Two-Level Hierarchical Linear Models

Using SAS, Stata, HLM, R, SPSS, and Mplus

Division of
Statistics + Scientific Computation
THE UNIVERSITY OF TEXAS AT AUSTIN

September 2012

Table of Contents

Introduction

This document serves to compare the procedures and output for two-level hierarchical linear models from six different statistical software programs: SAS, Stata, HLM, R, SPSS, and Mplus. We compare these packages using the *popular.csv* dataset from Chapter 2 of Joop Hox's *Multilevel Analysis* (2010), which can be downloaded from:

http://joophox.net/mlbook2/DataExchange.zip

The six models described below are all variations of a two-level hierarchical model, also referred to as a multilevel model, a special case of mixed model. This comparison is only valid for completely nested data (not data from crossed or other designs, which can be analyzed with mixed models). Although the website for the HLM software states that it can be used for crossed designs, this has not been confirmed. The procedures used in SAS, Stata, R, SPSS, and Mplus below are part of their multilevel or mixed model procedures, and can be expanded to non-nested data. But for the purposes of this comparison, we will only investigate a fully nested dataset.

The code/syntax used for each model is included below for all programs except HLM, which is completely run by a GUI. We have provided screen shots of HLM and SPSS for each model. In addition, each model is specified in a hierarchical format as well as a mixed format. Although these two expressions of the models are equivalent, some research fields prefer to visualize the hierarchical structure because it is easier to see the separation between levels, while others prefer the mixed format, where it is easier to distinguish between fixed and random effects.

Model Considerations

When adding predictors into the six models discussed in this document, we chose to **grand mean center** them, meaning that we subtracted the overall mean of that variable from each subject's score. Centering at the grand mean, as opposed to the group mean (where the mean of each group is subjected from the score of subjects within that group), will not be appropriate for all models, as discussed in detail by Enders & Tofighi (2007). **The choice of which centering method to use should be driven by the specific research question being asked**.

Another consideration is the method of estimation used by these programs to produce the parameter estimates, either **maximum likelihood** (*ML*) or **restricted maximum likelihood** (*REML*). Each has its own advantages and disadvantages. *ML* is better for unbalanced data, but it produces biased results. *REML* is unbiased, but it cannot be used when comparing two nested models with a likelihood ratio test. **Both methods will produce the same estimates for fixed effects, yet they do differ on the random effect estimates** (Albright & Marinova, 2010).

As we'll see in the models discussed below, the two methods produce very similar results, and do not greatly affect the p-values of the random factors. However, it is important to be aware that the choice of method can impact the estimate, standard error, and p-values of the random factors and could potentially impact the decision of declaring a random factor significant or not. SAS, HLM, R, and SPSS use *REML* by default, while Stata and Mplus use *ML*. In the Stata examples throughout this document, we tell Stata to use *REML* in order to compare the output with the other four programs. However, Mplus does not have such an option, but can only use *ML*, so you will see minor differences in the random variance estimates in the Mplus output compared to the other programs throughout this document.

Intraclass Correlation Coefficient

We have also reported the intraclass correlation coefficient (ICC), ρ, for each model. The ICC is the proportion of variance in the outcome variable that is explained by the grouping structure of the hierarchical model. It is calculated as a ratio of group-level error variance over the total error variance:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2} \; ,$$

where $\sigma_{u_0}^2$ is the variance of the level-2 residuals and σ_e^2 is the variance of the level-1 residuals. In other words, the ICC reports on the amount of variation *unexplained* by any predictors in the model that can be attributed to the grouping variable, as compared to the overall unexplained variance (within and between variance).

Example Dataset

The *popular* dataset consists of students from different classes and because each student belongs to one unique class, it is a nested design. The dependent variable is *Popular,* a self-rated popularity scale ranging from 0-10. Predictors include *Sex* (dichotomous) and *Extrav* (continuous self-rated extraversion score) at the student level and *Texp* (teacher experience in years, which is continuous) at the class level.

Intercept-only Model (Unconditional Model)

<u>Mixed Model</u>

$$Popular_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

<u>Hierarchical Model</u>

$$Popular_{ij} = \beta_{0j} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

The unconditional mixed model specification resembles a one-factor ANOVA with γ_{00} as the overall mean and u_{0j} as the class effect. However, we are considering u_{0j} as a random effect (a normally distributed variable with a mean of zero), not a fixed factor effect as in ANOVA. Thus, we interpret the estimate for u_{0j} as the variance of the mean for each class around the overall mean *Popular* score.

The estimate for γ_{00} is the mean of the means of *Popular* for each class, instead of the mean of all students in the study. If the data were completely balanced (i.e. same number of students in every class), then the results of the unconditional model will equal those from an ANOVA procedure.

SAS Results

```
proc mixed data=popdata covtest;
model popular = /solution;
random intercept /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.1	<0.001

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	1.221	0.040	6.46	<0.001
Intercept (u_{0j})	0.702	0.109	30.8	<0.001

The "covtest" option is needed to report the standard errors of the variance component estimates. Also, you need to specify the unstructured covariance matrix type, which is what HLM and R use by default, and we use here for comparison.

The output from SAS is equal to the results in Table 2.1 of Hox's book. We can conclude that mean *Popular* score among classes is 5.078, and that there is more variation within the classes (1.221) than among the different classes (0.702). This will be discussed further when we calculate the ICC for this model.

Stata Results

```
xtmixed popular || class: , variance reml
```

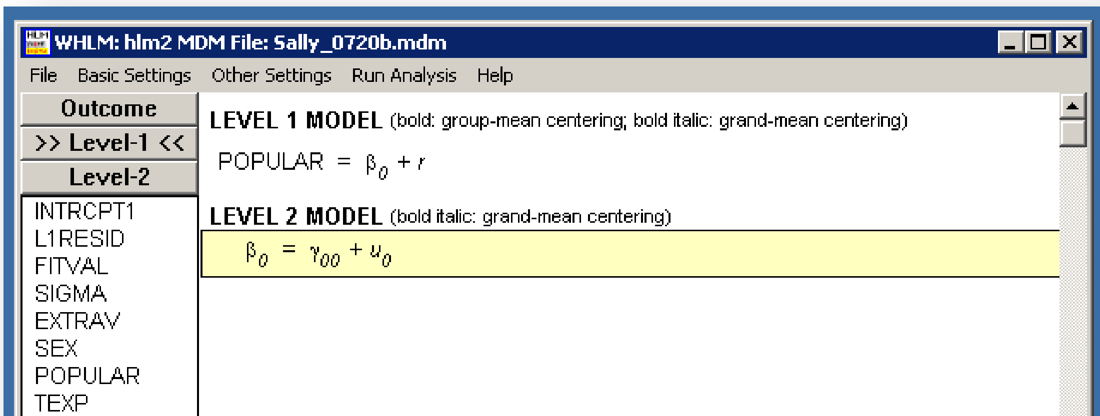
Fixed Effects	*Estimate*	*St. Error*	*z-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.1	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	1.221	0.040
Intercept (u_{0j})	0.702	0.109

Stata's *xtmixed* command requires the dependent variable followed by "||" which specifies the separation between the fixed and random variables. We must include the *variance* option to see the estimates for the variance components in the output, as well as the *reml* option to estimate using **restricted maximum likelihood**.

Also note that Stata does not output the p-values of the random component estimates, but significance can be determined by whether or not zero is contained in the confidence interval. These results exactly match those from SAS.

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.1	<0.001

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	1.222	1.105		
Intercept (u_{0j})	0.702	0.838	1227.3	<0.001

HLM reports the standard deviations, not the standard errors, of the variance components. Also, for the random effects, they report the Chi-squared statistic and p-value for the intercept only. These results equal those from the other programs.

R Results

```
library(lme4)
library(languageR)
lmer(popular ~ 1 + (1|class))
pvals.fnc(my_intonly)
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.1	<0.001

