Homework #6
Confirmatory Factor Analysis

Listed below is a homework exercise for *Confirmatory Factor Analysis*. This exercise's main purpose is to help students consolidate the material learned in this section. Use *Mplus* to determine the answers to the following questions. Make sure that you submit to the Instructor *via e-mail* by Tuesday at 4pm: your typed answers to the questions below, as well as your syntax and output.

Please refer to the attached Mplus syntax file, labeled "HW6Cfa.inp". In this example, we do not have access to the raw data. Rather, we have access to a covariance matrix, which we can read in Mplus in order to analyze the model. The syntax file thus has only the required information in order to run the covariance matrix, which is found in the attached "HW6Cfa.dat" file. Make sure these two files are in the same folder, in order for the syntax to run properly.

This dataset is taken from a social psychological experiment conducted by Reisenzein (1986). His full experiment was on helping behavior, where he used a hypothetical story about a person collapsing and lying on a subway floor. As part of the experimental manipulation, half of the subjects were told that the person was drunk, and the others were told the person was ill. The author tested whether a subject's feelings and sympathy and anger mediate the likelihood of their helping the victim. This example, however, only focuses on the sympathy and anger latent variables, which the author measured with three observed/indicator variables each ($N = 138$; see below for the measures items).

Let us assume that the data do *not* violate the assumptions of normality, independence of observations, homoscedasticity or homogeneity of variance, multicollinearity, singularity or linearity, and have no missing values.

Specify a confirmatory factor analysis model (ML estimation), in which the sympathy factor is measured by the sympathy, pity and concern variables, and the anger factor is measured by the anger, irritation and aggravation variables. Allow the two factors to covary (i.e., have an unanalyzed association).

1) Model specification
    a. Draw the model. Remember to include in your drawing all observed variables, paths, variances, and disturbances, and fix any parameters that require fixing.
2) Model identification
    a. How many observations are in this model, and how did you arrive at this number?
    b. How many parameters are in this model, and how did you arrive at this number?
    c. Is the model identified, and why?
3) Analysis of the model
    a. How well does the hypothesized model fit the observed data?
    b. Which variable loads highest on the sympathy factor? Which variable loads highest on the anger factor?
    c. How much variance in sym1 is not accounted for by the sympathy factor?
    d. Do any residual covariances look suspicious to you?

   i. What looks suspicious about them?
e. If you re-analyze the model, but using generalized least squares as the estimator (instead of ML), what happens to the fit statistics?

Measure's items (all variables measured 1=not at all to 9=very much)
Sym1: How much sympathy would you feel for that person?
Sym2: I would not feel pity for this person.
Sym3: How much concern would you feel for this person?
Ang1: How angry would you feel at this person?
Ang2: How irritated would you feel by that person?
Ang3: I would feel aggravated by that person.