Hierarchical Linear Modeling: A Primer[1]
(People within Groups)

R. C. Gardner
Department of Psychology

As discussed by Raudenbush and Bryk (2002, pp 5-6), Hierarchical Linear Modeling (HLM) has various names in different disciplines, *multilevel linear models* in sociology, *mixed-effects models* and *random-effects models* in biometrics, *random-coefficient regression models* in econometrics, and *covariance components models* in the statistical literature. Tabachnick and Fidell (2007) refer to it as *multilevel linear modeling*. There are also many computer programs that permit you to do the analysis, such as HLM (see Raudenbush, Cheong, Bryk, Congdon, & du Toit, 2004), SPSS Mixed Models, and SAS Proc Mixed. These programs offer different types of output (sometimes yielding different answers because of the different algorithms or tests of significance used) and use at least two different analytic approaches (REML (REstricted Maximum Likelhood, sometimes labelled MLR) and ML (Maximum Likelihood, sometimes labelled MLF)). Finally, they are applicable to data derived from sampling from independent observations or from repeated measures. This overview directs attention to data based on independent observations and focusses initially on the simplest form of HLM, the random-coefficients model. This involves a continuous dependent variable (e.g., school marks), a continuous independent variable (e.g., a measure of motivation), and a categorical independent variable (e.g., classes of students). Obviously, hierarchical linear modeling can be extended to include more variables in the random-coefficients model as well as in more complex models (e.g., an intercepts- and slopes-as-outcomes model). These extensions are discussed and examples of their analysis using HLM are presented.

Purpose and Rationale

In essence, the question being asked in the simplest form of the random-coefficients model is whether there is an association between the dependent variable and the continuous independent variable and whether the categorical independent variable influences this relationship. One way of answering this type of question is to perform a two-factor analysis of variance of the dependent variable with the two independent variables as factors (cf., Gardner, 2003). Using multiple regression where the categorical factor (classes) is effect coded, the continuous independent variable is centered, and product terms are made of the effect coded vectors and the centered continuous variable, tests of significance of the increment in R^2 over the other two sources of variance would test relevant hypotheses. Thus, a test of the increment in R^2 over the other two sources of variance tests hypotheses concerning variation of the intercepts over the different classes (Main effect for classes), variation of the slopes over classes (Interaction effect of continuous and categorical factors), and difference of the mean slope from

[1]Preparation of this manuscript was facilitated by grant number 410-2002-0810 from the Social Science and Humanities Research Council of Canada. I would like to express my appreciation to Roger Covin for his invaluable assistance in the initial phase of its preparation.

0 (Main effect for the continuous independent variable). A test of significance of the unstandardized regression coefficient for the constant is a test of whether the mean intercept differs significantly from 0. Hierarchical linear modeling tests comparable hypotheses, except that a slightly different approach is taken.[2]

The reason for performing an hierarchical linear modeling analysis is to test hypotheses about intercepts and slopes. For the simple example just described, four null hypotheses can be tested. One is that the mean intercept (across all groups) is 0; a second is that the mean slope (across all groups) is 0; the third is that the variance of the intercepts (over all groups) is 0; and the fourth is that the variance of the slopes (over all groups) is 0. Unless the dependent variable has a mean of 0 in the population, the first null hypothesis is of little interest. The others have psychological meaning, however. Thus, if the mean slope is found to differ significantly from 0, this indicates that there tends to be an average positive or negative relationship between the dependent variable and the continuous independent variable in the population (cf., the Main effect for the continuous independent variable discussed above). If the variance of the intercepts is significantly greater than 0 (note, this is a one tailed test), this would indicate that some groups tend to have higher intercepts (on the dependent variable) than others (cf., the Main effect for classes discussed above). Finally, if the variance of the slopes is significantly greater than 0 (again a one tailed test), this would indicate that the slopes differ in the various groups (cf., a test of the Interaction effects). As described above, this is akin to an analysis of variance where one factor is categorical and the other is continuous. Generally, in HLM, the categorical factor is random because the groups represent random, or at least representational sampling of a population of possibilities.

