Database Restructuring with the VARSTOCASES Command:
'Point and Click' in SPSS

In Appendix B, we presented the following hypothetical educational database:

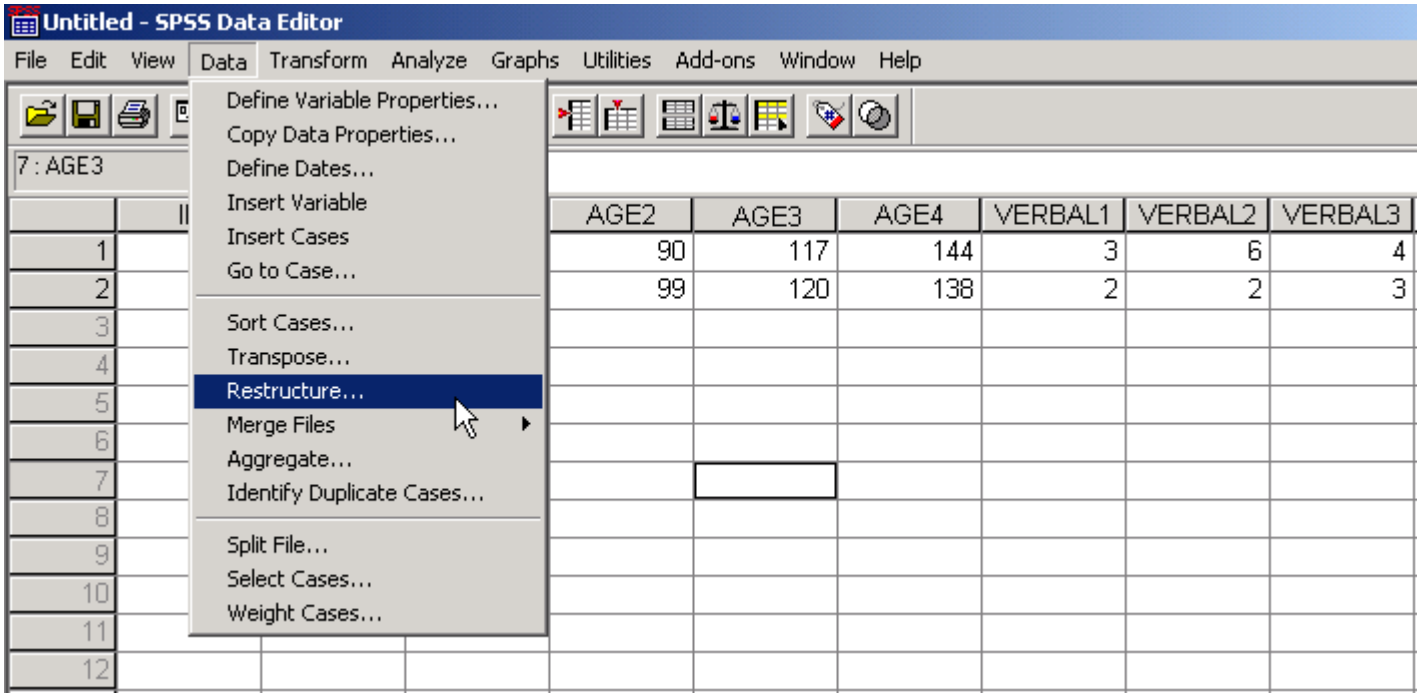
ID	GENDER	AGE1	AGE2	AGE3	AGE4	VERBAL1	VERBAL2	VERBAL3	VERBAL4	SCHOOL
1	1	60	90	117	144	3	6	4	5	1
2	0	63	99	120	138	2	2	3	1	1

This example dataset is in a typical "multivariate" format most familiar to applied researchers: two students provided responses to ten variables. Responses for each student are contained in a separate row in the database, and each response variable has its own column. For multilevel analyses performed in SPSS, "stacked" database is needed where:

- repeated measurements (i.e., Age1-Age4; Verbal1-Verbal4) are contained in 2 separate variable columns (i.e., Age, Verbal)
- each individual (i.e., ID #1 and ID #2) is assigned a unique identification number separating the repeated measurements of each participant
- a variable reflecting the timing of repeated assessments is created and included in the database.

The following steps illustrate the process of restructuring a database from a "multivariate" to a "stacked" format in SPSS via the "point-and-click" interface using the above hypothetical educational data.

Step 1: Assuming the above hypothetical educational data have been entered into SPSS, begin by clicking on the *Data* menu and selecting *Restructure…* as shown below.

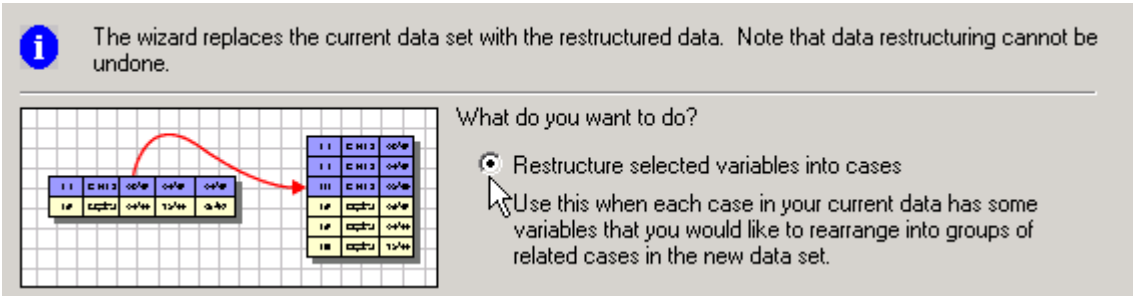


A selection window titled "Welcome to the Restructure Data Wizard!" will appear. This window gives researchers three database restructuring options:

- **Restructure selected variables to cases**
- **Restructure selected cases to variables**
- **Transpose all data**

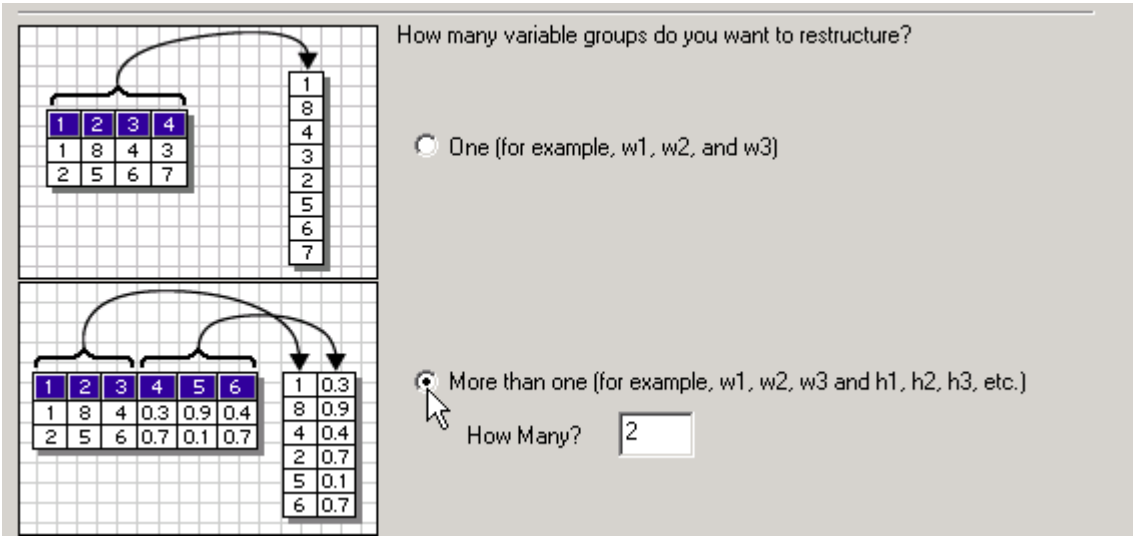
In this example, only the selected variables Age1-Age4 and Verbal1-Verbal4 need to be restructured from a "multivariate" to a "stacked" format.

<u>Step 2</u>: Click the option marked *'Restructure selected variables into cases'* as shown below, then click Next >



The *"Restructure Data Wizard – Step 2 of 7"* selection window will appear, titled "Variables to Cases: Number of Variable Groups". This window requires the researcher to specify how many variables are to be "stacked". In this example, two variable groups (Age1-Age4 and Verbal1-Verbal4) need to be restructured.