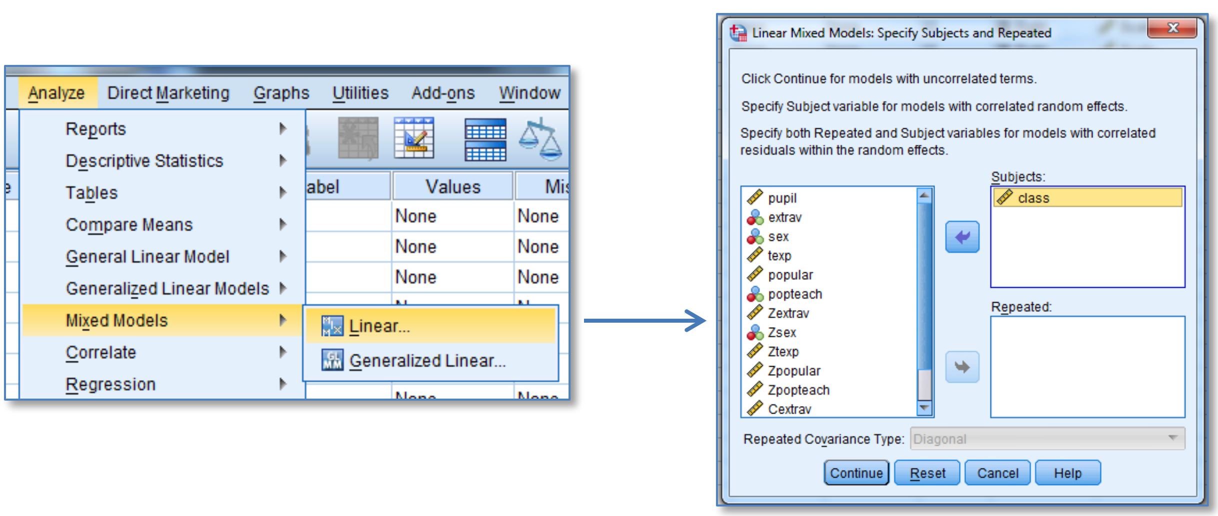
Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	1.221	1.105
Intercept (u_{0j})	0.702	0.838

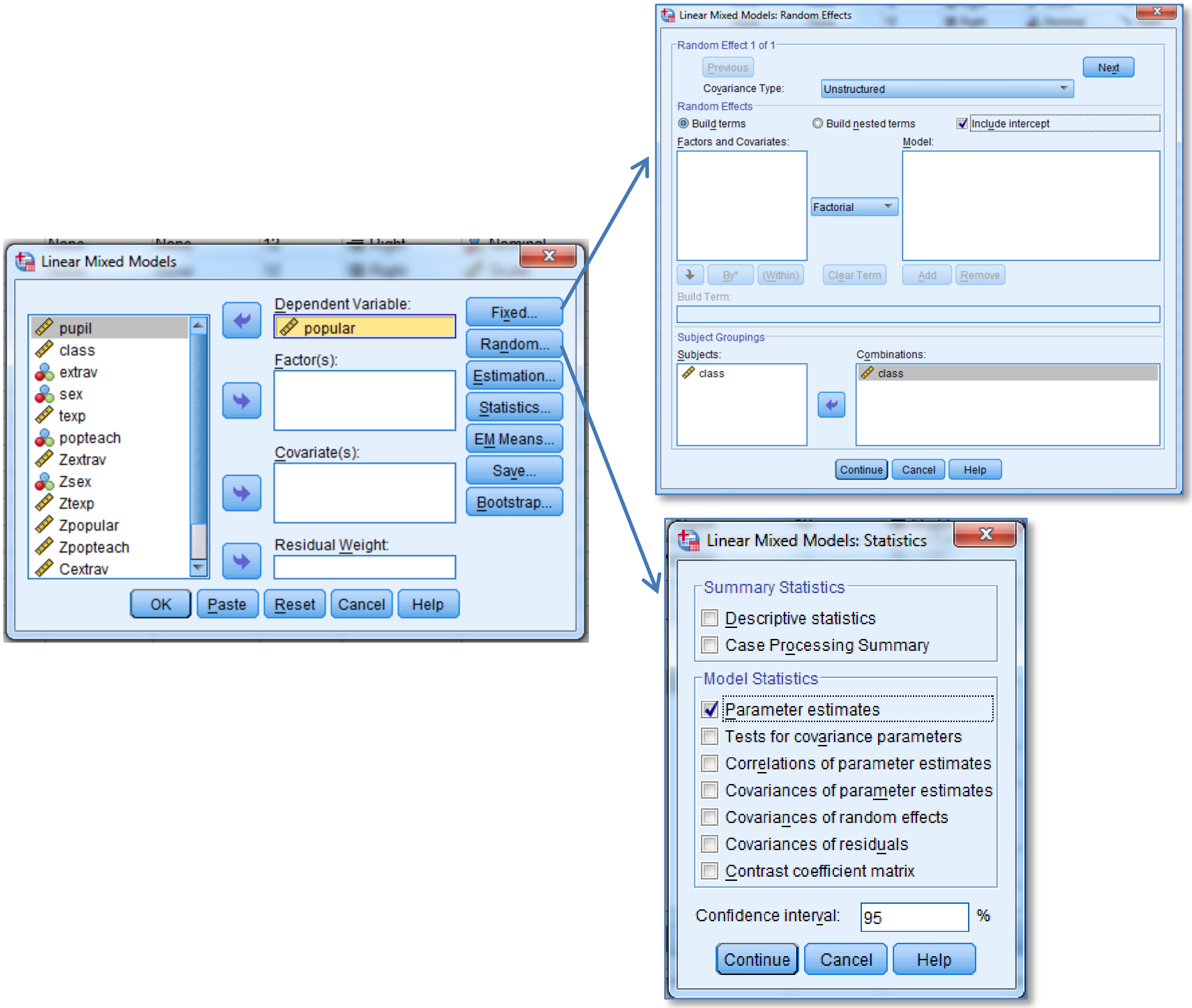
R reports the standard deviations of the variance components, like HLM, and the *lme4* package reports the t-statistic of the fixed effects. However, the *languageR* package is needed to get the **p-values** of the fixed effects. These results equal those from the other programs and the book.

SPSS Results

```
MIXED popular
 /FIXED=| SSTYPE(3)
 /METHOD=REML
 /PRINT=G  SOLUTION
 /RANDOM=INTERCEPT | SUBJECT(class) COVTYPE(UN).
```

Or follow the screen shots below:





Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.1	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	1.221	0.040
Intercept (u_{0j})	0.702	0.109

You need to specify the unstructured covariance type in the “Random” window. These results equal those from the other programs and the book. **Note that SPSS, like SAS and Mplus, reports the standard error of the variance components, while HLM and R report the standard deviation**. We have not been able to conclude which is more appropriate to report, but the difference does not affect the p-values for these parameters.

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
           Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
           USEVARIABLES ARE class popular;
           CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
         %BETWEEN%
OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.087	58.4	<0.001

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	1.222	0.047	26.2	<0.001
Intercept (u_{0j})	0.695	0.108	6.4	<0.001

Because this is an unconditional model, we don't need to specify any WITHIN or BETWEEN variables. The criteria for **listing variables in the MODEL statement** are below. We'll see examples of the first three in the following sections:

1. **%WITHIN%** – Level-1 fixed factors (non-random slope)
2. **%WITHIN% with latent slope variable** – Level-1 random factors
3. **%BETWEEN%** – Level-2 fixed factors
4. **Don't specify in either statement** – Variables measured at the student level but with a Level-1 and Level-2 variance estimate (we're not sure if/when this would be applicable for a multilevel model, and we won't see this in any of the models discussed in this document).

The above table shows the results from the "Model Results" section at the bottom of the Mplus output. Mplus does report p-values for each estimate, and all estimates match those from the other programs except for the variance estimate of the random intercept, which differs by about 0.007. This difference is due to the fact that Mplus uses *ML* estimation. Despite this difference, we do not see a change in the significance of any variables.

Model Summary

Overall, the six programs produce very similar results for the intercept-only model (with the only differences occurring in the Mplus estimate of the random effect). The only difference is how they report the precision of the random variance estimates.

The ICC for this model is equal to:

$$\rho = \frac{\sigma^2_{u_{0j}}}{\sigma^2_{u_{0j}} + \sigma^2_{e_{ij}}} = \frac{.702}{.702 + 1.221} = 0.365 \quad ,$$

which tells us that about one-third of the total variation in *Popular* can be accounted for by which class each student is in.