One might well ask what is the difference between a two factor analysis of variance with one continuous factor and a random-coefficients model in hierarchical linear modeling with one continuous predictor and subjects nested in groups. The purpose of this manuscript is to explain this difference. One difference between the analysis of variance approach and hierarchical linear modeling is that the former makes use of least squares estimates of effects, while the other uses maximum likelihood or restricted maximum likelihood. This is not the major distinction, but an interesting aspect of it is that with least squares, there are formulae that one can use to calculate the required statistics (i.e., sample values) which are then used to estimate the corresponding parameters. With maximum likelihood (and restricted maximum likelihood) in hierarchical linear modeling, there are no formulae to compute group statistics, but rather the solution makes direct estimates of the corresponding parameters and where applicable also the estimates of sampling error. The maximum likelihood procedures are iterative, beginning with the ordinary least squares estimates, and continually estimating the corresponding parameters (including the error terms) until the data fit the proposed model as closely as possible. It will be noted in many runs, that the fit of the model isn't always achieved, but that the fit approaches an asymptote at

[2] In the analysis of variance paradigm, the appropriate *F*-ratios would depend on the nature of the categorical factor and can be determined by considering the expected mean squares for the model. The continuous variable would be considered fixed by definition (because regression involves a fixed independent variable), but the categorical variable could be fixed or random depending on how it is formed.

some value, and the decision is made that this is an acceptable estimate of fit. In any event, none of the values estimated with the maximum likelihood approaches can be linked directly with the individual group statistics (such as those obtained with the least squares approach). Thus, the major distinction between analysis of variance and hierarchical linear modeling is in the nature of the models underlying the estimates of the parameters, their estimates of error, and the tests of significance..

The Random-Coefficients model

For this type of data, hierarchical linear modeling considers two levels of models. The model at level 1 is concerned with the regression of the dependent variable (school marks) on the continuous independent variable (motivation) within each group. The formula given by Raudenbush and Bryk (2002, p 19) is written as:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \overline{X}_{.j}) + r_{ij} \qquad (1)$$

In this equation, the β's refer to population values of *unstandardized* regression coefficients, where β_{0j} is the intercept and β_{1j} is the slope for group *j*, and r_{ij} is the error in predicting Y_{ij} with this equation. A more familiar form of the same equation for any one group is:

$$Y_i = b_0 + b_1 X_i + e_i \qquad (2)$$

That is the concern is with predicting values of the dependent variable from a knowledge of the independent variable within each group. In addition to the differences in notation, the first formula multiplies the slope by the deviation of each *X* from the mean of *X* in that group. The second formula does not subtract the sample mean. In any event, the value of b_1 would be the same as that for β_{1j}, but the values of b_0 and β_{0j} would differ. There is no loss in generality.

For the Random-Coefficients model to be discussed here, the model at level 2 is concerned with considering the values of β_{0j} and β_{1j} further from a knowledge of group membership. The formulae given by Raudenbush and Bryk (2002, p. 26) are:

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \qquad (3a)$$

where:

γ_{00} = mean intercept over the groups
μ_{0j} = deviation of the intercept for group j from the mean intercept (i.e., $\beta_{0j} - \gamma_{00}$)

and

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \qquad (3b)$$

where:

γ_{10} = mean slope over the groups
μ_{1j} = deviation of the slope for group j from the mean slope (i.e., $\beta_{1j} - \gamma_{10}$)

Note equations 3a and 3b are in essence tautological. If one were to expand these two equations based on their defining terms, they would find that $\beta_{0j} = \beta_{0j}$, and $\beta_{1j} = \beta_{1j}$. However, by substituting their defined terms into equation 1, we now define the original value of Y_{ij} in terms of five parametric values instead of the three involved in the direct least squares approach used in analysis of variance. That is, substituting the values for equations (3a) and (3b) into equation (1) yields a more complex form of equation (1) as follows in equation (4):

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{.j}) + \mu_{0j} + \mu_{1j}(X_{ij} \;\; \bar{X}_{.j}) + r_{ij}$$

Note this equation indicates that the obtained Y_{ij} is a function of four parameters, the mean intercept (γ_{00}) and the mean slope (γ_{10}) (both fixed parameters), plus two random parameters, variation of the intercepts (μ_{oj}), and variation of the slopes (μ_{1j}). The residual variation, (r_{ij}) is what is left over after these four parameters are estimated.

The difference between analysis of variance and hierarchical linear modeling lies in adding the level 2 equations when estimating the slopes and intercepts for each group. Using only equation (1) (which is the case if one uses multiple regression to perform an analysis of variance of these data) will yield the actual intercepts and slopes computed on the data for each group, and define the variance due to error in terms of the residual. Using equation (4) in hierarchical linear modeling will yield maximum likelihood estimates of the group intercepts and slopes taking into account the variation of the slopes and intercepts over the groups. The two sets of answers will not be the same. Note that hierarchical linear modeling is concerned with estimating the unique contribution of the various parameters in the final equation. Thus, the term “hierarchical” refers to the fact that the slopes and intercepts are considered from the two points of view, level 1 and level 2. It is not hierarchical in the same sense as hierarchical in multiple regression, where various predictors are entered in different steps. In hierarchical linear modeling, the equation is defined from the two points of view, but the final equation is solved in terms of the unique contribution of each parameter (i.e., SSTYPE III).