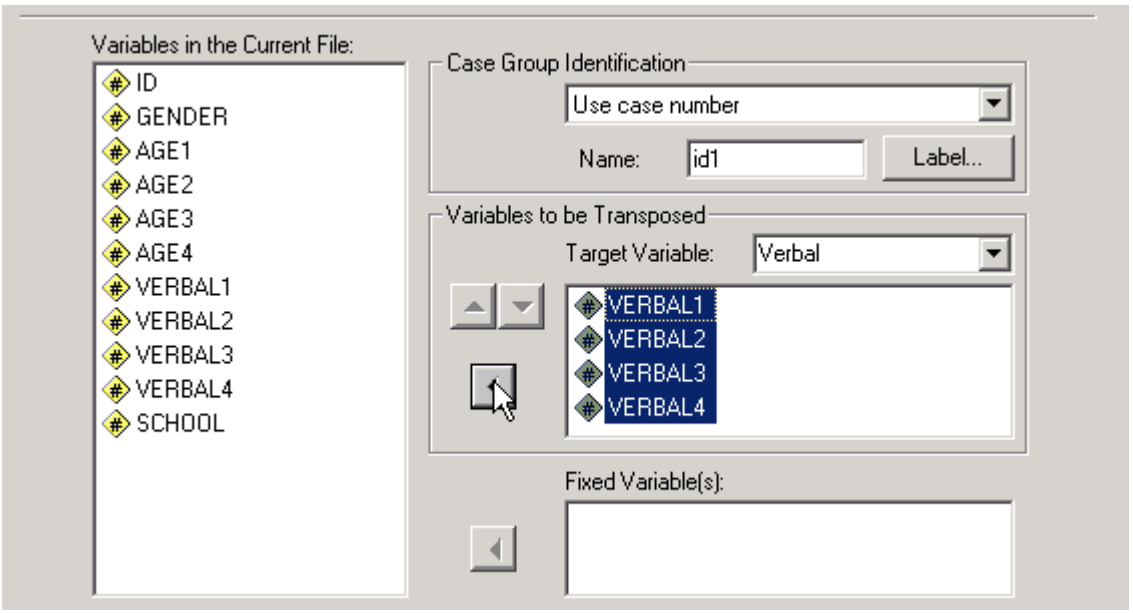
<u>Step 3</u>: Click on the *'More than one'* radio button, and type 2 in the *'How Many?'* box as shown below. Then click Next >



The *"Restructure Data Wizard – Step 3 of 7"* selection window will appear, titled "Variables to Cases: Selected Variables". This window requires the researcher to:

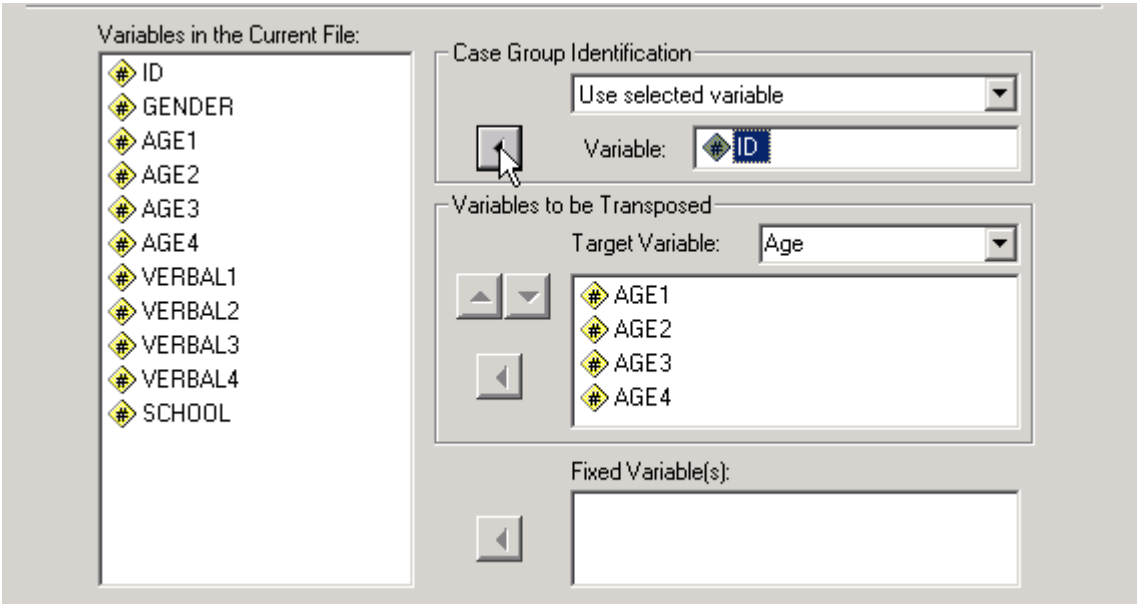
- **Specify and give a new name to each variable group to be "stacked"**
- **Specify how participant groups are to be identified**

<u>Step 4a</u>: In the *'Variables in the Current File:'* window, select Verbal 1, Verbal 2, Verbal 3, and Verbal 4 and move them into the *'Variables to be Transposed'* window. Next, overwrite *'trans1'* in the *'Target Variable'* box by highlighting it and typing Verbal. The repeated measures of Verbal1-Verbal4 will now be "stacked" into a single variable column named Verbal



<u>Step 4b</u>: Click on the *'Target Variable'* window arrow , select *'trans2'*, and overwrite it with the variable name Age. In the *'Variables in the Current File:'* window, select Age1, Age2, Age3, and Age4 and move them into the *'Variables to be Transposed'* window.

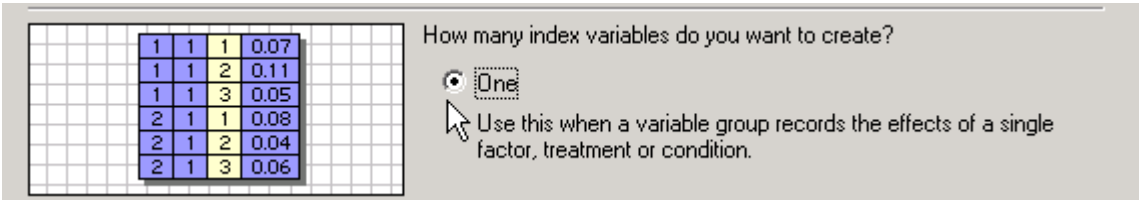
<u>Step 4c</u>: In the *'Case Group Identification'* window, click on the window arrow and choose the *'Use selected variable'* option. In the *'Variables in the Current File:'* window, select the variable ID and move it into the *'Variable:'* window as shown below. These commands indicate that the ID variable will be used to separate the repeated measurements values (i.e., Age1-Age4 and Verbal1-Verbal4) of each participant. Then click Next >



The *"Restructure Data Wizard – Step 4 of 7"* selection window will appear, titled "Variables to Cases: Create Index Variables". In the "multivariate" format, each repeated measure (i.e., Age1-Age4 and Verbal1-Verbal4) had its own variable column. In the "stacked" database, the variables Verbal and Age created in the previous step now contain all repeated measurements. An index variable, containing values from 1 to 4, is needed to indicate each of the four repeated measurements. For example, a participant's Verbal1 score would have an index variable value of 1; a Verbal2 score would have an index variable value of 2, and so on.

<u>Step 5</u>: Click on the *'One'* radio button when asked "How many index variables you want to create?" (In the article, we created one index variable called 'Timevar'. This point-and-click method will create an index variable called *'Index1'* by default).

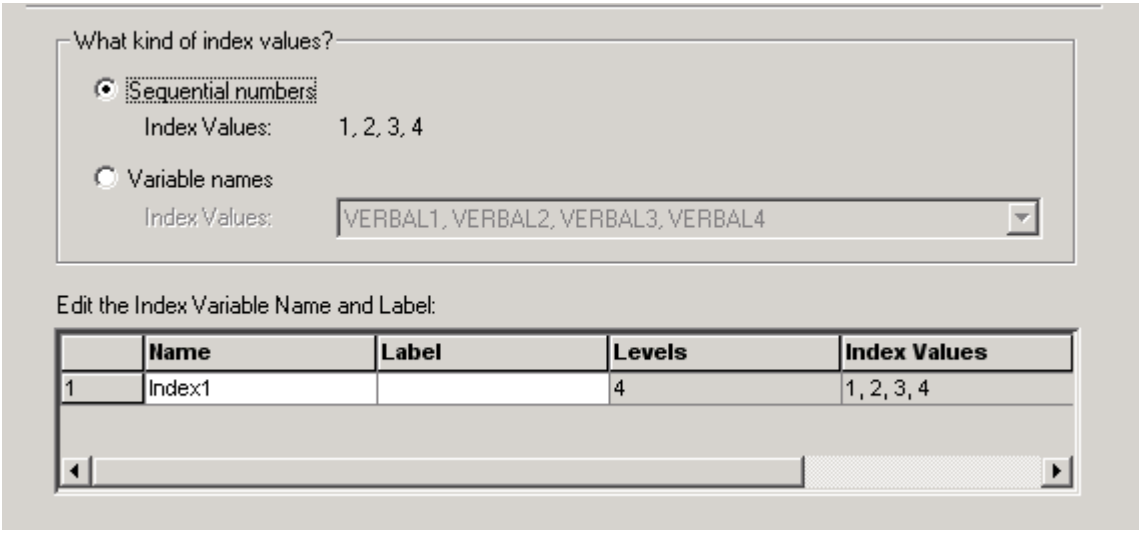
Then click Next >



The *"Restructure Data Wizard – Step 5 of 7"* selection window will appear, titled "Variables to Cases: Create One Index Variable". This window allows the researcher to specify the kind of values given in the index variable created in the previous step. The *'Sequential numbers'* radio button option assigns sequential numeric values for the index variable; the *'Variable names'* radio button option allows researchers to name (and give a label to) an alphanumeric value for the index variable.

