Mplus syntax for double cross-validation using latent class analysis (LCA) and comparing outcomes across classes; accompanies:

Merians, A. N., Baker, M. R., Frazier, P. A., & Lust, K. (2018, in press). Outcomes related to adverse childhood experiences in college students: Comparing latent class analysis and cumulative risk. *Child Abuse & Neglect.*

Description: Mplus syntax for double cross-validation using latent class analysis (LCA) and comparing outcomes across classes. This includes exploratory LCA to identify a best fitting model (syntax A below), cross-validating the model in separate halves of the study sample (B), and comparing outcomes (i.e., mental health, physical health, alcohol consequences, and GPA) across latent classes using a bias-adjusted, three-step analysis for comparing outcomes across latent classes (C).

A. Exploratory analysis in Half 1 to identify best fitting model

Note: We performed exploratory LCA in Half 2 as well, but only show Half 1 here for brevity.

```
TITLE:  LCA_split1
DATA: FILE IS Splithalf1_LCA_04112017.csv; ! CSV file containing raw data in wide
                                           ! format without a header column
VARIABLE: NAMES ARE ! identify all variables in CSV
IDnum mntill drink drug incarc divrc
domvio CPA CEA CSA filter split;
USEVARIABLES ARE mntill drink
drug incarc divrc domvio CPA CEA CSA ; ! identify the variables to be used in the LCA
CATEGORICAL  = mntill drink drug incarc divrc ! identify the variables that are categorical
domvio CPA CEA CSA;
MISSING = all(-999); ! identify what values in the CSV to be treated as missing
CLASSES = c(1); ! started with 1 class model, increased model number (2), (3), etc. until fit
                ! ceased improving or worsened
ANALYSIS: TYPE = MIXTURE; ! tell Mplus to perform mixture modeling
LRTSTARTS = 50 25 30 25;  ! specify the number of random starts for the parametric
                          ! bootstrapped likelihood ratio test displayed in the TECH14
                          ! output, relevant for comparing k to k - 1 class models, increased
                          ! when the number of random starts failed to replicate
OUTPUT: TECH10 TECH11 TECH14; ! request TECH10 output to print response patterns
                              ! frequencies and chi-square contributions, TECH11 output to print
                              adjusted Lo-Mendell-Rubin likelihood ratio test, TECH14 to print
                              parametric bootstrapped likelihood ratio test
```

```
SAVEDATA: FILE IS 04112017_LCA_splithalf1_useme.csv; SAVE IS CPROB;
        !SAVE CPROB to save the most likely class membership for each participant to the CSV
        to be then merged into SPSS for predicting outcomes using ANOVA for comparison
        purposes only (note: most likely class membership alone is not recommended to be used
        for predicting outcomes, see Feingold, Tiberio, & Capaldi, 2012)
```

B. Cross-validating the Half 1 4-class model in Half 2 sample

Note: We performed four cross-validations for the study (half 1 values in half 2 sample for the 4-class and 5-class models, and half 2 values in half 1 sample), but only show one cross-validation here for brevity.

```
TITLE:  LCA_split2
DATA: FILE IS Splithalf2_LCA_04112017.csv; ! CSV file containing raw data of Half 2
VARIABLE: NAMES ARE ! identify all variables in CSV
IDnum mntill drink drug incarc divrc
domvio CPA CEA CSA filter split;
USEVARIABLES ARE mntill drink  ! identify the variables to be used in the LCA
drug incarc divrc domvio CPA CEA CSA ;
CATEGORICAL = mntill drink drug incarc divrc ! identify the variables that are categorical
domvio CPA CEA CSA;
MISSING = all(-999); ! identify what values in the CSV to be treated as missing
CLASSES = c(4); ! specify a 4-class model because we are validating the 4-class model
ANALYSIS: TYPE = MIXTURE; ! tell Mplus to perform mixture modeling
MODEL:
%OVERALL%  ! describes the overall part of a mixture model
[C#1@-1.805]; ! set the latent class means in the current model to be equal to the
                ! estimated latent class means in the Half 1 4-model output created using
                ! syntax from A above. Latent means are estimated for k-1 classes.
  [C#2@-1.150];
  [C#3@-1.307];
%c#1%          ! describes the class specific part of a mixture model
[MNTILL$1@-1.427];  ! in each class, set thresholds for each indicator variable to be equal to
                        ! the estimated thresholds in the Half 1 4-model output created using
                        ! syntax from A above.
[DRINK$1@-1.571];
[DRUG$1@-0.888];
[INCARC$1@-0.045];
[DIVRC$1@-0.551];
```

