing the classes. Note, if you use the ML 3-step, these are used but adjusted for non-perfect assignment to class.

```
Class Counts and Proportions

Latent Class
Pattern

1  1              392              0.17602
1  2               52              0.02335
1  3              121              0.05433
1  4              169              0.07589
2  1              193              0.08666
2  2              136              0.06107
2  3               75              0.03368
2  4              187              0.08397
3  1              108              0.04850
3  2                8              0.00359
3  3               77              0.03458
3  4               75              0.03368
4  1              148              0.06646
4  2               76              0.03413
4  3               60              0.02694
4  4              350              0.15716
```

Marginal distribution of the classes based on modal class assignment. These are not commonly reported as they do not account for the classification error present in the model. For reporting class distribution in manuscript, Table X is most commonly used.

```
NOTE " Maybe remove this section not relevant to applied researchers? " NOTE

FINAL CLASS COUNTS AND PROPORTIONS FOR EACH LATENT CLASS VARIABLE
BASED ON THEIR MOST LIKELY LATENT CLASS PATTERN

  Latent Class
     Variable    Class

     C1           1             1015             0.32827
                  2              700             0.22639
                  3              462             0.14942
                  4              915             0.29592
     C2           1             1246             0.40298
                  2              235             0.07600
                  3              388             0.12549
                  4             1223             0.39554
```

Below are the average posterior probability matrix $AvePP_k$ estimates, which are used as a metric of class-specific classification precision (diagonal of matrix) and mis-classification or error (off-diagonal of matrix). What often is reported is the diagonals of these in the LCA context (see Masyn, 2013).

	1	2	3	4	5	6	7	8	9
1	0.659	0.046	0.078	0.063	0.038	0.016	0.009	0.012	0.047
2	0.040	0.576	0.006	0.071	0.002	0.214	0.000	0.017	0.003
3	0.124	0.075	0.563	0.020	0.006	0.021	0.059	0.005	0.008
4	0.000	0.069	0.013	0.666	0.000	0.026	0.002	0.125	0.000
5	0.059	0.001	0.003	0.000	0.669	0.070	0.067	0.001	0.018
6	0.005	0.030	0.002	0.004	0.023	0.679	0.036	0.059	0.006
7	0.008	0.002	0.027	0.000	0.084	0.175	0.556	0.021	0.003
8	0.010	0.003	0.003	0.013	0.115	0.166	0.068	0.501	0.008
9	0.072	0.004	0.008	0.007	0.118	0.041	0.026	0.039	0.402
10	0.004	0.074	0.000	0.018	0.003	0.310	0.000	0.026	0.021
11	0.012	0.006	0.047	0.002	0.017	0.052	0.116	0.007	0.057
12	0.000	0.007	0.002	0.065	0.000	0.051	0.003	0.227	0.000
13	0.000	0.000	0.000	0.000	0.116	0.011	0.011	0.000	0.027
14	0.000	0.000	0.000	0.000	0.005	0.118	0.007	0.007	0.001
15	0.000	0.000	0.000	0.000	0.011	0.035	0.104	0.004	0.002
16	0.000	0.000	0.000	0.003	0.020	0.030	0.013	0.088	0.002

	10	11	12	13	14	15	16
1	0.004	0.016	0.009	0.002	0.000	0.000	0.000
2	0.057	0.001	0.012	0.000	0.000	0.000	0.000
3	0.006	0.103	0.003	0.001	0.001	0.005	0.001
4	0.007	0.004	0.086	0.000	0.000	0.000	0.002
5	0.001	0.003	0.000	0.095	0.005	0.008	0.000
6	0.048	0.005	0.010	0.003	0.068	0.007	0.015
7	0.001	0.028	0.000	0.008	0.013	0.066	0.006
8	0.004	0.009	0.027	0.007	0.007	0.004	0.055
9	0.032	0.140	0.087	0.011	0.003	0.002	0.008
10	0.434	0.000	0.090	0.000	0.018	0.000	0.002
11	0.038	0.614	0.015	0.002	0.003	0.011	0.001
12	0.043	0.020	0.558	0.000	0.002	0.000	0.022
13	0.001	0.004	0.000	0.715	0.049	0.067	0.001
14	0.002	0.001	0.000	0.051	0.657	0.048	0.102
15	0.002	0.040	0.001	0.076	0.131	0.548	0.046
16	0.001	0.002	0.009	0.073	0.075	0.047	0.636

This section presents the class-specific item threshold parameters which are estimated in the logit scale. These are often converted to probabilities. Note that since we assume measurement invariance in this example, the thresholds for consistent patterns (e.g., 1 1) the thresholds are constrained to be the equal. Logits estimated close to the extreme ends of the logit scale are fixed by Mplus to 15 and -15 to avoid estimation complications.