<u>Step 6</u>: The *'Sequential numbers'* option is highlighted by default. The hypothetical educational dataset involves *Age* and *Verbal* data collected at four time points. So, *'Index1'* will be a variable column having values from 1-4.

So, just click Next >



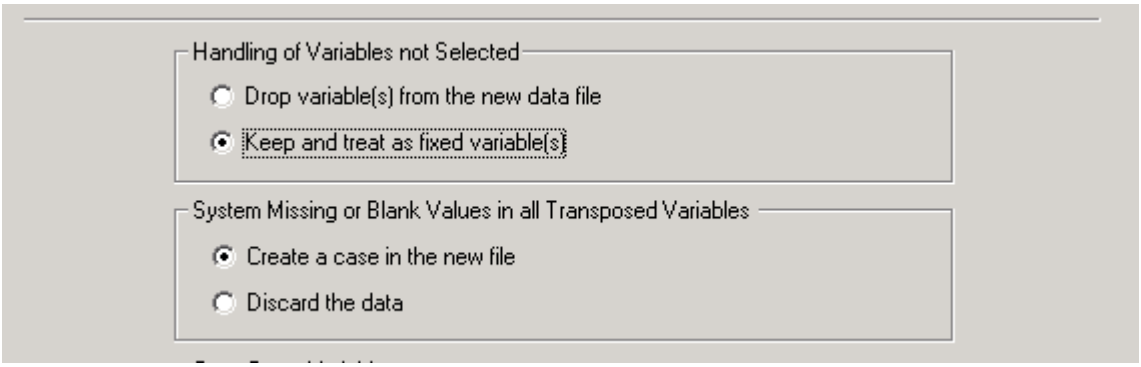
The *"Restructure Data Wizard – Step 6 of 7"* selection window will appear, titled "Variables to Cases: Options". This window requires researchers to:

- **specify what is to be done with variables in the database not selected to be "stacked" in Steps 4a-c**
- **specify how missing values are to be handled.**

<u>Step 6a</u>: In the *'Handling of Variables not Selected'* option, we transposed Verbal and Age, but did not specify how SPSS should handle the remaining variables ID, GENDER, and SCHOOL. Select the *'Keep and treat as fixed variable(s)'* radio button option to retain them in the database.

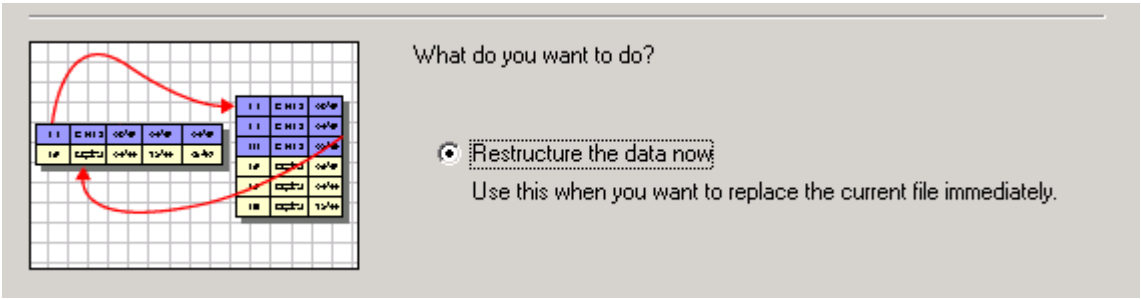
<u>Step 6b</u>: In the *'System Missing or Blank Values in all Transposed Variables'*, this option asks how to treat missing data in the transposed variables (i.e., Age and Verbal). Select *'Create a case in the new file'* to keep missing values in the database (e.g., see Peugh & Enders, 2004).

Then click 



The *"Restructure Data Wizard – Finish"* selection window will appear, titled "Finish". To immediately restructure the database, select the *'Restructure the data now'* radio button. To paste the syntax commands for Steps 1-7 and make modifications (e.g., change the name of the index variable to "Timevar") select the *'Paste the syntax generated by the wizard into the syntax window'* radio button.

<u>Step 7</u>: Click *'Restructure the data now'*



Then click 

.

ID	GENDER	SCHOOL	Index1	Verbal	Age
1	1	1	1	3	60
1	1	1	2	6	90
1	1	1	3	4	117
1	1	1	4	5	144
2	0	1	1	2	63
2	0	1	2	2	99
2	0	1	3	3	120
2	0	1	4	1	138

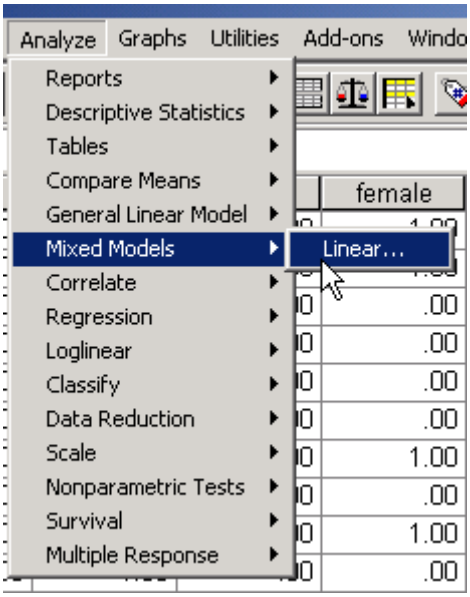
The resulting database is similar (i.e., the index variable *'Timevar'* was created with the syntax command, the point-and-click method assigns the index variable the name *'Index1'* by default) to the database example in the lower portion of Appendix B generated via the SPSS VARSTOCASES syntax.

Cross-Sectional Analysis with Level-1 and Level-2 Covariates:
'Point-and-Click' in SPSS

In the final cross-sectional model example, the group mean centered level-1 covariate *cses* and the level-2 fixed-effect covariates *meanses* (i.e., average school SES) and sector (i.e., 0 = public school, 1 = Catholic school) were added to predict intercept and slope variability in the dependent variable *mathach* (i.e., mean mathematics achievement scores) using the 'High School and Beyond' data.

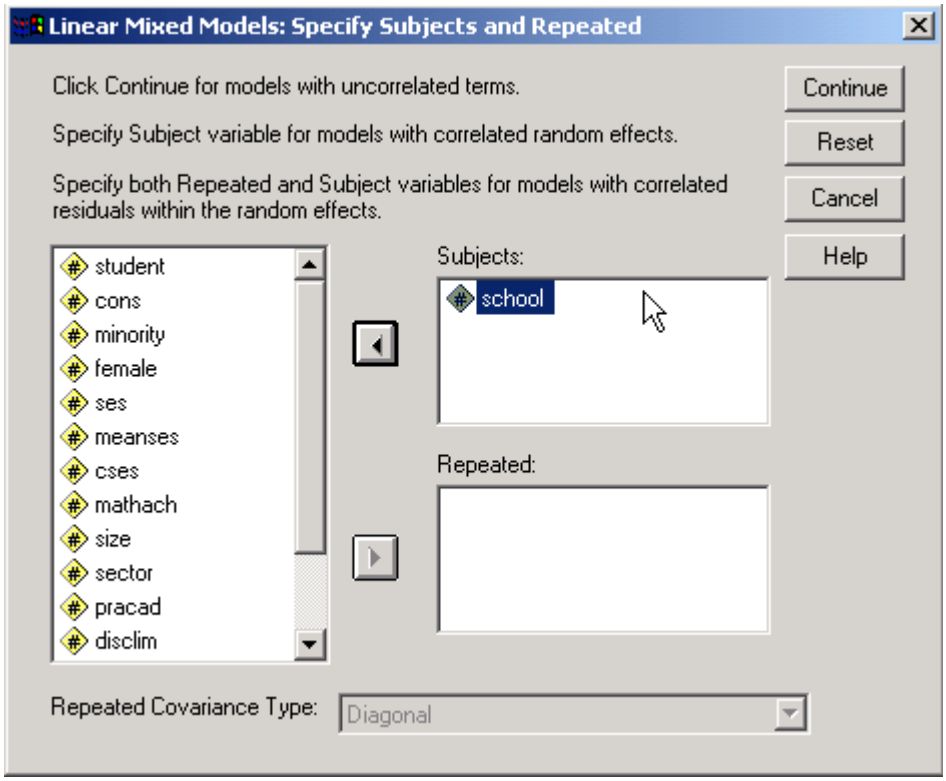
The following steps illustrate the process of performing a cross-sectional multilevel analysis with level-1 and level-2 covariates in SPSS via the "point-and-click" interface using the 'High School and Beyond' database.