Random Intercept with One Fixed Level-1 Factor (Non-Random Slope)

Mixed Model

$$Popular_{ij} = \gamma_{00} + \gamma_{10}\boldsymbol{Extrav_{ij}} + u_{0j} + e_{ij}$$

Hierarchical Model

$$Popular_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{Extrav_{ij}} + e_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10}$$

This model has added one student-level fixed factor, *Extrav*, the self-reported extraversion score. The mixed model looks like an ANCOVA based on class with the covariate *Extrav*, but remember we still consider u_{0j} to be a random effect, not a fixed effect. Thus, the estimate for γ_{10} differs from what would be found by an ANCOVA procedure.

In the real application of this data, it does not make sense that *Extrav* should have a fixed effect instead of a random effect, since levels of student extraversion should vary by class. However, for the purposes of comparing the four programs, we still want to investigate a case with one student-level fixed factor.

SAS Results

```
proc mixed data=popdata covtest;
model popular = extrav_c /solution;
random intercept /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	53.9	<0.001
Extraversion (γ_{10})	0.486	0.020	24.1	<0.001

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	0.930	0.030	30.8	<0.001
Intercept (u_{0j})	0.841	0.127	6.64	<0.001

We now have an estimate for the fixed effect of *Extrav*. For every one unit increase in a student's reported extraversion score, there is a 0.486 increase in their popularity score. These results equal those from the other programs which use *REML*.

Stata Results

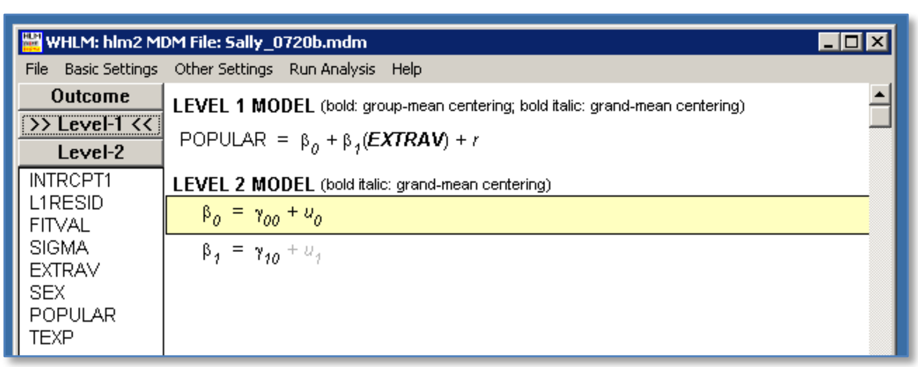
xtmixed popular cextrav || class: , variance cov(un) reml

Fixed Effects	*Estimate*	*St. Error*	*z-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	53.9	<0.001
Extraversion (γ_{10})	0.486	0.020	24.1	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.930	0.030
Intercept (u_{0j})	0.841	0.127

As we add predictors to the model in Stata, we add the *cov(un)* option, specifying an unstructured covariance matrix. We placed the centered *Extraversion* variable before the "||" to indicate that it is a fixed factor (with a non-random slope). These results equal those from the other programs.

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	53.9	<0.001
Extraversion (γ_{10})	0.486	0.020	24.1	<0.001

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	0.930	0.965		
Intercept (u_{0j})	0.841	0.917	1865.7	<0.001

R Results

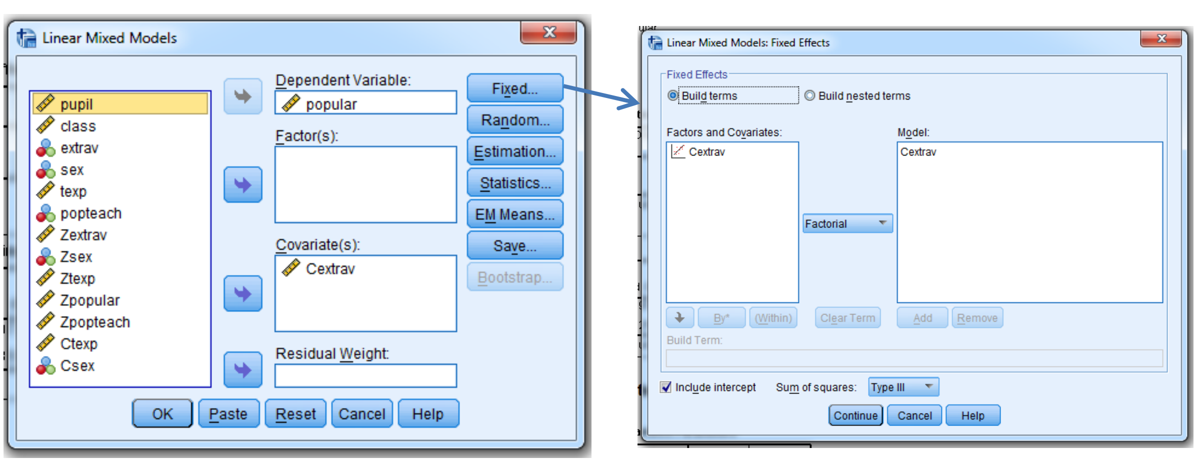
```
my_1fixed <- lmer(popular ~ 1 + c_extrav + (1|class))
pvals.fnc(my_1fixed)
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	53.9	<0.001
Extraversion (γ_{10})	0.486	0.020	24.1	<0.001

Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	0.930	0.965
Intercept (u_{0j})	0.841	0.917

SPSS Results

```
MIXED popular WITH Cextrav
 /FIXED=INTERCEPT Cextrav | SSTYPE(3)
 /METHOD=REML
 /PRINT=SOLUTION
 /RANDOM=INTERCEPT  | SUBJECT(class) COVTYPE(UN).
```



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	53.9	<0.001
Extraversion (γ_{10})	0.486	0.020	24.1	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.930	0.030
Intercept (u_{0j})	0.841	0.127

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
           Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
           USEVARIABLES ARE class popular Cextrav;
           WITHIN = Cextrav;
           CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
         popular ON Cextrav;
         %BETWEEN%
         popular;
OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.078	0.094	54.2	<0.001
Extraversion (γ_{10})	0.486	0.027	18.3	<0.001

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	0.930	0.028	33.5	<0.001
Intercept (u_{0j})	0.831	0.126	6.6	<0.001

We now include the centered *Extrav* variable in the WITHIN option of the VARIABLE statement. We must use the "ON" option for the within-level MODEL specification to tell Mplus that *Extrav* is a fixed level-1 factor. Again, you can see that there are minor differences in many of the estimates and standard error of the estimates (and therefore the t-statistics) due to using *ML* estimation instead of *REML*. Because the estimate for the variance of u_{0j} is different than the other programs, the ICC reported by Mplus differs from what is reported below.

Model Summary

For this model, the first five programs have exactly the same results, while the estimates from Mplus are off by very small margins. The ICC for this model is larger than for the unconditional model (as expected, since we are **controlling for some student-level variation by adding a fixed factor**):

$$\rho = \frac{.841}{.841 + .930} = 0.475$$

With one student-level fixed factor, almost one-half of the total variation in *Popular* can be accounted for by both the class of the student and the student-level fixed factor *Extraversion*.

Random Intercept and Slope for One Level-1 Factor

Mixed Model

$$Popular_{ij} = \gamma_{00} + \gamma_{10}\boldsymbol{Extrav_{ij}} + u_{1j}\boldsymbol{Extrav_{ij}} + u_{0j} + e_{ij}$$

Hierarchical Model

$$Popular_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{Extrav_{ij}} + e_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$

This model contains a random slope for *Extrav*, which means that we are **allowing the slope of our regression equation to vary by class**. This model is more appropriate than the previous model for the variables being used since it is intuitive to assume that extraversion varies from class to class.