MODEL RESULTS

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

Latent Class Pattern 1 1

Thresholds
    AB39M$1   0.964   0.097   9.914   0.000

```
    AB39T$1           3.758        1.016        3.698        0.000
    AB39U$1           3.291        0.878        3.750        0.000
    AB39W$1           2.828        0.175       16.117        0.000
    AB39X$1           4.328        0.770        5.621        0.000
    GA33A$1           0.964        0.097        9.914        0.000
    GA33H$1           3.758        1.016        3.698        0.000
    GA33I$1           3.291        0.878        3.750        0.000
    GA33K$1           2.828        0.175       16.117        0.000
    GA33L$1           4.328        0.770        5.621        0.000

Latent Class Pattern 1 2

 Thresholds
    AB39M$1           0.964        0.097        9.914        0.000
    AB39T$1           3.758        1.016        3.698        0.000
    AB39U$1           3.291        0.878        3.750        0.000
    AB39W$1           2.828        0.175       16.117        0.000
    AB39X$1           4.328        0.770        5.621        0.000
    GA33A$1          -0.718        0.101       -7.093        0.000
    GA33H$1           0.830        0.147        5.636        0.000
    GA33I$1           0.287        0.146        1.965        0.049
    GA33K$1           0.289        0.174        1.663        0.096
    GA33L$1          -0.802        0.563       -1.424        0.154

# OUTPUT APPENDED

Latent Class Pattern 2 1

 Thresholds
    AB39M$1          -0.718        0.101       -7.093        0.000
    AB39T$1           0.830        0.147        5.636        0.000
    AB39U$1           0.287        0.146        1.965        0.049
    AB39W$1           0.289        0.174        1.663        0.096
    AB39X$1          -0.802        0.563       -1.424        0.154
    GA33A$1           0.964        0.097        9.914        0.000
    GA33H$1           3.758        1.016        3.698        0.000
    GA33I$1           3.291        0.878        3.750        0.000
    GA33K$1           2.828        0.175       16.117        0.000
    GA33L$1           4.328        0.770        5.621        0.000

Latent Class Pattern 2 2

 Thresholds
    AB39M$1          -0.718        0.101       -7.093        0.000
    AB39T$1           0.830        0.147        5.636        0.000
    AB39U$1           0.287        0.146        1.965        0.049
    AB39W$1           0.289        0.174        1.663        0.096
    AB39X$1          -0.802        0.563       -1.424        0.154
    GA33A$1          -0.718        0.101       -7.093        0.000
    GA33H$1           0.830        0.147        5.636        0.000
    GA33I$1           0.287        0.146        1.965        0.049
    GA33K$1           0.289        0.174        1.663        0.096
    GA33L$1          -0.802        0.563       -1.424        0.154
```

The following output are the parameters of the latent class variable model expressed as logits. The first section of parameters are the multinomial logistic regression or structural paths between the latent variables. These parameters are used to compute the transition probability matrix shown below (b11...b13). The section labeled 'Means' presents the logit intercepts for the latent class variable which are used to compute class size counts & proportions for the estimated model. These intercept parameters are used to compute the transition probability matrix shown in table below (a1...a3).