<u>Step 1</u>: Begin by clicking on the *Analyze* menu and selecting *Mixed Models, Linear* as shown below.



A selection window titled, "Linear Mixed Models: Specify Subjects and Repeated" will appear. This window requires the researcher to specify variables with correlated random effects. In the 'High School and Beyond' database, students are nested within schools. So students within the same school will have correlated responses (i.e., will have correlated random effects).

<u>Step 2</u>: Select the variable *school* from the variable list window and move it into the *'Subjects:'* window as shown below. Then click Continue

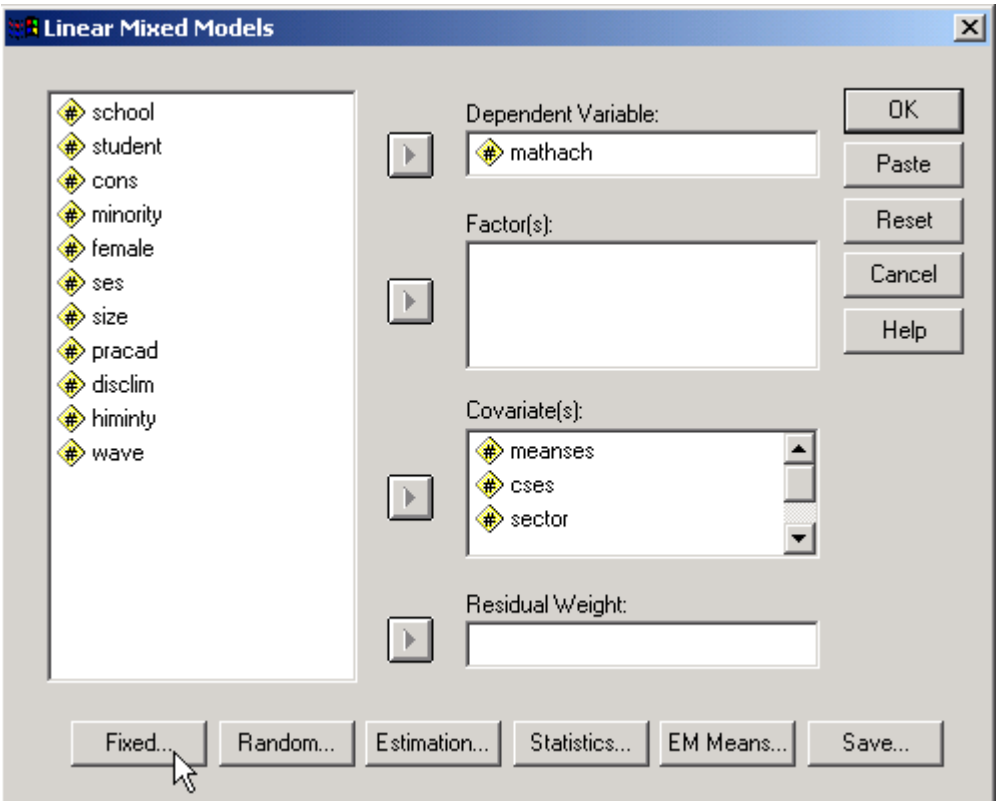


A selection window titled, *'Linear Mixed Models'* will appear. Here the researcher specifies the linear mixed model to be analyzed.

<u>Step 3a:</u> Select *mathach* from the variable list window and move it into the *'Dependent Variable:'* window.

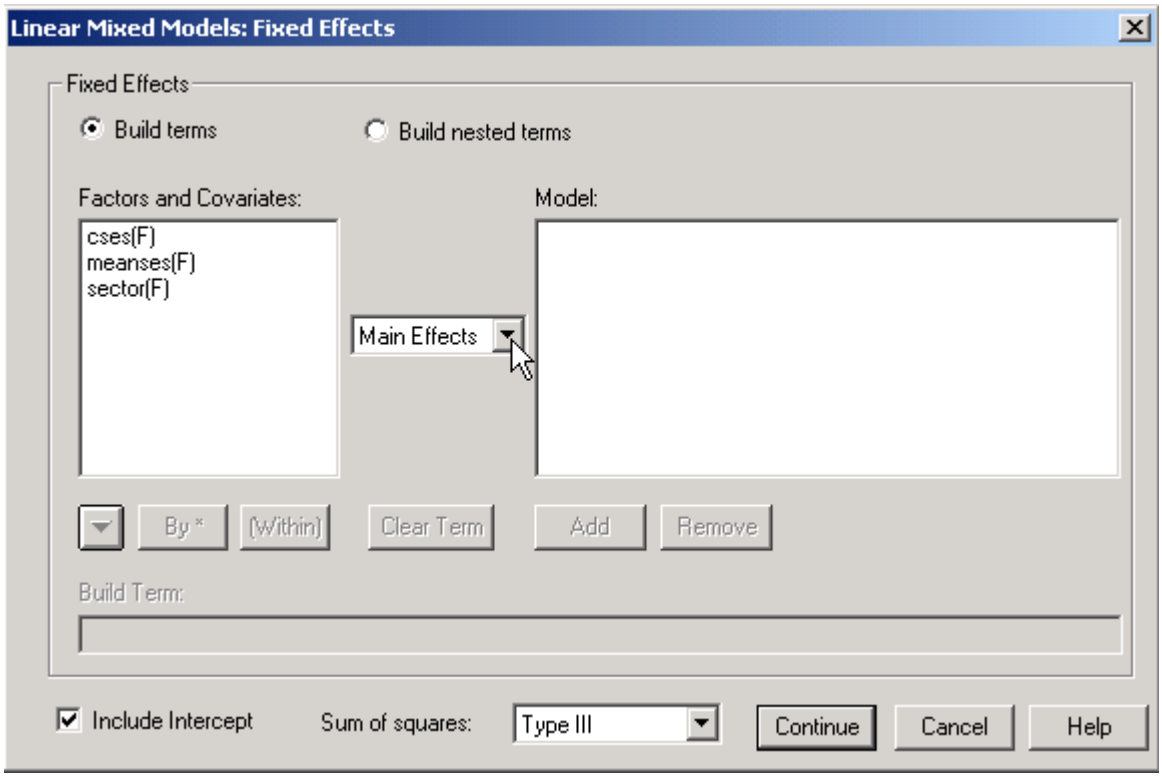
<u>Step 3b</u>. Select *cses* (the level-1 covariate) and both *meanses* and *sector* (the level-2 covariates) from the variable list window and move them into the *'Covariate(s):'* window. As discussed in the article, SPSS makes no distinction between level-1 and level-2 covariates.