SAS Results

```
proc mixed data=popdata covtest;
model popular = extrav_c /solution;
random intercept extrav_c /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.031	0.097	51.9	<0.001
Extraversion (γ_{10})	0.493	0.025	19.4	<0.001

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	0.895	0.030	30.0	<0.001
Intercept (u_{0j})	0.892	0.135	6.6	<0.001
Extraversion (u_{1j})	0.026	0.009	2.8	0.003

The estimate for the random *Extrav* slope is significant (p-value of 0.003), and therefore we would say that the student extraversion scores do vary by class. These results exactly match those from the other programs, except for some small discrepancies in the t-statistics for the fixed effects.

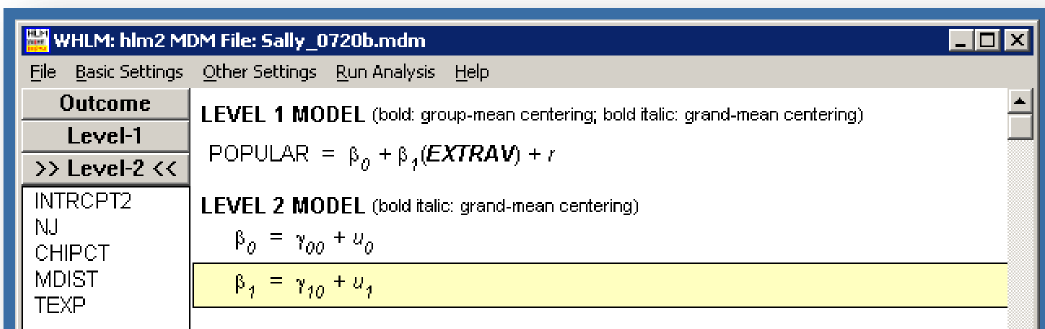
Stata Results

xtmixed popular cextrav || class: cextrav, variance cov(un) reml

Fixed Effects	*Estimate*	*St. Error*	*z-stat*	*p-value*
Intercept (γ_{00})	5.031	0.097	51.9	<0.001
Extraversion (γ_{10})	0.493	0.025	19.4	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.895	0.030
Intercept (u_{0j})	0.892	0.135
Extraversion (u_{1j})	0.026	0.009

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.031	0.097	52.1	<0.001
Extraversion (γ_{10})	0.493	0.025	19.5	<0.001

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	0.895	0.946		
Intercept (u_{0j})	0.892	0.944	1,589.4	<0.001
Extraversion (u_{1j})	0.026	0.162	180.6	<0.001

R Results

```
my_1rnd <- lmer(popular ~ 1 + c_extrav + (1 + c_extrav|class))
p.values.lmer(my_1rnd)
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.031	0.097	51.9	<0.001
Extraversion (γ_{10})	0.493	0.025	19.4	<0.001

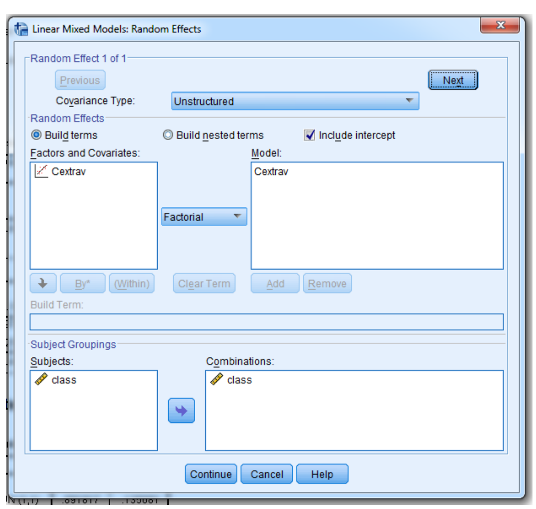
Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	0.895	0.946
Intercept (u_{0j})	0.892	0.944
Extraversion (u_{1j})	0.026	0.161

The R package “languageR” will not output the p-values for models with random coefficients (other than the intercept), so in order to get p-values for the fixed effects in this model (and all models below), you need to run the “p.values.lmer” function, which can be found at:

http://www.biostat.umn.edu/~julianw/courses/pubh7430/code/p.values.lmer.r

SPSS Results

```
MIXED popular WITH Cextrav
  /FIXED=INTERCEPT Cextrav | SSTYPE(3)
  /METHOD=REML
  /PRINT=SOLUTION
  /RANDOM=INTERCEPT Cextrav | SUBJECT(class) COVTYPE(UN).
```



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.031	0.097	51.9	<0.001
Extraversion (γ_{10})	0.493	0.025	19.4	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.895	0.030
Intercept (u_{0j})	0.892	0.135
Extraversion (u_{1j})	0.026	0.009

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
                Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
                USEVARIABLES ARE class popular Cextrav;
                WITHIN = Cextrav;
                CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
             randoms1 | popular ON Cextrav;
             %BETWEEN%
             popular;
OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.045	0.095	52.9	<0.001
Extraversion (γ_{10})	0.485	0.026	18.3	<0.001

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	0.894	0.027	33.2	<0.001
Intercept (u_{0j})	0.892	0.130	6.5	<0.001
Extraversion (u_{1j})	0.029	0.010	2.8	0.005

This time we **included a latent slope variable in the WITHIN statement** to specify *Extrav* as a random factor, which tells Mplus not to look for "randoms1" in the dataset because it is not observed. You can interpret the output for this variable as the random slope component of *Extrav*. We must do this because Mplus is designed for structural equation models, and its multilevel model capability is an adaptation of its underlying latent analysis procedures.

Model Summary

Overall, the first five programs produce the same results for this model, while Mplus again differs by a small amount due to *ML* estimation. The ICC for this model is:

$$\rho = \frac{.892}{.892 + .895} = 0.500$$

By changing the effect of *Extrav* from fixed to random, **the ICC increases slightly since we are considering more random variation at the student level**.

Random Slope for Two Level-1 Factors

<u>Mixed Model</u>

$$Popular_{ij} = \gamma_{00} + \gamma_{10}\boldsymbol{Extrav_{ij}} + \gamma_{20}\boldsymbol{Sex_{ij}} + u_{1j}\boldsymbol{Extrav_{ij}} + u_{2j}\boldsymbol{Sex_{ij}} + u_{0j} + e_{ij}$$

<u>Hierarchical Model</u>

$$Popular_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{Extrav_{ij}} + \beta_{2j}\boldsymbol{Sex_{ij}} + e_{ij}$$
$$\beta_{0j} = \gamma_{00} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + u_{2j}$$

For this model, we are including a second student-level variable, *Sex*, which also has a random slope, u_{2j}. This means that we are accounting for both the gender of the students as well as their extraversion score, and we are allowing the slopes of both of these factors to vary by class.