Running HLM

To analyze data using the program HLM, consider the following example, where:
(a) the dependent variable is GRS (grades in English)
(b) the continuous independent variable is FMOT (Motivation),
(c) the subjects are nested in classes (ENV). Because there is no variable describing differences between groups on some attribute, other than the group itself, this is referred to as a Random-Coefficients Regression model.

The example given assumes the data are in the SPSS data format (.SAV). The analysis proceeds as follows:

1. Enter HLM

2. Click on ***File*** in the tool bar. To create a new file, click on ***Make New MDM file***, and then select ***Stats Package Input***.

3. This presents a window. Select ***HLM2*** if it is not already selected. Click on ***OK***.

4. This presents another window (labelled ***MAKE MDM - HLM2***). In this window:

a. Type in a file name with the extension ***.MDM*** in the pane at the upper right. This is the file that will contain the instructions for this run. If you save it as suggested later in the sequence, it will be available for future runs that might be edited, etc.

b. Click the drop down window for ***Input File Type*** and select ***SPSS/Windows***. There is a section indicating the type of data nesting. For the type of model considered here, the **Nesting of Input Data** should have **persons within groups** indicated.,

c. Click ***Browse*** for Level 1**.**

5. This will open another window labelled **Choose Level-1 File** that presents a file source (e.g., C:\, H:\, E:\, etc, depending on your configuration). At the bottom of this window, **Files of Type** should indicate ***SPSS/Windows files [*.SAV]***. Click on **Open**.

6. This will return you to the window **Make MDM-HLM2**. Click on **Choose variables** (for Level-1 Specification).

7. This will produce another window labelled **Choose variables-HLM2.** This window lists all the variables in your file with two columns of boxes next to them. One is labelled **ID**, and the other **in MDM**. Put a check in the box for **ID** for the variable which indicates which group the subjects are in, and checks in the boxes in the **in MDM** columns for any variable that you might wish to consider as a predictor at level 1. Click on **OK**.

8. This will return you to the window, **Make MDM-HLM2**. This window also allows you to indicate whether or not you have missing data. The default checked is **no missing data**. If you have missing data, indicate this and then click on either delete missing data when **making MDM** or on **running analyses**. Click on **Browse** for Level 2 Specification.

9. This will produce another window labelled **Choose Level-2 File** that presents a file source (e.g., C:\, H:\, E:\, etc, depending on your configuration). At the bottom of this window click the drop down arrow for **Files of Type** and select ***SPSS Windows files [.SAV*]**. Select the appropriate folder and the data file from this folder. Note this could be the same file as in 5 above, as long as any group variables are indicated for each subject. In this example, the only group variable is class number (ENV), hence we can read in the same file again for this purpose. That is, in many cases the file of interest is the existing data file. If this is the case for you, select that file again and click **OK**. In some instances, a researcher might describe the groups in a separate file with columns for the relevant group variable and one row for each group. In this case, one of the columns must have the same values (and label) as that indicating groups in the original data file, but then the other variables can be group-defined variables. If you use this option, then at this point you would select this file, and click **Open**. This will return you to the **Make MDM-HLM2** window. Click on **Choose Variables** (for Level-2 Specification).

10. In any event, this will reintroduce the window **Choose variables-HLM2**, which lists all the

variables in this file with the two columns, **ID**, and **in MDM** after each one. You are to put a check in the box for **ID** for the variable which indicates which group the subjects are in, and checks in the boxes in the **in MDM** columns for any level 2 variable that you might wish to consider as a predictor at level 2. Note there must be at least one variable checked in the MDM column, even if it is not used later, so you must put a check in a column for at least one more variable (one of the previous ones is fine if you selected the same file as before). Then click on **OK**. This returns you to the window, **Make MDM - HLM2**.

11. Click on **Save MDMT file** and save the file you just created in a folder of your choice**.** The computer presents a window to save files. If the default file source is not the one you want, click on the down arrow in the **Save In** window, select the source and folder of interest, click on the **File Name** pane and type in the file name. Click on **Save**. This returns you to the **Make MDM-HLM2** window.