<u>Step 3c</u>: Click on the Fixed... **button as shown below.**

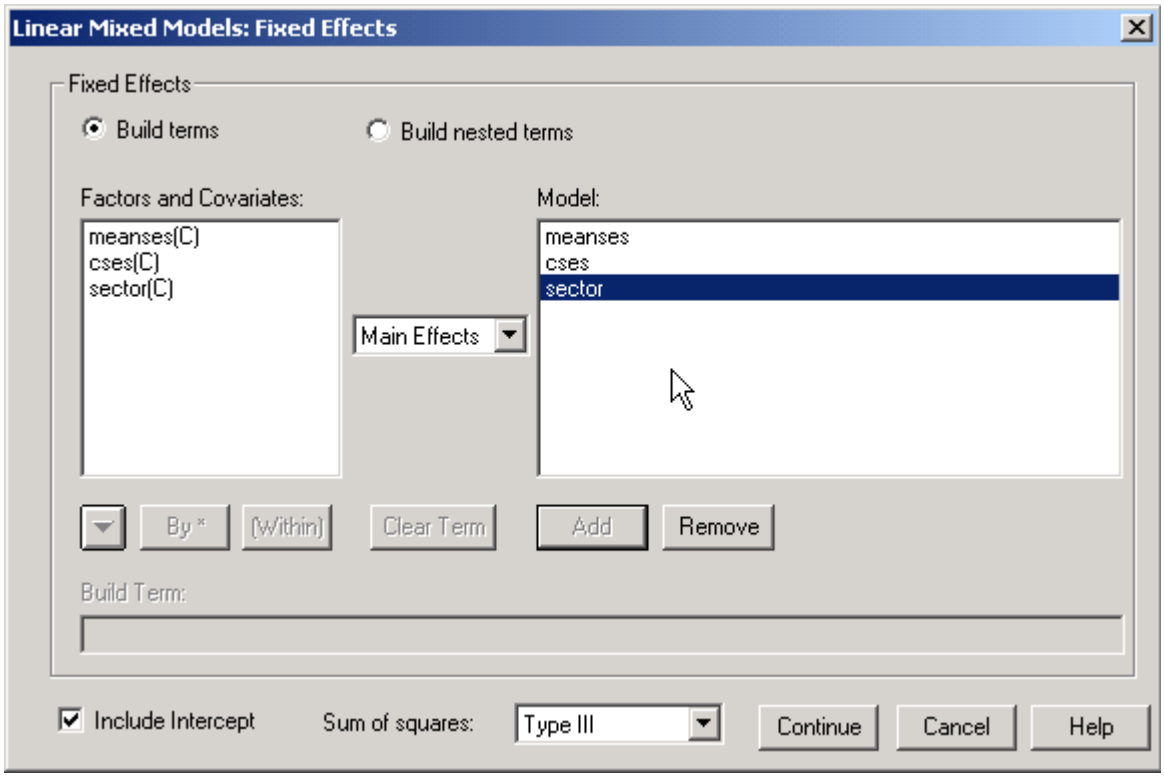


A selection widow titled, "Linear Mixed Models: Fixed Effects" appears. This window requires the researcher to specify all fixed effects in the model to be analyzed. Three main effects (*cses*, *meanses*, and *sector*) and two cross-level interactions (*cses*meanses*, *cses*sector*) are specified.

Check to ensure that: 1) the *'Include Intercept'* checkbox in the lower left corner is checked, 2) the *'Build terms'* radio button is selected, and 3) the window between *'Factors and Covariates:'* and *'Model'* shows *'Main Effects'* as illustrated below.



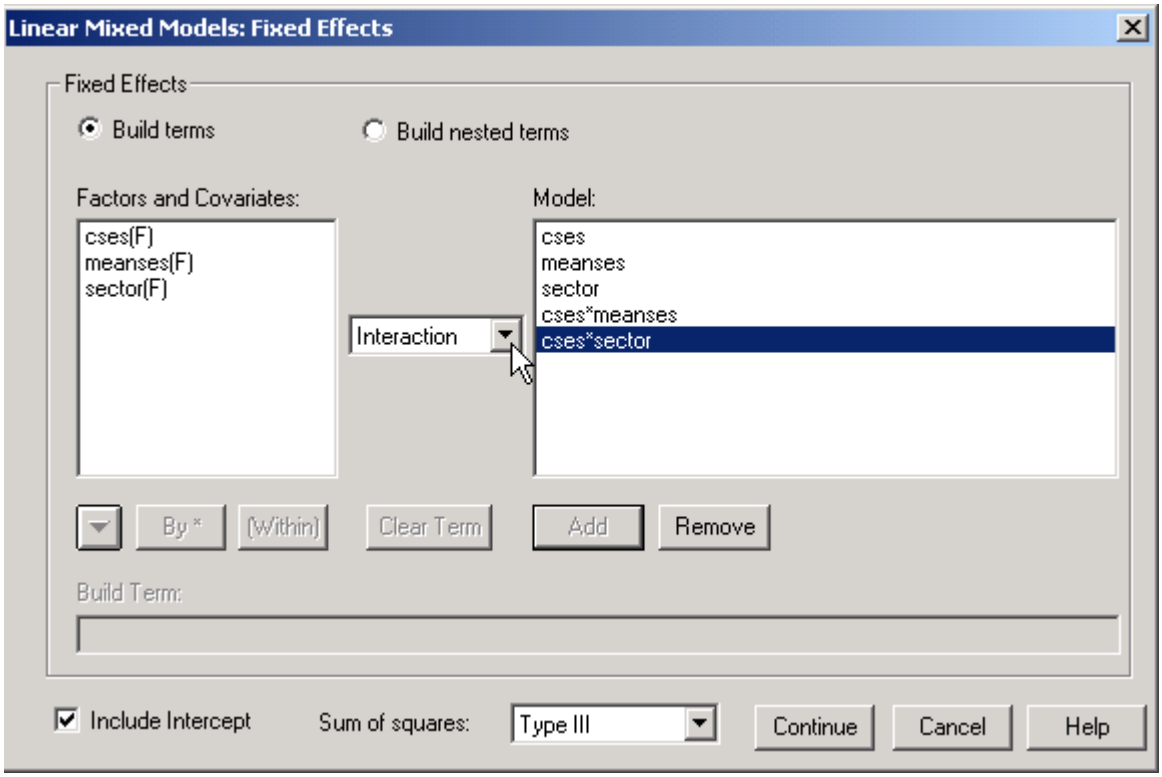
<u>Step 4a</u>: Select *meanses*, *cses*, and *sector* from the *'Factors and Covariates:'* window and move them to the *'Model:'* window as shown below.



<u>Step 4b</u>: Change the *'Main Effects'* option in the middle window to *'Interactions'*.
Include the two cross-level interactions (i.e., *cses*meanses*, *cses*sector*) to the *'Model:'* window as follows:

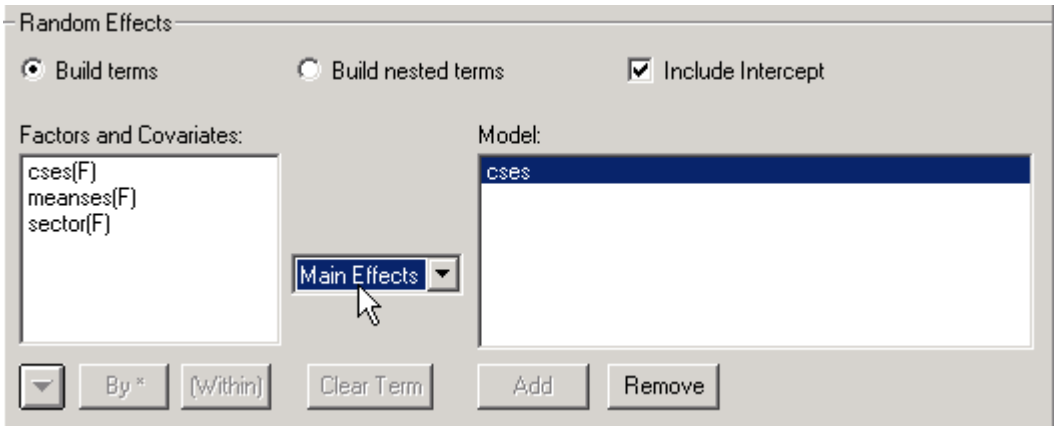
- **select *cses*, press and hold the 'Ctrl' key, select *meanses*, click the Add button**
- **select *cses*, press and hold the 'Ctrl' key, select *sector*, click the Add button**

<u>Step 4c</u>: Click Continue, then click the Random... button.

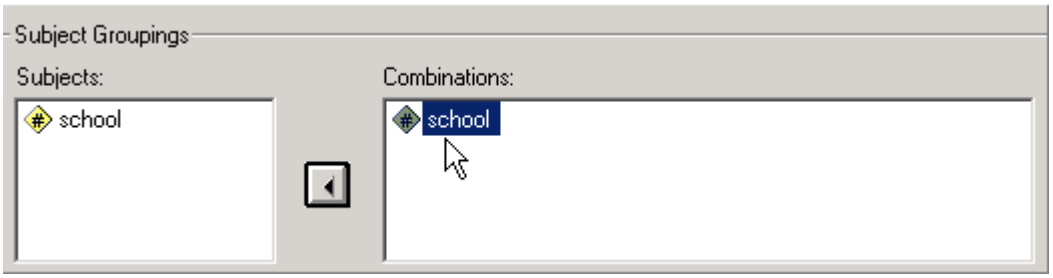


A selection window titled, "Linear Mixed Models: Random Effects" will appear. This window requires researchers to specify the variables that vary randomly in the mixed model. In the 'High School and Beyond' example, only *cses* is allowed to vary across schools.