SAS Results

```
proc mixed data=popdata covtest;
model popular = extrav_c sex_c /solution;
random intercept extrav_c sex_c /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.028	0.084	59.8	<0.001
Extraversion (γ_{10})	0.443	0.023	18.9	<0.001
Sex (γ_{20})	1.244	0.036	34.1	<0.001

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	0.554	0.019	29.9	<0.001
Intercept (u_{0j})	0.674	0.102	6.6	<0.001
Extraversion (u_{1j})	0.030	0.008	3.8	<0.001
Sex (u_{2j})	<0.001	-	-	-

In this output, we can see that gender does have a significant effect on a student's self-reported popularity (p-value < 0.001). The fixed estimate for *Sex*, γ_{20}, means that female students (*Sex* = 1) have a *Popular* score that is 1.244 higher than male students (the baseline group, *Sex* = 0), holding *Extrav* constant.

SAS did not like that the estimated variance for *Sex* was so close to zero in this model, and therefore did not output a standard error or p-value. Because u_{2j} is extremely close to zero, we would conclude that gender does not vary significantly by class. These results equal the other programs within a few thousandths of a unit.

Stata Results

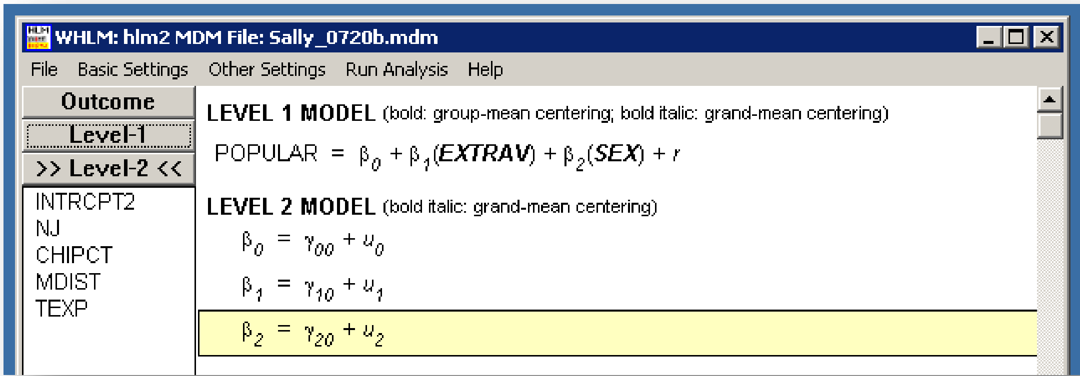
xtmixed popular cextrav csex || class: cextrav csex, variance cov(un) reml

Fixed Effects	*Estimate*	*St. Error*	*z-stat*	*p-value*
Intercept (γ_{00})	5.028	0.084	59.8	<0.001
Extraversion (γ_{10})	0.443	0.023	18.9	<0.001
Sex (γ_{20})	1.244	0.036	34.1	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.553	.
Intercept (u_{0j})	0.674	.
Extraversion (u_{1j})	0.030	.
Sex (u_{2j})	0.005	.

Stata cited an error when running this model: **standard error calculation failed**, meaning that the standard errors for the random effects were not calculated. We found that by removing the *cov(un)* option, this error did not appear. However, all of the estimates in that output differed from the other programs, so we choose to report the output with the unstructured covariance matrix specification. We are not sure if this is a common problem with running this type of model in Stata, but it is important to be aware that it can happen.

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.027	0.084	60.1	<0.001
Extraversion (γ_{10})	0.443	0.023	19.0	<0.001
Sex (γ_{20})	1.244	0.036	34.7	<0.001

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	0.553	0.743		
Intercept (u_{0j})	0.673	0.820	1,331.2	<0.001
Extraversion (u_{1j})	0.030	0.172	168.2	<0.001
Sex (u_{2j})	0.007	0.081	80.3	>0.5

These estimates roughly equal the results from the other programs, except in the estimate for the random gender effect. Since this effect was so close to zero, the programs do no report exactly the same value, but all show that it is far from significant.

R Results

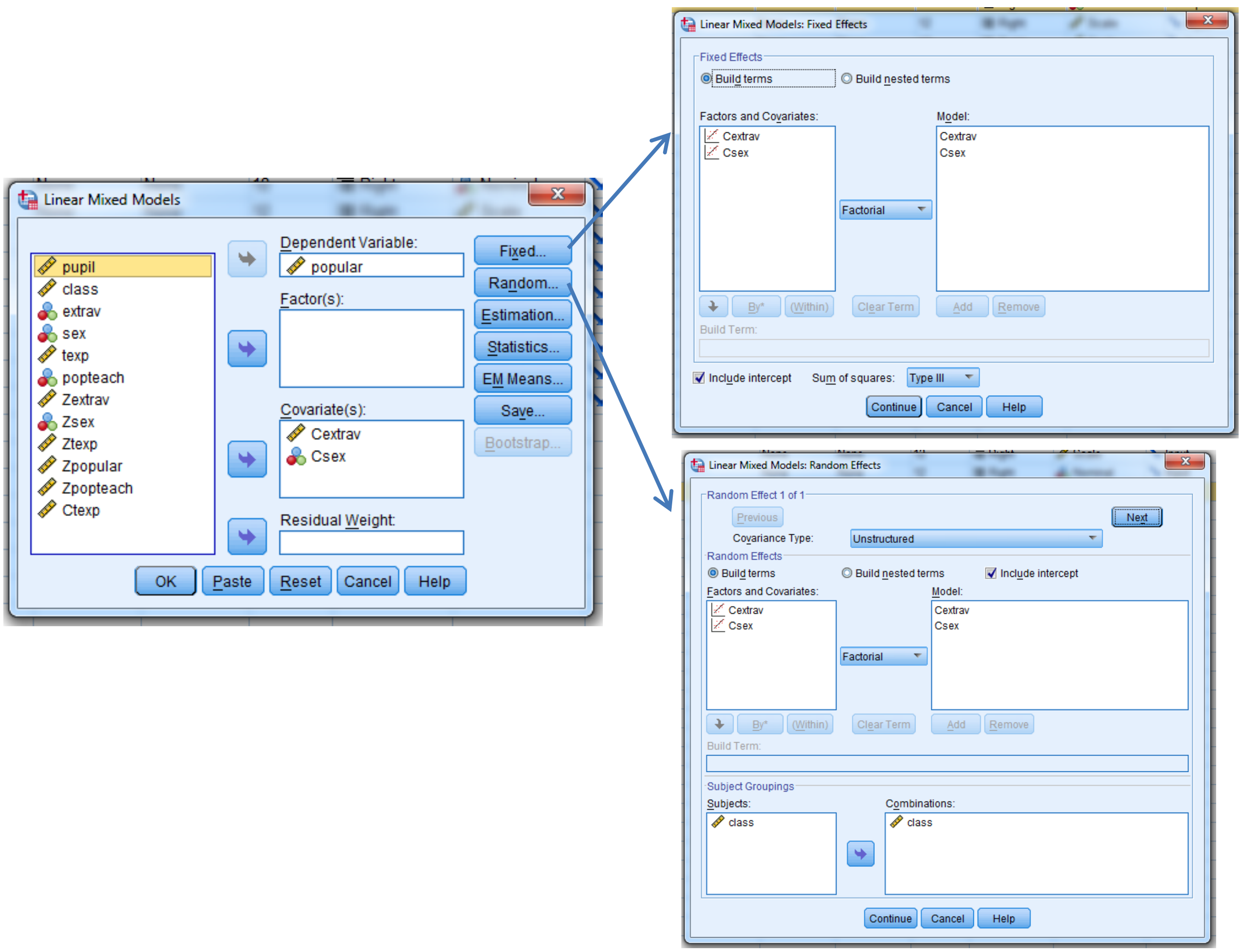
```
my_2fixedrnd <- lmer(popular ~ 1 + c_extrav + c_sex + (1 + c_extrav + c_sex |class))
p.values.lmer(my_2rnd)
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.021	0.841	59.7	<0.001
Extraversion (γ_{10})	0.443	0.023	18.9	<0.001
Sex (γ_{20})	1.245	0.037	33.4	<0.001

Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	0.553	0.744
Intercept (u_{0j})	0.674	0.821
Extraversion (u_{1j})	0.030	0.173
Sex (u_{2j})	0.005	0.073

SPSS Results

```
MIXED popular WITH Cextrav Csex
  /FIXED=INTERCEPT Cextrav Csex | SSTYPE(3)
  /METHOD=REML
  /PRINT=SOLUTION
  /RANDOM=INTERCEPT Cextrav Csex | SUBJECT(class) COVTYPE(UN).
```



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.021	0.083	60.6	<0.001
Extraversion (γ_{10})	0.442	0.023	18.8	<0.001
Sex (γ_{20})	1.246	0.038	33.0	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.553	0.019
Intercept (u_{0j})	0.654	0.097
Extraversion (u_{1j})	0.030	0.008
Sex (u_{2j})	0.008	0.022

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
           Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
           USEVARIABLES ARE class popular Cextrav Csex;
           WITHIN = Cextrav Csex;
           CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
         randoms1 | popular ON Cextrav;
         randoms2 | popular ON Csex;
         %BETWEEN%
         popular;
OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.031	0.083	60.6	<0.001
Extraversion (γ_{10})	0.441	0.024	18.3	<0.001
Sex (γ_{20})	1.254	0.036	35.3	<0.001

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	0.552	0.021	25.7	<0.001
Intercept (u_{0j})	0.647	0.083	7.6	<0.001
Extraversion (u_{1j})	0.031	0.008	4.1	<0.001
Sex (u_{2j})	0.005	0.024	0.2	0.8

This time we **included two latent slope variables in the WITHIN statement** to specify *Extrav* and *Sex* as a random factors. We can interpret the output for "randoms1" as the estimates for *Extrav* and "randoms2" as the estimates for *Sex.*

The output from Mplus for this model has estimates that are further away from the other programs that in the previous models. We see that as the model must estimate more random parameters, the difference in estimation procedures (*ML* vs. *REML*) become more apparent. However, Mplus agrees with the other programs that all estimates except the random variance component of *Sex* are highly significant.

Model Summary

For a random effect with a variance very close to zero, the six programs handle the estimate in different ways. SAS and Stata were unable to report the standard errors or p-values of the random effects, while the others had fairly different values for both the estimates and the standard errors. The Mplus results also show greater differences than in the previous models.