[DOMVIO$1@-0.825];
[CPA$1@-0.581];
[CEA$1@-2.744];
[CSA$1@0.589];
%c#2%
[MNTILL$1@-0.147];
[DRINK$1@-0.215];
[DRUG$1@0.535];
[INCARC$1@1.325];
[DIVRC$1@0.325];
[DOMVIO$1@2.271];
[CPA$1@15];
[CEA$1@0.269];
[CSA$1@2.105];
%c#3%
[MNTILL$1@0.447];
[DRINK$1@1.226];
[DRUG$1@3.201];
[INCARC$1@3.187];
[DIVRC$1@0.667];
[DOMVIO$1@0.41];
[CPA$1@-0.404];
[CEA$1@-2.225];
[CSA$1@1.383];
%c#4%
[MNTILL$1@1.741];
[DRINK$1@3.17];
[DRUG$1@4.496];
[INCARC$1@5.376];
[DIVRC$1@1.72];
[DOMVIO$1@4.006];
[CPA$1@4.332];
[CEA$1@1.139];
[CSA$1@3.027];

C. Bias-adjusted, three-step analysis for comparing outcomes across latent classes in Half 1

Note: We compared outcomes across latent classes in half 1 and half 2, and show only half 1 here for brevity.

```
TITLE:  LCA_outcomes1
DATA:FILE IS LCA_revisions_half1_2017.11.17.csv;  ! CSV file containing raw data of Half 1
VARIABLE: NAMES ARE                    ! identify all variables in CSV
IDnum mntill drink drug incarc divrc
domvio CPA CEA CSA
mntht phsht alc GPA;
USEVARIABLES are mntill drink
drug incarc divrc domvio CPA CEA CSA ! the latent class indicators are the same as before
mntht phsht alc GPA;          ! include the 4 continuous outcomes that will be predicted by class
CATEGORICAL  = mntill drink drug incarc divrc  ! identify the variables that are categorical
domvio CPA CEA CSA;
MISSING = all(-999);          ! identify what values in the CSV to be treated as missing
CLASSES = c(4);       ! specify a 4-class model because we are validating the 4-class model
AUXILIARY = mntht(DCONTINUOUS) phsht(DCONTINUOUS)
alc(DCONTINUOUS) GPA (DCONTINUOUS);
        ! include the 4 outcomes as auxiliary variables and identify them as continuous variables
ANALYSIS: TYPE = MIXTURE; ! tell Mplus to perform mixture modeling
MODEL:
%OVERALL%          ! set latent class means and item thresholds to the values that resulted from
                   ! the exploratory analysis; this ensures that the outcome variables are
                   compared across the latent class model that was decided on in previous
                   steps of analysis; results will be displayed in the output under Equality of
                   Means/Probabilities Across Classes
[C#1@-1.805];
[C#2@-1.150];
[C#3@-1.307];
%c#1%
[MNTILL$1@-1.427];
[DRINK$1@-1.571];
[DRUG$1@-0.888];
[INCARC$1@-0.045];
[DIVRC$1@-0.551];
[DOMVIO$1@-0.825];
[CPA$1@-0.581];
[CEA$1@-2.744];
[CSA$1@0.589];
%c#2%
[MNTILL$1@-0.147];
```

```
[DRINK$1@-0.215];
[DRUG$1@0.535];
[INCARC$1@1.325];
[DIVRC$1@0.325];
[DOMVIO$1@2.271];
[CPA$1@15];
[CEA$1@0.269];
[CSA$1@2.105];
%c#3%
[MNTILL$1@0.447];
[DRINK$1@1.226];
[DRUG$1@3.201];
[INCARC$1@3.187];
[DIVRC$1@0.667];
[DOMVIO$1@0.41];
[CPA$1@-0.404];
[CEA$1@-2.225];
[CSA$1@1.383];
%c#4%
[MNTILL$1@1.741];
[DRINK$1@3.17];
[DRUG$1@4.496];
[INCARC$1@5.376];
[DIVRC$1@1.72];
[DOMVIO$1@4.006];
[CPA$1@4.332];
[CEA$1@1.139];
[CSA$1@3.027];
OUTPUT: TECH10;
```