		C2 = 1	C2 = 2	C2 = 3	C2 = 4
Time 1	C1 = 1	a1+b11	a2+b21	a3+b31	0
	C1 = 2	a1+b12	a2+b22	a3+b32	0
	C1 = 3	a1+b13	a2+b23	a3+b33	0
	C1 = 4	a1	a2	a3	0

```
Categorical Latent Variables


# NOTE: Output has been altered due to rendering conflict, added symbols (%) and labels (!)

%C2#1%       ON
   %C1#1%          !b11!        1.948       0.193      10.109       0.000
   %C1#2%          !b12!        0.707       0.234       3.027       0.002
   %C1#3%          !b13!        1.183       0.282       4.194       0.000


%C2#2%       ON
   %C1#1%          !b21!        0.524       0.376       1.394       0.163
   %C1#2%          !b22!        1.102       0.300       3.669       0.000
   %C1#3%          !b23!        0.082       1.132       0.073       0.942


%C2#3%       ON
   %C1#1%          !b31!        1.285       0.330       3.898       0.000
   %C1#2%          !b32!        0.738       0.362       2.040       0.041
   %C1#3%          !b33!        1.634       0.542       3.015       0.003


 Means
   %C1#1%                       0.073       0.115       0.634       0.526
   %C1#2%                       0.098       0.154       0.635       0.525
   %C1#3%                      -0.638       0.464      -1.373       0.170
   %C2#1%          !a1!        -0.892       0.154      -5.786       0.000
   %C2#2%          !a2!        -1.282       0.265      -4.836       0.000
   %C2#3%          !a3!        -1.508       0.402      -3.747       0.000
```

Conditional item probabilities are derived from the threshold parameters using equation (2) in the paper. These are the values that are used when graphing the posterior probability plots as in Figure 2 (bottom). Note that since we assume measurement invariance in this analysis, the conditional item probabilities are constrained to be equal across time when the classes are equal.

```
RESULTS IN PROBABILITY SCALE


# The first pattern refers to C1=1, C2=1
# The conditional item probabilities are equal across C1 and C2 (invariance assumed).
```

Latent Class Pattern 1 1

C1=1, item probabilities for 7th grade

```
 AB39M
    Category 1          0.724        0.019      37.251       0.000
    Category 2          0.276        0.019      14.207       0.000
 AB39T
    Category 1          0.977        0.023      43.177       0.000
    Category 2          0.023        0.023       1.007       0.314
 AB39U
    Category 1          0.964        0.030      31.746       0.000
    Category 2          0.036        0.030       1.182       0.237
 AB39W
    Category 1          0.944        0.009     102.050       0.000
    Category 2          0.056        0.009       6.037       0.000
 AB39X
    Category 1          0.987        0.010      99.736       0.000
    Category 2          0.013        0.010       1.316       0.188
```

C2=1, item probabilities for 10th grade

```
 GA33A
    Category 1          0.724        0.019      37.251       0.000
    Category 2          0.276        0.019      14.207       0.000
 GA33H
    Category 1          0.977        0.023      43.177       0.000
    Category 2          0.023        0.023       1.007       0.314
 GA33I
    Category 1          0.964        0.030      31.746       0.000
    Category 2          0.036        0.030       1.182       0.237
 GA33K
    Category 1          0.944        0.009     102.050       0.000
    Category 2          0.056        0.009       6.037       0.000
 GA33L
    Category 1          0.987        0.010      99.736       0.000
    Category 2          0.013        0.010       1.316       0.188
```

C1=1, C2=2

Latent Class Pattern 1 2
C1=2, item probabilities for 7th grade

```
 AB39M
    Category 1          0.724        0.019      37.251       0.000
    Category 2          0.276        0.019      14.207       0.000
 AB39T
    Category 1          0.977        0.023      43.177       0.000
    Category 2          0.023        0.023       1.007       0.314
 AB39U
    Category 1          0.964        0.030      31.746       0.000
    Category 2          0.036        0.030       1.182       0.237
 AB39W
```

```
   Category 1              0.944      0.009    102.050        0.000
   Category 2              0.056      0.009      6.037        0.000
AB39X
   Category 1              0.987      0.010     99.736        0.000
   Category 2              0.013      0.010      1.316        0.188

# C2=2, item probabilities for 10th grade

GA33A
   Category 1              0.328      0.022     14.685        0.000
   Category 2              0.672      0.022     30.122        0.000
GA33H
   Category 1              0.696      0.031     22.361        0.000
   Category 2              0.304      0.031      9.750        0.000
GA33I
   Category 1              0.571      0.036     15.973        0.000
   Category 2              0.429      0.036     11.990        0.000
GA33K
   Category 1              0.572      0.043     13.441        0.000
   Category 2              0.428      0.043     10.068        0.000
GA33L
   Category 1              0.310      0.120      2.571        0.010
   Category 2              0.690      0.120      5.734        0.000

# OUTPUT APPENDED
```