The ICC for this model is:

$$\rho = \frac{.654}{.654 + .553} = 0.542$$

Again, the **ICC has increased slightly as we add another student-level effect**, including a random slope, into the model.

One Level-2 Factor and Two Random Level-1 Factors (No Interactions)

<u>Mixed Model</u>

$$Popular_{ij} = \gamma_{00} + \gamma_{01}\boldsymbol{Texp_j} + \gamma_{10}\boldsymbol{Extrav_{ij}} + \gamma_{20}\boldsymbol{Sex_{ij}} + u_{1j}\boldsymbol{Extrav_{ij}} + u_{2j}\boldsymbol{Sex_{ij}} + u_{0j} + e_{ij}$$

<u>Hierarchical Model</u>

$$Popular_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{Extrav_{ij}} + \beta_{2j}\boldsymbol{Sex_{ij}} + e_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}\boldsymbol{Texp_j} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + u_{2j}$$

This is the first model we have seen that has a level-2 (class-level) variable: teacher's experience in years (*Texp*), which is also **grand mean centered**. As noted by Enders and Tofighi (2007), the only centering option for level-2 variables is grand mean centering. You cannot group mean center *Texp* because it is already measured at the class level, meaning the "group mean" would equal the original value.

In the hierarchical format, you can see that it has a fixed slope coefficient, γ_{01}, and is unique for every class *j*. This model does not have any interaction between teacher's experience and the student-level variables. We would use this model if we had reason to believe that *Texp* does not moderate the effects of *Sex* and *Extrav* on *Popular*, meaning that **the slopes for our student-level variables are the same whether or not the students have a teacher that is new or one that has many years of experience**.

SAS Results

```
proc mixed data=popdata covtest;
model popular = extrav_c sex_c texp_c /solution;
random intercept extrav_c sex_c /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.030	0.057	89.2	<0.001
Extraversion (γ_{10})	0.453	0.025	18.4	<0.001
Sex (γ_{20})	1.250	0.037	34.2	<0.001
Teach Experience (γ_{01})	0.089	0.009	10.4	<0.001

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	0.552	0.018	30.0	<0.001
Intercept (u_{0j})	0.285	0.046	6.2	<0.001
Extraversion (u_{1j})	0.034	0.009	4.0	<0.001
Sex (u_{2j})	0	-	-	-

We now see *Texp* in the fixed effects table, with an estimate of 0.089 and a significant p-value. This means that, holding the students gender and extraversion score constant, for every additional year's experience the teacher has, that student's *Popular* score increases by 0.089.

Again, we see that SAS can't handle the very small variation of the random gender effect. Therefore, no standard error, z-statistic, or p-value is reported.

Stata Results

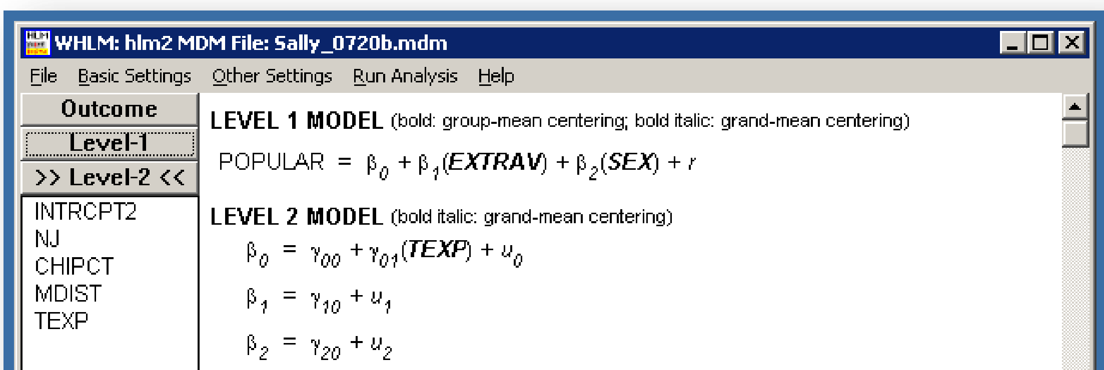
xtmixed popular ctexp cextrav csex || class: cextrav csex, variance cov(un) reml

Fixed Effects	*Estimate*	*St. Error*	*z-stat*	*p-value*
Intercept (γ_{00})	5.022	0.056	89.0	<0.001
Extraversion (γ_{10})	0.453	0.025	18.4	<0.001
Sex (γ_{20})	1.250	0.037	33.9	<0.001
Teach Experience (γ_{01})	0.090	0.009	10.4	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.551	.
Intercept (u_{0j})	0.285	.
Extraversion (u_{1j})	0.035	.
Sex (u_{2j})	0.002	.

As with the previous model, we got an error telling us that Stata could not calculate the standard errors of the variance components. However, these estimates roughly match those from the other programs.

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.033	0.056	90.1	<0.001
Extraversion (γ_{10})	0.453	0.025	18.5	<0.001
Sex (γ_{20})	1.251	0.035	35.8	<0.001
Teach Experience (γ_{01})	0.089	0.009	10.4	<0.001

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	0.551	0.742		
Intercept (u_{0j})	0.284	0.533	733.3	<0.001
Extraversion (u_{1j})	0.035	0.186	169.1	<0.001
Sex (u_{2j})	0.003	0.058	80.5	>0.5

These estimates differ slightly from the results of the other programs by very small amounts (in the hundredths or thousandths place).

R Results

```
my_model5 <- lmer(popular ~ c_extrav + c_sex + c_texp + (1 + c_extrav +
        c_sex|class))
p.values.lmer(my_model5)
```

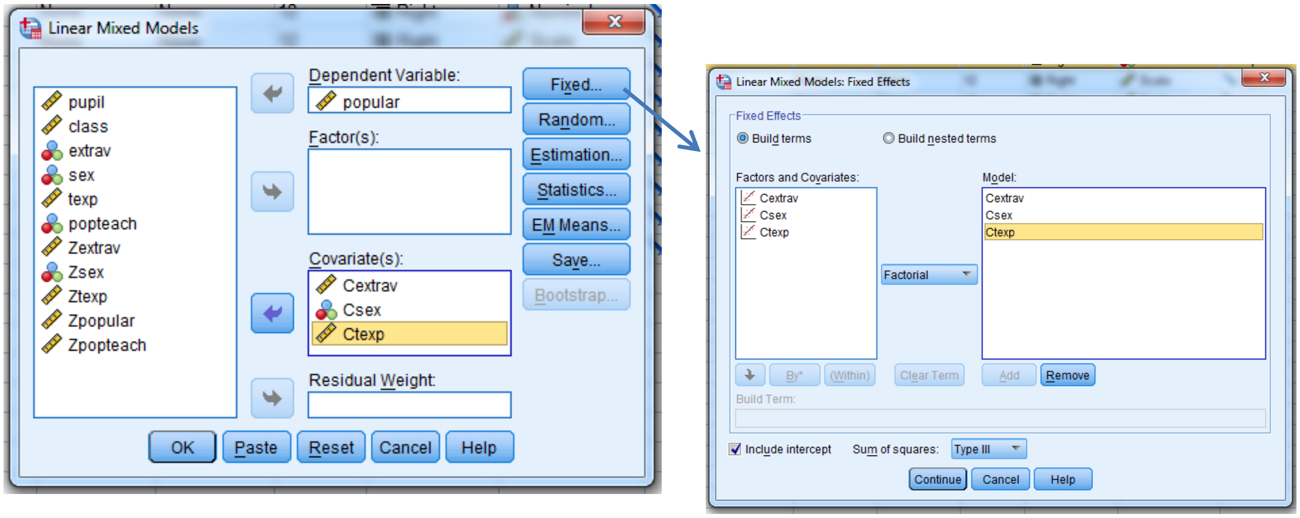
Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.022	0.056	89.0	<0.001
Extraversion (γ_{10})	0.453	0.025	18.4	<0.001
Sex (γ_{20})	1.251	0.037	33.9	<0.001
Teach Experience (γ_{01})	0.090	0.009	10.4	<0.001

Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	0.551	0.742
Intercept (u_{0j})	0.285	0.533
Extraversion (u_{1j})	0.035	0.186
Sex (u_{2j})	0.002	0.049

SPSS Results