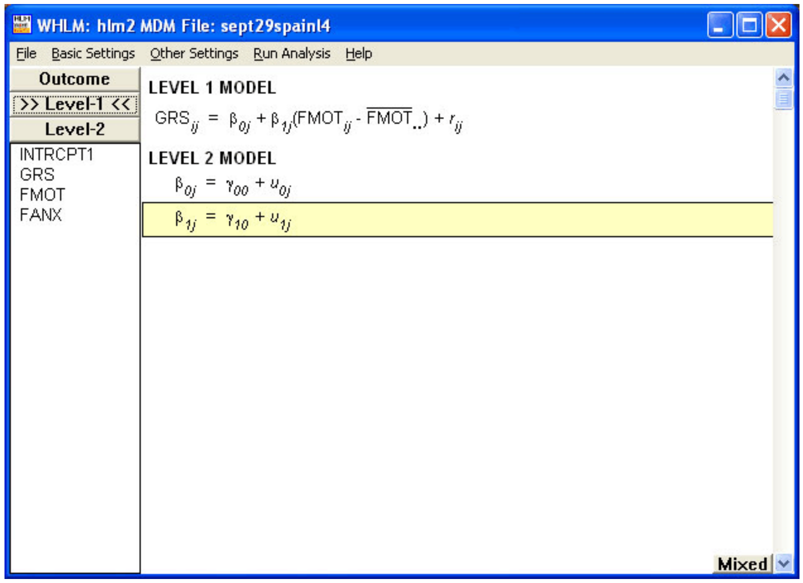
12. Click on ***Make MDM***. This presents a black screen with white writing showing the descriptive statistics for all the variables chosen.

13. Click on ***Done***. The computer advizes you to check the statistics. Click on **OK** and then on **Done** again. This leads to the next window.

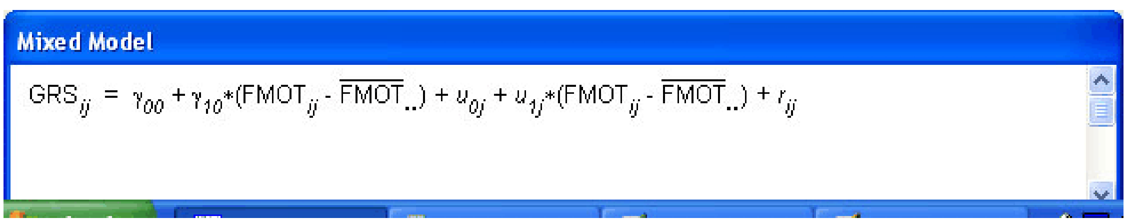
14. This presents a blank window in which you can build the model. First, click on **Level 1**, then click on the dependent variable (GRS in our example) and then on **Outcome variable**. This will produce the beginning of the Level 1 equation. Next click on the continuous independent variable (FMOT in our example), and then on **add variable grand mean centered**. There are actually three choices you can make here, uncentered, group mean centered, and grand mean centered (see Raudenbush & Bryk, 2002, pp. 31-35).

15. Click on Level 2, and build the Level 2 model. Our interest here is simply to study the effects of groups, thus there are no variables to add. At this point, there are two equations at Level 2. It is possible that one of the terms, µij is presented in a faded manner in the second equation. If so, click on that term (to toggle it), and it will become darker. This will cause the estimate of the variance in the group slopes to be estimated, which you will want in order to determine whether the within group slopes differ. If you do not want to estimate these variances, click on it to toggle it off (it becomes a lighter font).

Following is a window showing the final form of the equations.



16. To view the final equation (i.e., the full model), click on ***Mixed*** in the lower right hand side of the window. This will present the following window. This is the equation containing all the parameters that are being estimated. Note in this example that, except for differences in notation, it is identical to equation 4. It shows FMOT to be grand mean centered because that is what we selected at step 14.



17. HLM does not output the maximum likelihood regression coefficients or the ordinary least squares ones by default. If the maximum likelihood estimates are desired, an **SPSS.SAV** file with the estimates (and other estimates) can be obtained by clicking on **Basic Settings** in the tool bar. This will produce another window. Click on **Level 2 Residual File**. This produces another window that shows the dependent variable in a pane labelled **Possible choices**. Double click on the dependent variable and it will move it to the pane **Variables in the Residual file**. Note that at the bottom, it indicates the file (e.g., **resfil2.sav**).(In more complex models there may other variables to move. Also, if you wanted the Level 1 residual file, you could obtain this by following comparable instructions after having checked **Level 1 Residual File**). Click on **OK**.