<u>Step 5a</u>: Check the *'Include Intercept'* box, select *'Main* Effects', then click on *cses* in the *'Factors and Covariates:'* window and move it into the *'Model:'* window as shown below.



<u>Step 5b</u>: Click on *school* in the *'Subjects:'* window under *'Subject Groupings'* and move it into the *'Combinations:'* window as shown below. This indicates *cses* varies across schools.

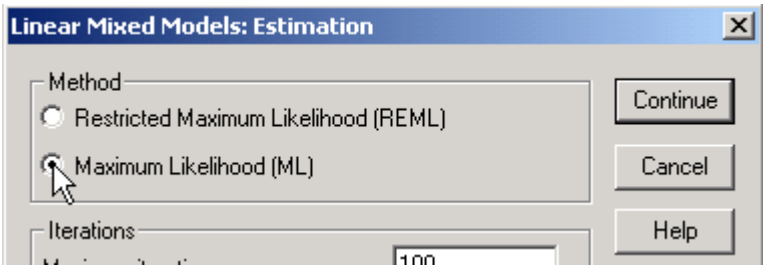


<u>Step 5c</u>: Click on the *'Covariance Type:'* window at the top to change the *'Variance Components'* default to *'Unstructured'* as shown below.
As indicated in the article, an *'Unstructured'* covariance matrix means that intercept variance, *CSES* slope variance, and the covariance between the intercept and slope values are estimated.

<u>Step 5d</u>: Click  **, then click the**  **button.**

A selection window titled, "Linear Mixed Models: Estimation" will appear. This window allows researchers to modify the estimation default settings.

<u>Step 6a</u>: Since this analysis was conducted using maximum likelihood (ML) estimation instead of the default restricted maximum likelihood (REML; see Singer & Willett, 2003, pp. 87-90), click on the *'Maximum Likelihood (ML)'* radio button as shown below (All article analyses were performed with ML estimation).



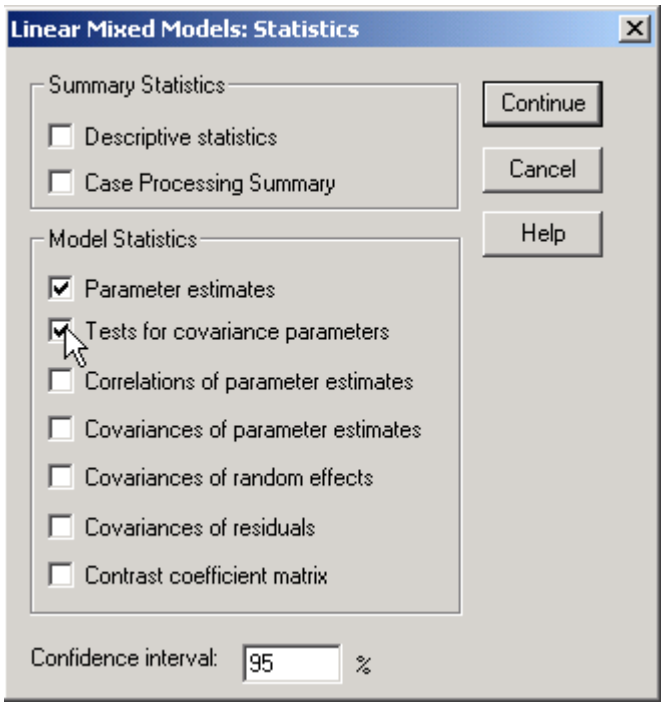
<u>Step 6b</u>: Click 

, then click on the 

button.

A selection window titled, "Linear Mixed Models: Statistics" will appear. This window allows researchers to control the printed output.

<u>Step 7a</u>: Under *'Model Statistics'* click the checkboxes for *'Parameter Estimates'* and *'Tests for covariance parameters'*. These selections will print fixed effect estimates and standard errors, as well as significance tests for the variance components, respectively.



<u>Step 7b</u>: Click Continue

<u>Step 8</u>: Click OK

Estimates of Fixed Effects[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	12.11360	.1969132	162.972	61.517	.000	11.7247725	12.5024329
meanses	5.3379502	.3656776	153.728	14.597	.000	4.6155484	6.0603520
cses	2.9393565	.1534947	140.203	19.150	.000	2.6358931	3.2428200
sector	1.2169438	.3033674	152.323	4.011	.000	.6175930	1.8162947
meanses * cses	1.0423705	.2960358	162.756	3.521	.001	.4578043	1.6269367
cses * sector	-1.64386	.2373583	145.075	-6.926	.000	-2.1129911	-1.1747369

a. Dependent Variable: mathach.

Estimates of Covariance Parameters[a]

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		36.72115	.6261133	58.649	.000	35.5142727	37.9690485
Intercept + cses [subject = school]	UN (1,1)	2.3188948	.3610217	6.423	.000	1.7090709	3.1463138
	UN (2,1)	.1881010	.1985144	.948	.343	-.2009800	.5771821
	UN (2,2)	.0652362	.2071072	.315	.753	.0001295	32.8717526

a. Dependent Variable: mathach.

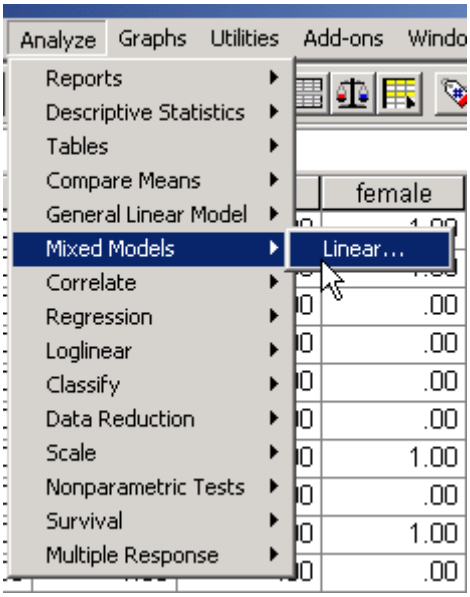
This output matches the output produced by the SPSS MIXED syntax.

Longitudinal Analysis Testing the Error Covariance Matrix:
'Point-and-Click' in SPSS

The final longitudinal model presented in the article involved testing an alternate error covariance structure for Willett's (1988) simulated data. Specifically, "opposites-naming" scores (i.e., *Y*), and a grand mean centered covariate (i.e., *ccovar*) were simulated at four time points for 35 cases. Briefly, testing an alternate error covariance structure examines whether all elements of the error covariance matrix need to be freely estimated (i.e., require an "unstructured" covariance matrix) or could be more parsimoniously modeled with a specific error covariance structure.

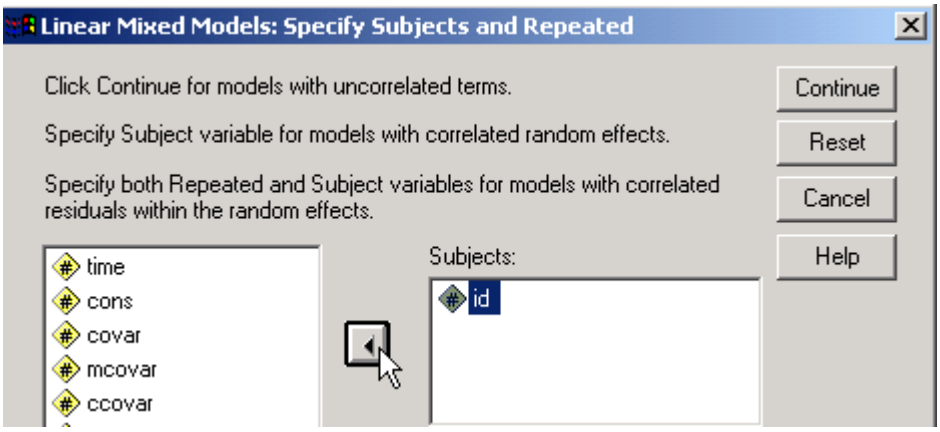
The following steps illustrate the process of performing an alternate error covariance test in SPSS via the "point-and-click" interface using the Willett's (1988) simulated database.

