**Mplus**

Mplus is a lovely software package, useful for advanced statistical modeling of various kinds. If you expect that structural equations models, mediational models, multilevel data, or latent classes form any part of your future, you should invest in learning (and possibly purchasing) Mplus. It is quickly overtaking LISREL, AMOS, EQS, etc. as the industry standard.

There is quite a bit of excellent support available for Mplus. The website statmodel.net contains a massive wealth of resources, including not only a thorough User’s Guide with lots and lots of examples, but also example scripts you can run yourself and even **video short courses complete with handouts** for a wide variety of complex statistical techniques. Working through them is a great way to spend a summer before you propose your dissertation.

The website also features excellent discussion boards, in which **the authors of the software** (Bengt & Linda Muthén)answer the questions of confused users around the world. If there is something you want to do in Mplus but don’t know how, search the boards; someone has almost certainly had that question before you and the Muthéns have almost certainly answered it.

Today, we’ll only scratch the surface of what we can do in Mplus: we’ll get some data in and calculate basic descriptive statistics.

**Getting Data into Mplus**

Mplus is on the whole quite simple and elegant, but it has one small flaw: a quirky, cumbersome and potentially frustrating data input process.

**Mplus requires a .csv file with no header information.**

If there are variable names in the first row of your data file, you have to delete them or it will not read the file. You have to add the variable names later within your script file.

The challenge here is that you risk making errors in the renaming process that can have major consequences for your results and interpretation, because, let’s face it, one column of numbers looks a lot like another. My usual solution to this (you are welcome to develop your own) is to keep one copy of the data **with** headers and a second nearly identical copy without, titled “Lab 7 Data for Mplus.csv” or similar, so I know what’s going on if I open the file and wonder why there are no headers.

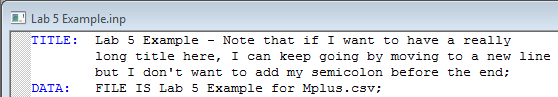
**The Logic of Mplus**

Open up the Mplus editor, and download the “Lab 4 Data for Mplus” file from Canvas. You will start with a blank script or input file, which looks like a simple text file. We’re going to build a very, very simple analysis here.

A few things to understand about how Mplus thinks:

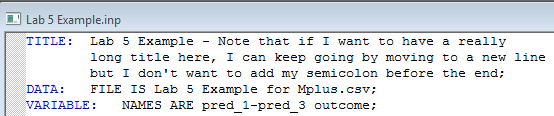
* In Mplus, you enter all of your script at once and then run it all (not line by line as in R or other programs). This means that **every analysis needs its own script file** – you can’t just keep adding new analyses in the same script, or Mplus becomes very confused. It also makes line-by-line troubleshooting a little more difficult, but
* Mplus uses **commands** (keywords) to figure out what you want it to do. There are major commands (TITLE, DATA, VARIABLE, ANALYSIS, MODEL, OUTPUT are some of the common ones) and then there are subcommands (loosely analogous to arguments in R functions) that fill in the details. For example, under VARIABLE you might have NAMES ARE, MISSING ARE, GROUPING IS, etc. There is a (fairly) complete list of possible commands in Chapter 20 of the User’s Guide.
* The combination of commands you need changes with the type of analysis; as with arguments in R, some are optional and some only apply to certain types of analyses. The best way to figure out which commands you need is to look at the examples in the User’s Guide.
* Mplus is **NOT** case sensitive.
* Mplus **IS** picky about how many characters are in a line. If there are more than 50 characters on one line, it **will not run**. You will need to add a line break (i.e., hit enter) in the middle of your command.
* Also, every command needs to end with a semicolon. Typically, this will mean a semicolon at the end of every line, **unless** you have a command that spills over onto a second (or third!) line. Do **not** type that semicolon until you are at the end of the command, but make sure you have a semicolon before you move on to a new command!

Give your analysis a title and then tell Mplus what data file to use:

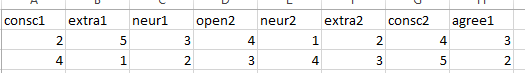


* Mplus will look for the data file in whatever folder the script is in. If you save this file in the same folder as your data, all you need is the file name. It does not need to be in quotes, but you can use them if you want to.

Now, we have to tell Mplus the names of our variables in the script, because they aren’t in the data file.

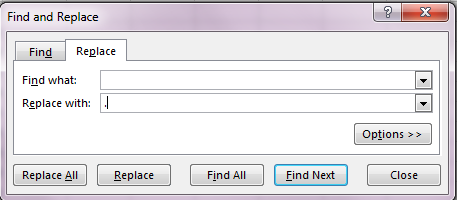


* Variable names in Mplus can only be 8 characters (they can be longer, but only the first 8 will be used, which can make it challenging to read the results). They can contain underscores (\_) and hyphens (-) but not periods or other punctuation. They can end with a number but cannot begin with a number.
* There are actually 3 predictors in this file, labeled pred\_1, pred\_2, and pred\_3. Mplus uses shorthand to simplify entering lots of variables. Instead of typing “item1 item2 item3 item4 item5 item6,” we can just type item1-item6. Mplus will fill in the sequence – **as long as there are no gaps** and the part before the number is constant for all of the items.
* This does not work if your variables are mixed up in your data file, e.g.:

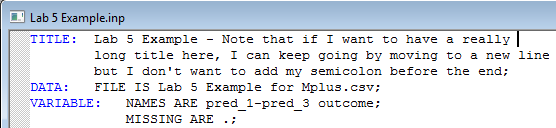


Remember, this is the **only** information Mplus has to tell which variable is which! Doublecheck this and make sure you’ve got it right.

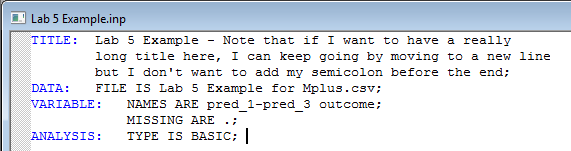
Now, tell Mplus how to tell missing data when it sees it. There are no missing data in the current file, but that won’t always be the case. ☺ You can code missing data in several ways (e.g., -999); my favorite in Mplus is “.”, because there’s no way I’m going to accidentally interpret that as a number. If your missing data are just blank cells, you need to go back to your Excel file and do a find & replace like this:



Then tell Mplus that’s your code for missing data:



And finally, ask Mplus for the BASIC analysis, which will give you your descriptive statistics:



Now, go to the RUN command on the Mplus menu and run your script!



In the output, Mplus will first repeat your input script back to you, along with some technical details about the analysis. Then it will give you results. It is a good idea to check things like the number of groups, observations, and variables to make sure the data was imported correctly.

**Lab Exercise:**

1. Download the file “Lab 4 Exercise Data.csv” from Canvas and save it somewhere on your computer.
2. Format the file properly so that you can read it into Mplus.
3. Use Mplus to obtain the descriptive statistics.
4. Turn in your Mplus output.