18. To have the ordinary least squares regression coefficients printed as part of the output, Click on **Other Settings** in the tool bar and when a new window is presented turn off **Reduced Output** in that window. This will result in the OLS regression coefficients being presented in the output file along with other intermediate results. The maximum likelihood results will be the

last stage presented (there are intermediate results before them).

19. To choose the type of analysis (i.e., REML or ML), click on **Other Settings** in the tool bar, and then on **Estimation Settings**. The default setting is Restricted Maximum Likelihood (REML), which we will use in this example, but you could choose Maximum Likelihood (ML) if you prefer. The answers differ as well as some of the output depending on which one you select.

20. When the model is ready, click on ***Run Analysis***. If you have not saved the run, the computer will remind you of this. If you choose **Save and Run**, it will produce another window, **Save Command File**. Type in the file name and it will save it.

21. To view the output, click ***File*** and ***View Output***.

The Output

A data set of 136 participants distributed in 6 classes was analyzed following these directions. The dependent variable was GRS (grades in English at the end of the school year), the continuous independent variable was Motivation assessed at the beginning of the year (FMOT), and Classes were identified as ENV. The values of the six relevant statistics calculated using REML are as follows:

1. Mean Intercept	55.433175
2. Mean Slope	2.866550
3. Variance of intercepts	41.33254
4. Variance of slopes	2.01738
5. Covariance of slopes and intercepts	3.26090
6. Variance of Residuals at level 1	213.22126

A sample of some of the output including tests of significance is printed below.

The sample size, intercept, and slope for each of the six classes are presented below. Both the OLS and the REML estimates are given. If the analysis was performed using ML, the maximum likelihood values would differ. These values can be obtained in HLM (see items 17 and 18 above), but note, that the OLS values are simply those that you would obtain if you were to compute them separately for each group using the bivariate regression program in SPSS or any other package.

		Ordinary Least Squares (OLS)		Restricted Maximum Likelihood (REML)	
Group	N	Intercept	Slope	Intercept	Slope
1	22	56.065	1.400	55.994	1.829
2	21	50.967	0.682	53.328	1.486
4	21	48.820	5.710	51.838	3.820
5	23	57.821	4.102	57.036	3.617
6	25	47.912	4.392	48.851	4.153

7	24	67.816	2.470	65.550	2.293
Mean		54.900	3.126	55.433	2.866
Variance		55.5959	3.7240	33.183	1.288

Inspection of the OLS and REML estimates for each of the 6 classes will reveal that there are differences in the corresponding individual values in the two sets, though the rank order of both the intercepts and the slopes are similar. The rank order correlation is 1.0 for the intercepts and .94 for the slopes. It will be noted too that the means of the two sets of coefficients are similar, but that the variances are much lower for the REML estimates than the OLS ones. (If the ML solution were used, the answers would be slightly different).

The output from HLM contains the following information. Some explanatory comments are made to identify the nature of the information.

The tau matrix is a covariance matrix for the group intercepts and slopes. The diagonal values are the estimates of the population variances of the intercepts and slopes respectively, while the value in the off-diagonal is the estimate of the covariance of the intercepts and slopes.[3] This matrix is also presented in standard score form and hence is the correlation matrix obtained from the covariance matrix by dividing by the product of the associated standard deviations. Following are the two matrices:

Tau Matrix

INTRCPT1, B0	41.33254	-3.26090
FMOT, B1	-3.26090	2.01738

Tau (as correlations)

INTRCPT1, B0	1.000	-0.357
FMOT, B1	-0.357	1.000

The next two tables are tests of the mean intercept and slope across the various groups and the tests of the variances of the intercepts and slopes for the various groups respectively. The first table is identified as tests of the final estimation of fixed effects. In this case, there is one test of the null hypothesis that the mean intercept is 0, and another that the mean slope is 0. These are single sample *t*-tests with degrees of freedom equal to the number of groups minus 1. Note too that the values given for the Coefficient are the mean intercept and slope as given in the summary of the REML values presented above; the values of Standard Error, however, are also

[3] In the full maximum likelihood solution, this is followed by a matrix of standard errors but these are not given in the REML solution. In other programs such as SAS Proc Mixed and SPSS Mixed Models, these standard errors are used to test the significance of these variances and covariances, but Raudenbush and Bryk (2002, p. 64) indicate that the approximation provided by this test can be extremely poor, and recommend and actually make use of another test of these hypotheses in the HLM program (in both the REML and ML solutions).

REML estimates and are not the square roots of the variances shown above.