```
MIXED popular WITH Cextrav Csex Ctexp
  /FIXED=INTERCEPT Cextrav Csex Ctexp | SSTYPE(3)
  /METHOD=REML
  /PRINT=SOLUTION
  /RANDOM=INTERCEPT Cextrav Csex | SUBJECT(class) COVTYPE(UN).
```

(See the screen shot below)



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.022	0.057	88.9	<0.001
Extraversion (γ_{10})	0.453	0.025	18.4	<0.001
Sex (γ_{20})	1.251	0.037	33.8	<0.001
Teach Experience (γ_{01})	0.089	0.009	10.3	<0.001

Variance Components	*Estimate*	*St. Error*
Residual (e_{ij})	0.548	0.018
Intercept (u_{0j})	0.285	0.046
Extraversion (u_{1j})	0.035	0.009
Sex (u_{2j})	0.004	0.000

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
           Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
           USEVARIABLES ARE class popular Cextrav Ctexp Csex;
           WITHIN = Cextrav Csex;
           BETWEEN = Ctexp;
           CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
         randoms1 | popular ON Cextrav;
```

```
            randoms2 | popular ON Csex;
            %BETWEEN%
            popular ON Ctexp;
    OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.020	0.056	90.1	<0.001
Extraversion (γ_{10})	0.452	0.024	18.5	<0.001
Sex (γ_{20})	1.254	0.035	35.3	<0.001
Teach Experience (γ_{01})	0.094	0.009	10.9	<0.001

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	0.551	0.021	26.0	<0.001
Intercept (u_{0j})	0.276	0.050	5.5	<0.001
Extraversion (u_{1j})	0.034	0.007	4.6	<0.001
Sex (u_{2j})	0.005	0.024	0.2	0.8

For the Level-2 factor, we include *Ctexp* in the BETWEEN statements. We again see small departures in these estimates from the output of the other five programs.

Model Summary

The outputs from the five programs that use *REML* are essentially equal, separated by only a few thousandths of a unit. As with the previous model, the biggest discrepancy occurs in the variance estimate of the random gender effect, since it is so close to zero.

Notice that the ICC for this model has **decreased** from the previous model (ρ = 0.542):

$$\rho = \frac{.285}{.285 + .551} = 0.341$$

Remember, the ICC is a measure of how much of the *unexplained* variation can be accounted for by which class you are in. **By adding a class-level predictor, we are accounting for a larger portion of the variation among the different classes**. Therefore, less variation exists in the random intercept, u_{0j}, for this model than those without any level-2 predictors, and thus the ICC is also lower.

One Level-2 Factor and Two Random Level-1 Factors with Interaction

Mixed Model

$$Popular_{ij} = \gamma_{00} + \gamma_{01}\boldsymbol{Texp_j} + \gamma_{10}\boldsymbol{Extrav_{ij}} + \gamma_{20}\boldsymbol{Sex_{ij}} + \gamma_{11}\boldsymbol{Texp_j} * \boldsymbol{Extrav_{ij}} + \gamma_{21}\boldsymbol{Texp_j} * \boldsymbol{Sex_{ij}} + u_{1j}\boldsymbol{Extrav_{ij}} + u_{2j}\boldsymbol{Sex_{ij}} + u_{0j} + e_{ij}$$

Hierarchical Model

$$Popular_{ij} = \beta_{0j} + \beta_{1j}\boldsymbol{Extrav_{ij}} + \beta_{2j}\boldsymbol{Sex_{ij}} + e_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}\boldsymbol{Texp_j} + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}\boldsymbol{Texp_j} + u_{1j}$$
$$\beta_{2j} = \gamma_{20} + \gamma_{21}\boldsymbol{Texp_j} + u_{2j}$$

This is the only model in which we have cross-level interactions between the class-level variable, *Texp*, and both student-level variables, *Sex* and *Extrav*. We would use this model if, for instance, we wanted to find out if teachers with more experience have a different impact on the relationship between student's extraversion or gender and their self-reported popularity than newer teachers. In other words, **does teacher's experience moderate the effect of extraversion or gender on popularity?**

You can see that in the hierarchical format, *Texp* has a slope coefficient within each of the three β equations. This relates to the interaction terms in the mixed model for teacher's experience by extraversion and well as teacher's experience by gender.

SAS Results

```
proc mixed data=popdata covtest;
model popular = extrav_c sex_c texp_c extrav_c*texp_c
        sex_c*texp_c /solution;
random intercept extrav_c sex_c /subject=class type=un;
run;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.000	0.057	88.2	<0.001
Extraversion (γ_{10})	0.451	0.018	25.8	<0.001
Sex (γ_{20})	1.240	0.036	34.2	<0.001
Teach Experience (γ_{01})	0.097	0.009	11.2	<0.001
Texp*Extrav (γ_{11})	-0.025	0.003	-9.6	<0.001
Texp*Sex (γ_{21})	-0.002	0.006	-0.3	0.770

Variance Components	*Estimate*	*St. Error*	*z-stat*	*p-value*
Residual (e_{ij})	0.553	0.018	30.0	<0.001
Intercept (u_{0j})	0.287	0.045	6.3	<0.001
Extraversion (u_{1j})	0.006	0.005	1.26	0.104
Sex (u_{2j})	0	-	-	-

In the fixed effects table, there are two interaction terms, one of which (γ_{21}) is far from significant, with a p-value of >0.5. However, $\gamma_{11}$is significant, meaning that teacher's experience moderates the relationship between *Extrav* and *Popular*, but not the relationship between *Sex* and *Popular*.

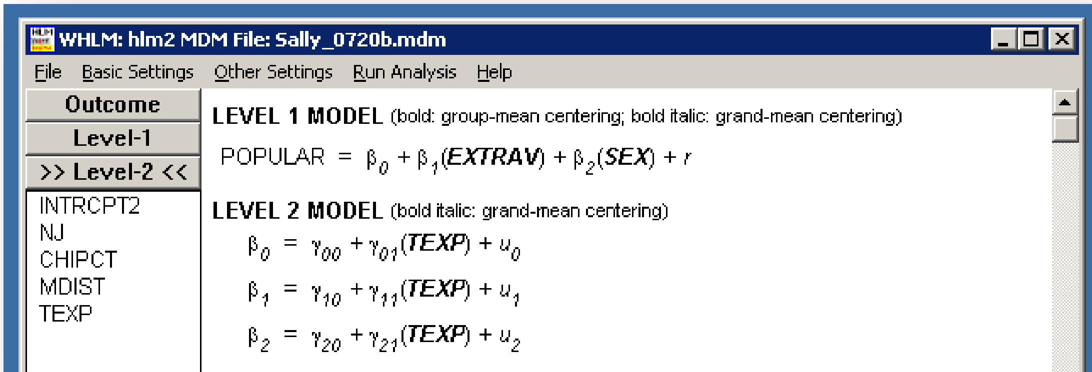
In the random variance components table, we see that the estimates for the extraversion random slope, u_{1j}, and the sex random slope, u_{2j}, are not significantly different from zero. This means that there is no evidence to suggest that these two factors actually vary by class in this model.

Stata Results

```
gen texp_extrav = ctexp*cextrav
gen texp_sex = ctexp*csex
xtmixed popular ctexp cextrav csex texp_extrav texp_sex || class: cextrav csex,
        variance cov(un) reml
```

Stata does not have the capability to recognize interaction terms between variables automatically, so we must manually create variables for both of our cross-level interactions (see the *gen* statements in the above code). When we ran this *xtmixed* command with the unstructured covariance matrix option, Stata gave an error saying **Hessian is not negative semidefinite, conformability error** and produced no output. The code would run without the *cov(un)*, but because we are comparing the outputs among the six programs, we do not list that output here.

HLM Results



Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	5.002	0.056	88.2	<0.001
Extraversion (γ_{10})	0.450	0.017	26.2	<0.001
Sex (γ_{20})	1.240	0.036	34.8	<0.001
Teach Experience (γ_{01})	0.097	0.009	11.2	<0.001
Texp*Extrav (γ_{11})	-0.025	0.002	-10.3	<0.001
Texp*Sex (γ_{21})	-0.002	0.006	-0.3	0.762

Variance Components	*Estimate*	*St. Dev.**	*Chi-square*	*p-value*
Residual (e_{ij})	0.552	0.743		
Intercept (u_{0j})	0.286	0.535	743.5	<0.001
Extraversion (u_{1j})	0.006	0.075	97.7	0.182
Sex (u_{2j})	0.006	0.076	80.4	>0.500

These estimates roughly equal the results from the other programs.