<u>Step 1</u>: Begin by clicking on the *Analyze* menu and selecting *Mixed Models, Linear* as shown below.



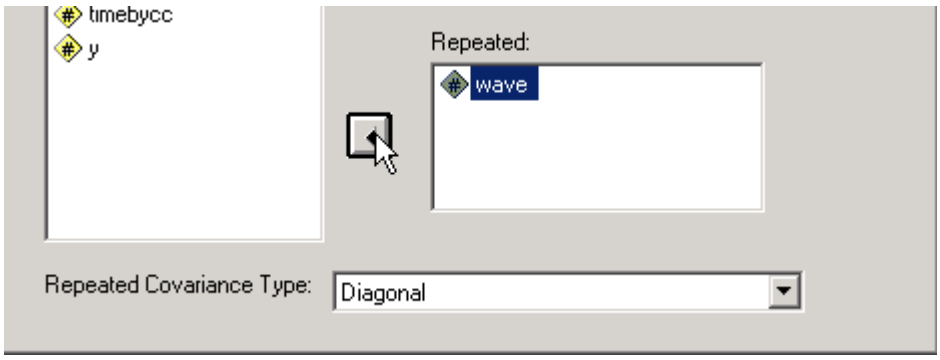
A selection window titled, "Linear Mixed Models: Specify Subjects and Repeated" will appear. This window requires the researcher to specify variables with correlated random effects. In Willett's (1988) database, repeated measures are nested within individuals, and individuals are coded by *ID*. So repeated observations within each individual are correlated responses (i.e., will have correlated random effects).

Step 2a: Select *ID* from the variable list and move it into the *'Subjects:'* window as shown below.

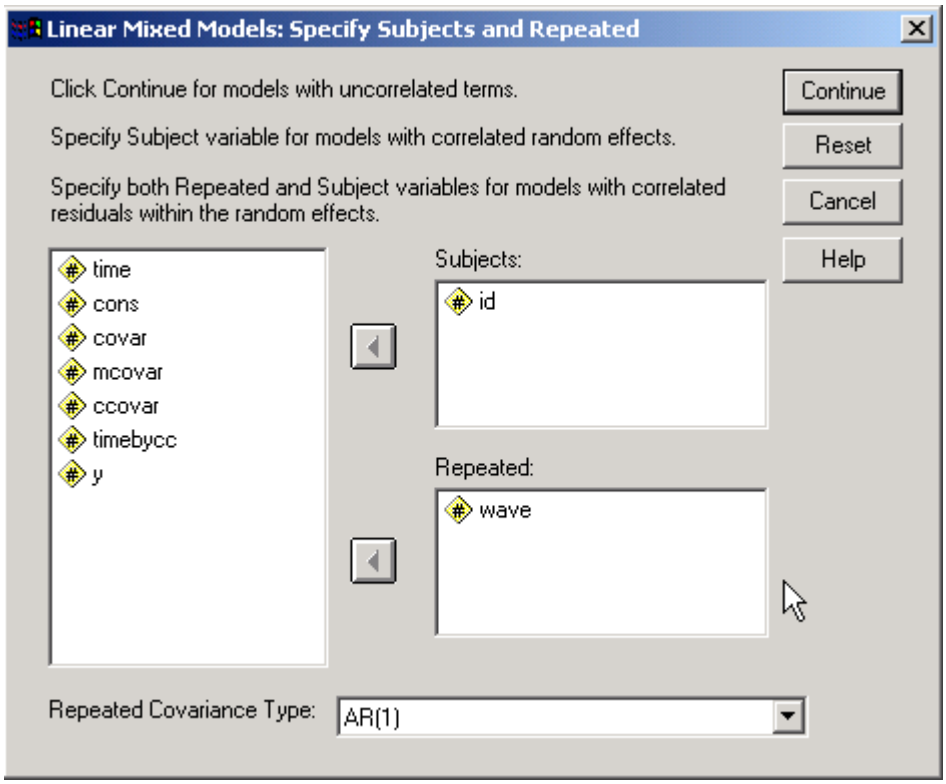


As discussed in the article, *wave* is a categorical variable that assigns an integer value to each repeated measurement (e.g., 1, 2, …, *t*, where *t* is the number of repeated measures).

<u>Step 2b</u>: *Wave* is moved into the *'Repeated:'* window as shown below.

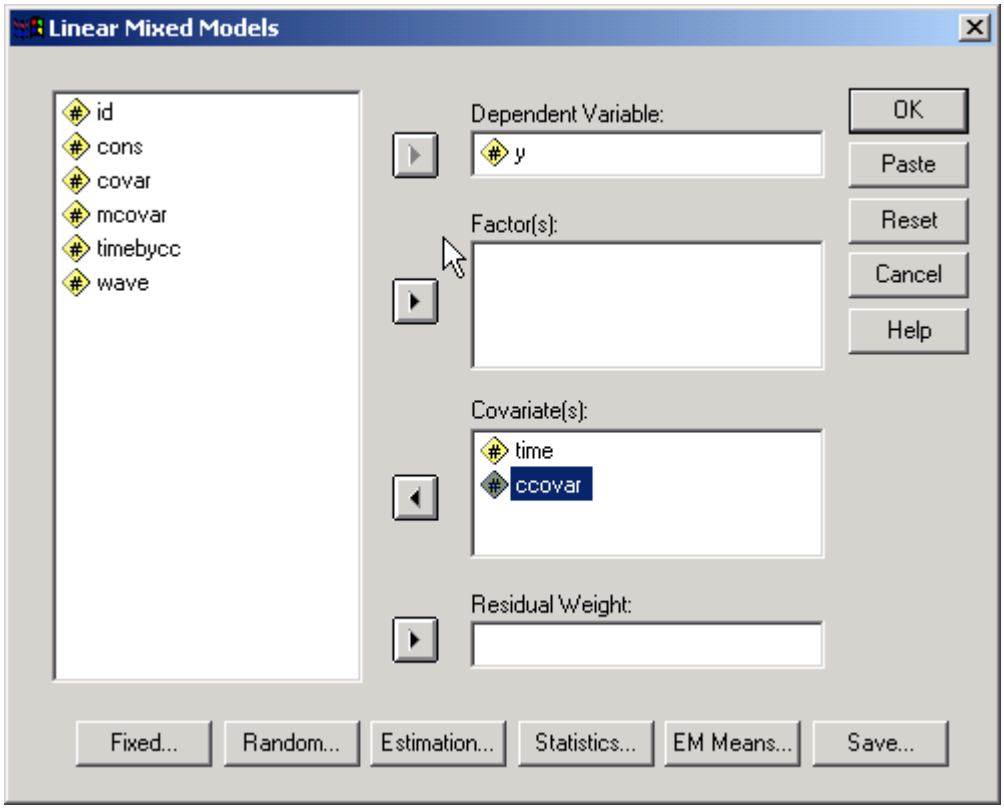


<u>Step 2c</u>: Select AR(1) (i.e., first order autoregressive) from the *'Repeated Covariance Type'* window since it is the error covariance structure to be tested. Click Continue



The *'Linear Mixed Models'* selection window appears that allows the researcher to specify the model to be analyzed.