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	d.f.	Approx. P-value
For INTRCPT1, B0					
INTRCPT2, G00	55.433175	2.928026	18.932	5	0.000
For FMOT slope, B1					
INTRCPT2, G10	2.866550	0.727615	3.940	5	0.015

The second table presents tests of variability of the intercepts and slopes over the classes. These are considered random effects in that classes are considered a random factor, and the generalization of interest is whether these variances are greater than 0 in the population of all possible classes. Note that the variance components are in fact the diagonal values in the tau matrix presented above. The tests of significance in this case make use of a Chi-square statistic with degrees of freedom equal to the number of classes minus 1.

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	6.42904	41.33254	5	28.67176	0.000
FMOT slope, U1	1.42034	2.01738	5	16.17150	0.007
level-1, R	14.60210	213.22126			

Interpretation

Interpretation of these data would focus on the two tables given immediately above, Final estimation of fixed effects, and Final estimation of variance components. In the first table, it is shown that the mean intercept is 55.43 (note this is in fact the mean of the REML intercepts shown in the table of the regression coefficients presented earlier) and that it differs significantly from 0. This is of little interest, however, because all the values of the dependent variable were greater than 0. It is akin to the test of the constant in traditional analysis of variance, which would only have meaning if the dependent variable could take positive and negative values. Second, the mean slope is shown to be 2.87 (i.e., the mean of the REML slopes given in the table presented earlier), and it too differs significantly from 0 ($p<.015$). This indicates that the mean slope across the different classes is significantly positive. That is, on average, there is a positive association between grades and motivation across the classes. These two tests of significance are performed using *t*-tests with standard errors and degrees of freedom based on the REML estimates. The output in the run itself (but not presented here) also shows tests using robust standard errors, but then warns that these data do not satisfy the conditions necessary to use these tests.

The second set of output of interest presents the final estimation of variance components. Three estimates are shown. The first is the variance of the intercepts across the classes estimated at 41.33, and tested for significance using a chi-square statistic. It is significant, indicating that the intercepts vary across the six classes. The intercept for any given group is that value when the continuous independent variable is 0 (the grand mean when it is centered), thus this can be interpreted as indicating that these measures of central tendency (adjusted means) for the various group differ more than can be reasonably attributed to chance. As can be seen in the estimates given earlier, the lowest intercept was obtained in class 6 (48.85) and the highest in class 7 (65.55). The second is the variance of the slopes, estimated at 2.02, and tested for significance using a chi-square test. It too is significant, indicating that the slopes in the various groups differ more than one could reasonably attribute to chance. That is, for some groups the slopes are larger than for others. A close examination of the REML estimated slopes presented earlier shows in fact that these slopes vary from 1.49 (class 2) to 4.15 (class 6). The final statistic estimated is the residual variance shown to be 213.22. This value is a measure of the variance not accounted for by the analysis, and is not tested for significance in HLM.

Generalizing the Analyses

The Random-Coefficients Model with two Level 1 Predictors

Of course, it is possible to perform much more complex analyses. Thus, within the Random-Coefficients model, you could have more than one Level 1 predictor. For example, you might add another variable (e.g., FANX) to the equation and investigate the intercepts and slopes for the two predictors (i.e., FMOT and FANX). Following the instructions for using HLM and including both FMOT and FANX as level 1 predictors yields the following equation:

Mixed Model

$$GRS_{ij} = \gamma_{00} + \gamma_{10}*(FMOT_{ij} - \overline{FMOT}_{..}) + \gamma_{20}*(FANX_{ij} - \overline{FANX}_{..}) + u_{0j} + u_{1j}*(FMOT_{ij} - \overline{FMOT}_{..}) + u_{2j}*(FANX_{ij} - \overline{FANX}_{..}) + r_{ij}$$

Note that this equation is a direct extension of Equation 4 (and the corresponding equation for the random-coefficients model with one predictor presented in the first example) except that it has two additional elements. The equivalent of equation 1 then is:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{1ij} - \overline{X_{1.j}}) + \beta_{2j}(X_{2ij} - \overline{X_{2.j}}) + r_{ij} \qquad (1a)$$

and following the substitution for level 2 equations for β_{oj}, β_{1j}, and β_{2j}, the general form of the equation can be written as:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{1ij} - \overline{X_{1.j}}) + \mu_{1j}(X_{1ij} - \overline{X_{1.j}}) + \mu_{0j} + r_{ij}$$

$$+ \gamma_{20}(X_{2ij} - \overline{X_{2.j}}) + \mu_{2j}(X_{2ij} - \overline{X_{2.j}})$$

Note that the first part of the equation is identical to that of equation (4) except that the subscript 1 has been introduced to refer to the first set of X values. The two new terms are introduced on the second line. One is the cross-product of the mean slope with the second centered variable and the other is the cross-product of the deviations of the group slopes from the mean slope with the second centered variable. These correspond to the terms, γ_{20}*FANX $_{ij}$ deviations and μ_{2j}*FANX$_{ij}$ deviations in the equation shown in HLM, and can be used to assess the mean slope and the variance in these slopes for the second variable.