R Results

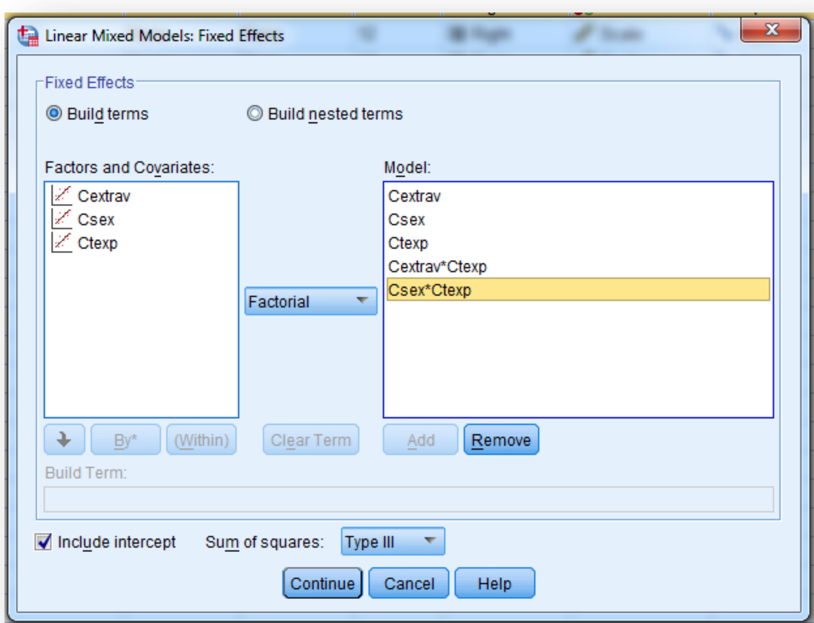
```
my_model6 <- lmer(popular ~ c_extrav + c_sex + c_texp + c_extrav*c_texp +
                c_sex*c_texp + (1 + c_extrav + c_sex|class))
p.values.lmer(my_model6)
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	4.991	0.058	88.0	<0.001
Extraversion (γ_{10})	0.450	0.017	25.8	<0.001
Sex (γ_{20})	1.240	0.037	33.7	<0.001
Teach Experience (γ_{01})	0.097	0.009	11.2	<0.001
Texp*Extrav (γ_{11})	-0.025	0.002	-9.6	<0.001
Texp*Sex (γ_{21})	-0.002	0.006	-0.3	0.766

Variance Components	*Estimate*	*St. Dev.**
Residual (e_{ij})	0.552	0.743
Intercept (u_{0j})	0.287	0.536
Extraversion (u_{1j})	0.006	0.075
Sex (u_{2j})	0.004	0.064

SPSS Results

```
MIXED popular BY Cextrav Csex Ctexp
  /FIXED=INTERCEPT Cextrav Csex Ctexp Cextrav*Ctexp Csex*Ctexp | SSTYPE(3)
  /METHOD=REML
  /PRINT=SOLUTION
  /RANDOM=INTERCEPT Cextrav Csex | SUBJECT(class) COVTYPE(UN).
```



This model was too much for SPSS 19 to handle. It is possible that for more complicated models with an unstructured covariance matrix, the other programs run a more efficient algorithm and therefore are preferred over SPSS.

Mplus Results

```
TITLE:  HLM Popular Data - Unconditional Model
DATA:  FILE IS C:\popular_mplus.csv;
VARIABLE:  NAMES ARE pupil class extrav sex texp popular popteach Zextrav
            Zsex Ztexp Zpopular Zpopteach Cextrav Ctexp Csex;
            USEVARIABLES ARE class popular Cextrav Ctexp Csex;
            WITHIN = Cextrav Csex;
            BETWEEN = Ctexp;
            CLUSTER = class;
ANALYSIS:  TYPE = twolevel random;
MODEL:   %WITHIN%
          randoms1 | popular ON Cextrav;
          randoms2 | popular ON Csex;
          %BETWEEN%
          randoms1 ON Ctexp;
          randoms2 ON Ctexp;
OUTPUT: sampstat;
```

Fixed Effects	*Estimate*	*St. Error*	*t-stat*	*p-value*
Intercept (γ_{00})	4.989	0.056	89.1	<0.001
Extraversion (γ_{10})	0.045	0.017	25.9	<0.001
Sex (γ_{20})	1.242	0.036	34.4	<0.001
Teach Experience (γ_{01})	0.097	0.009	11.3	<0.001
Texp*Extrav (γ_{11})	-0.025	0.002	-10.2	<0.001
Texp*Sex (γ_{21})	-0.001	0.006	-0.2	0.8

Variance Components	*Estimate*	*St. Error*	*t-stat*	*p-value*
Residual (e_{ij})	0.551	0.021	25.9	<0.001
Intercept (u_{0j})	0.279	0.049	5.7	<0.001
Extraversion (u_{1j})	0.005	0.004	1.1	0.3
Sex (u_{2j})	0.007	0.022	0.3	0.8

We now include two ON statements in the BETWEEN model section to indicate the cross-level interactions with teacher's experience. Again, we see minor discrepancies with the other outputs, but Mplus agrees that the fixed interaction between *Texp* and *Sex* is not significant, as well as the random components for *Extrav* and *Sex*.

Model Summary

With the addition of two cross-level interaction terms, Stata and SPSS were unable to run the model with an unstructured covariance option. That is not to say that they shouldn't be used for this type of analysis, but some caution should be used when adding more complicated parameters to an model with an unstructured covariance matrix.

As with previous models, the results from SAS, HLM, and R are relatively close to being equal, while the Mplus estimates differ slightly. Also, the ICC is nearly exactly the same as with Model 5, meaning that the interaction terms did not change the proportion of variance accounted for by class:

$$\rho = \frac{.287}{.287 + .552} = 0.342$$

Overall Summary

The purpose of this comparison was to investigate the possible differences in procedures and results for a nested two-level hierarchical model from six different statistical software programs. Overall, we have found that there is not much difference in the actual estimates produced by SAS, Stata (with the *reml* option), HLM, R, and SPSS. **Mplus uses a different method of estimation, *ML*, which causes its estimates to differ somewhat from the others**. In addition, it is important to note the following:

1. For random effects with a variance estimate very close to zero, SAS was unable to produce standard errors or p-values. The other three programs differed in their estimates for these parameters to a greater extent that for other effects.
2. Stata and SPSS were unable to handle the most complicated model, which contained two cross-level interactions. The other programs are recommended for analyses dealing with complicated models and specifying an unstructured covariance matrix.

Additionally, we investigated the value of ρ, the intra-class correlation coefficient, in each model. By adding level-1 predictors, the ICC increased. However, when we **added a level-2 predictor, the ICC dramatically decreased to an even lower value than the unconditional**

model. This is due to a decrease in the unexplained Level-2 variation, the random intercept term u_{0j}, when a predictor was added at the class level.

Although this document can be used as a guide for running various two-level hierarchical models for nested datasets, we strongly urge readers to only use these models when they are appropriate for answering your specific research questions. Caution must be used when deciding between fixed and random factors, and with whether to grand mean or group mean center level-1 factors. For further information on multilevel modeling, we recommend Hox's book, referenced below.

If you have questions about model selection, appropriate uses of mixed models, or interpretation of the results from any of these programs, schedule an appointment to meet with one of the consultants in the Division of Statistics and Scientific Computation at:

http://ssc.utexas.edu/consulting/free-consulting

References

Albright, Jeremy J. and Marinova, Dani M. (2010). "Estimating Multilevel Models using SPSS, Stata, SAS, and R" [PDF document]. Retrieved from http://www.indiana.edu/~statmath/stat/all/hlm/hlm.pdf .

Enders, Craig K. and Tofighi, Davood (2007). "Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue." *Psychological Methods*, vol. 12, pg. 121-138.

Hox, Joop J. (2010). *Multilevel Analysis* (2nd ed.). New York: Routledge.