The run on HLM using this model produced the following results:

Tau Matrix

INTRCPT1, B0	43.43178	1.42032	9.17814
FMOT, B1	1.42032	2.27971	1.25093
FANX, B2	9.17814	1.25093	2.34899

Tau (as correlations)

INTRCPT1, B0	1.000	0.143	0.909
FMOT, B1	0.143	1.000	0.541
FANX, B2	0.909	0.541	1.000

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	d.f.	Approx. P-value
For INTRCPT1, B0					
INTRCPT2, G00	55.106946	2.969004	18.561	5	0.000
For FMOT slope, B1					
INTRCPT2, G10	2.476595	0.771614	3.210	5	0.028
For FANX, slope, B2					
INTRCPT2, G20	-1.923500	0.913282	-2.106	5	0.087

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	6.59028	43.43178	5	31.24741	0.000
FMOT slope, U1	1.50987	2.27971	5	17.51254	0.004
FANX, slope, U2	1.53264	2.34899	5	5.38891	0.370
level-1, R	13.94577	194.48452			

There are a number of things that can be noted in these results, the major one of which is that there are a few more estimates and many of the values that we obtained in the previous analysis are slightly different. As before, the degrees of freedom for the t-tests of the fixed effects are 5, but the estimates are slightly different, as are the *t*-values and the significance levels. Nonetheless, it is still the case that the mean intercept is significantly different from 0 as is the mean slope for FMOT. The *t*-test for the mean slope for FANX is not significant, $t(5) = -2.11$, *ns*, indicating that in the presence of FMOT, FANX does not contribute to the prediction of grades. The tests of the random effects, similarly have 5 degrees of freedom as before, and although the variances and chi-square values are different, the variances of the intercepts and the slopes for FMOT are still significantly greater than 0, indicating that they both differ among the groups. The variance of FANX over the six groups is not significant, indicating that these slopes do not differ across the groups, Chi-square = 5.39,*ns*.. Note, in this case the slopes for FMOT and FANX are slopes in a model that has only these two predictors and, just as in multiple regression, the values of the slopes of each variable are a function of the relationship of the two predictors to each other as well as with the criterion (grades). In fact, the OLS estimates of these slopes are the actual regression coefficients you would obtain if you were to perform the multiple regression analysis separately for each group. Because of the rationale discussed earlier, they would not be identical to the maximum likelihood estimates obtained with this model. Clearly, this could be made more complex by adding more Level 1 predictors. As is the case with multiple regression, however, the interpretation of the results becomes even more complex.

The Intercepts- and Slopes-as-Outcomes Model

It is also possible to test a model with **non-random varying slopes** (Raudenbush & Bryk, 2002, p. 28). In this case, one has at least one Level 2 variable and the question is whether there is a linear relationship between the Level 2 variable(s) and the slopes (and intercepts) for the Level 1 variables. In this case, the characteristic of the level 2 variable is defined as W_j , a value that describes some aspect of group j (i.e., is common to all members in group j) and the fundamental level 1 equation is as before:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \overline{X}_{.j}) + r_{ij} \qquad (1)$$

but now the two level 2 equations are:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}$$

and

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + \mu_{1j}$$

Combining the two sets of equations leads to the general form:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j} + \gamma_{10}(X_{ij} - \overline{X}_{.j}) + \gamma_{11}W_j(X_{1ij} - \overline{X}_{1.j})$$

$$+ \mu_{1j}(X_{ij} - \overline{X}_{.j}) + r_{ij}$$

Thus, assume there was one Level 1 variable (e.g., FMOT, identified as X), and one Level 2 variable (e.g., ATTCLAS, identified as W). For sake of illustration, assume that W was the mean attitude toward the learning situation for each class. This could be entered into the data file for each individual in that class (for use in step 9 when setting up the HLM run), or could be contained in another file where each level of class is indicated using the same code as in the original data file, the score for ATTCLAS for each class, and some other variable (anything, even nonsense numbers) and then read in at step 9 in the HLM run). For the simplest form where there is the grouping variable (ENV), the Level 1 variable (FMOT), and the Level 2 variable (ATTCLAS), the full equation would appear in the HLM output as follows:

Mixed Model

$$GRS_{ij} = \gamma_{00} + \gamma_{01}*(ATTCLAS_{j} - \overline{ATTCLAS}_{.}) + \gamma_{10}*(FMOT_{ij} - \overline{FMOT}_{..}) + \gamma_{11}*(ATTCLAS_{j} - \overline{ATTCLAS}_{.})*(FMOT_{ij} - \overline{FMOT}_{..}) + u_{0j} + u_{1j}*(FMOT_{ij} - \overline{FMOT}_{..}) + r_{ij}$$

Note this model has two additional terms. The first is the regression of the class intercepts on ATTCLAS (i.e., γ_{01}*ATTCLAS deviation) and the second is the regression of the class slopes on ATTCLAS (i.e., γ_{11}*ATTCLAS*FMOT deviations). This is achieved by adding ATTCLAS as a predictor at Level 2 to the equations for both the intercepts and the slopes. If it were felt desirable to not estimate one of these values, that term could be left out of the model by not including it in its equation. The results are:

Tau Matrix

INTRCPT1,B0	13.79450	-4.32908
FMOT,B1	-4.32908	2.44741

Tau (as correlations)

INTRCPT1,B0	1.000	-0.745
FMOT,B1	-0.745	1.000

Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	d.f.	Approx. P-value
For INTRCPT1, B0					
INTRCPT2, G00	55.412869	1.998498	27.727	4	0.000
ATTCLAS, G01	-8.430881	3.202816	-2.632	4	0.056
For FMOT slope, B1					
INTRCPT2, G10	2.905690	0.780974	3.721	4	0.033
ATTCLAS, G11	-0.067059	1.279706	-0.052	4	0.961

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P-value
INTRCPT1, U0	3.71409	13.79450	4	9.62274	0.047
FMOT slope, U1	1.56442	2.44741	4	16.17302	0.003
level-1, R	14.62204	213.80410			

It is interesting to compare these results with those presented earlier. The models are both comparable in that the only level 1 predictor is FMOT. In the first example, there was no Level 2 predictor but in this case there is, ATTCLAS. Comparison of the Tau matrices will reveal a very large decrease this time in the variance of the intercept from 32.45 to 13.79, but the other values are comparable. The table of final estimation of the fixed effects is comparable as well, though this table has an added set of tests for the Level 2 variable, ATTCLAS. As before, the results demonstrate that the mean intercept is significantly greater than 0, $t(4) = 27.73$, as is the mean slope, $t(4) = 3.72$. There is also a possible indication that the intercepts of the various classes is negatively associated with ATTCLAS, suggesting that higher grades tend to be associated with classes that have less positive attitudes toward the class, but the *t*-value just fails to be significant, $t(4) = -2.63$, $p < .06$, aqnd that there is no association between the slopes in the various classes and ATTCLAS, *t*(4) = .-05. As before, furthermore, the results indicate that both the intercepts and slopes differ among the classes, $\chi^2(4) = 9.62$ and 16.17, respectively.

The Issue of Centering

In the examples shown here, the predictor variables were all grand-mean centered. That is, FMOT, FANX, and ATTCLAS were all centered such that their means were 0. As such, the various intercepts referred to are the values of the dependent variable, GRS, when each of the predictors is considered at its grand mean. In their general discussion of the location of the predictors, Raudenbush and Bryk (2002, pp. 31-35) discuss the implications of using the original metric (i.e., not centering), centering the predictor at the grand mean, centering the predictor at the group level (so that the intercept corresponds to the value of the dependent variable at the mean for each group), or other potentially interesting values. Obviously, the values obtained (particularly those involving the intercepts) will differ depending on the choice one makes, and on the associated interpretation. There are many possibilities, and the reader is referred to Raudenbush and Bryk (2002) for a fuller discussion of some of the implications.

References

Gardner, R.C. (2003). Analysis of Variance with a Continuous Independent Variable: Model I, the Unique Approach. Unpublished manuscript, University of Western Ontario, London, Canada..

Kirk, R. E. (1995). *Experimental Design: Procedures for the Behavioral Sciences. Pacific Grove, CA: Brooks/Cole.*

Raudenbush, S.W. and Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods.* Thousand Oaks, CA: Sage

Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R. & du Toit, M. (2004). *HLM6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL: SSI Scientific Software International.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*. (Fifth Edition). Boston, MA: Allyn & Bacon.