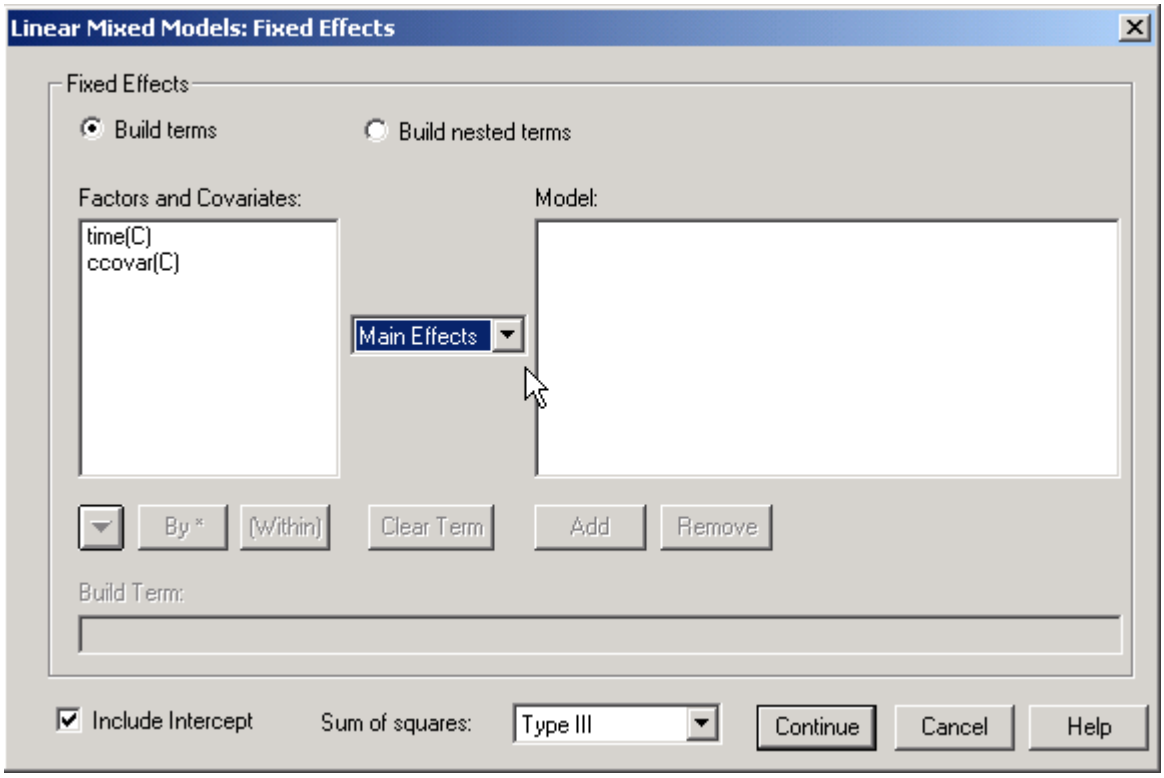
<u>Step 3a</u>: From the variable list, move *Y* to the *'Dependent Variable'* window, and move *time* and *ccovar* to the *'Covariate(s):'* window as shown below.



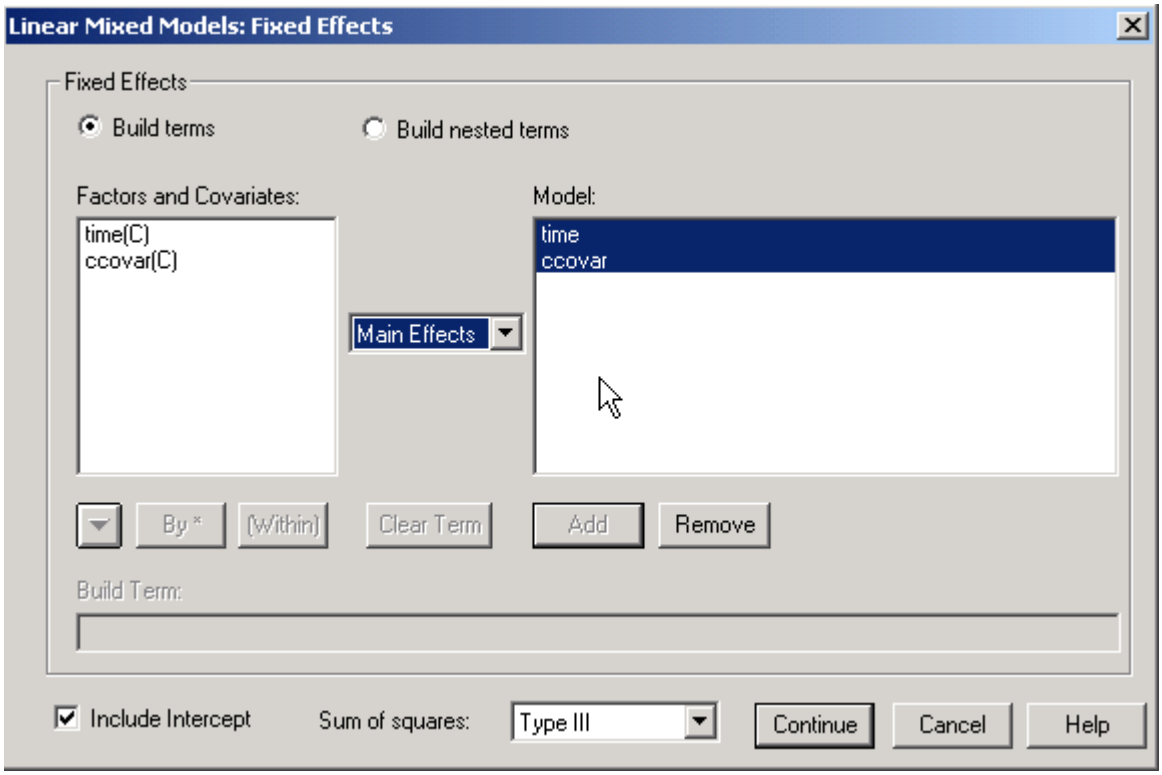
<u>Step 3b</u>: Click on the  **button.**

A selection widow titled, "Linear Mixed Models: Fixed Effects" appears. This window requires the researcher to specify all fixed effects in the model to be analyzed. Two main effects (*time* and *ccovar*) and one interaction (*time*ccovar*) are specified.

Check to ensure that: 1) the *'Include Intercept'* checkbox in the lower left corner is checked, 2) the *'Build terms'* radio button is selected, and 3) the window between *'Factors and Covariates:'* and *'Model'* shows *'Main Effects'* as illustrated below.

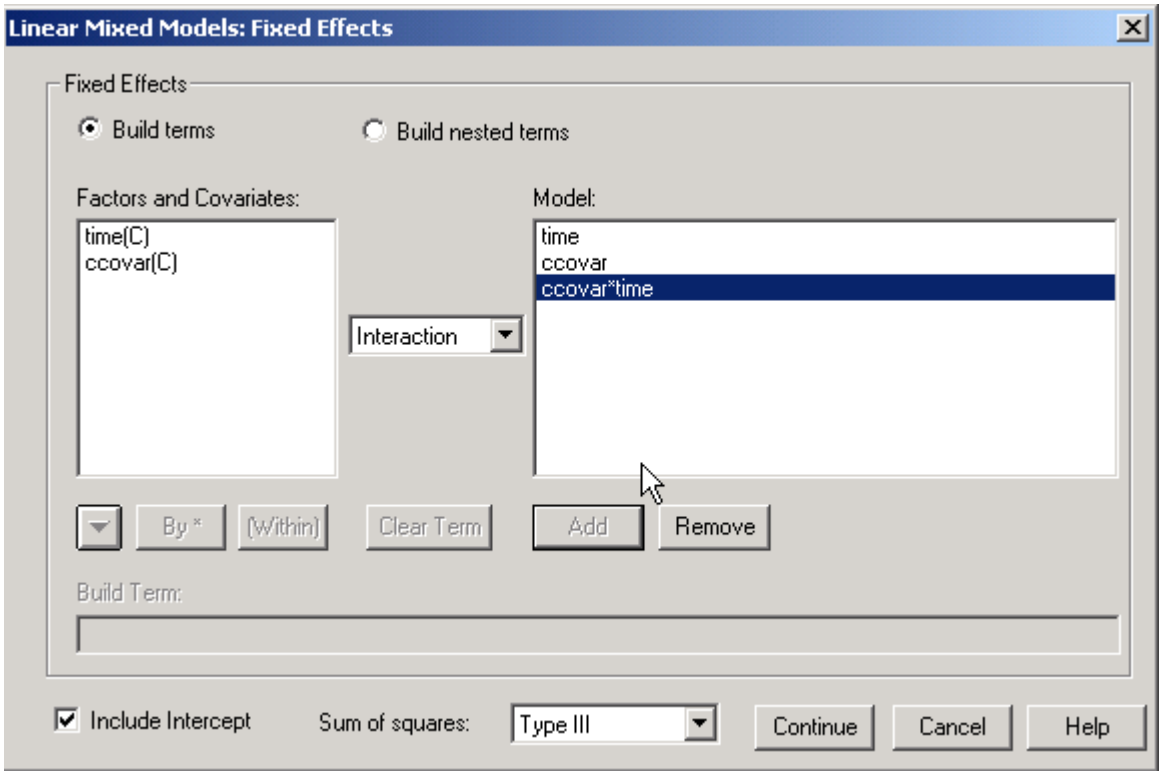


4a. <u>Step 4a</u>: Select *time* and *ccovar* from the *'Factors and Covariates:'* window and move them to the *'Model:'* window as shown below.



<u>Step 4b</u>: Change the *'Main Effects'* option in the middle window to *'Interactions'*. Include the interaction (i.e., *time*cocvar*) to the *'Model:'* window as follows:

- **select *time*, press and hold the 'Ctrl' key, select *ccovar*, click the Add button**



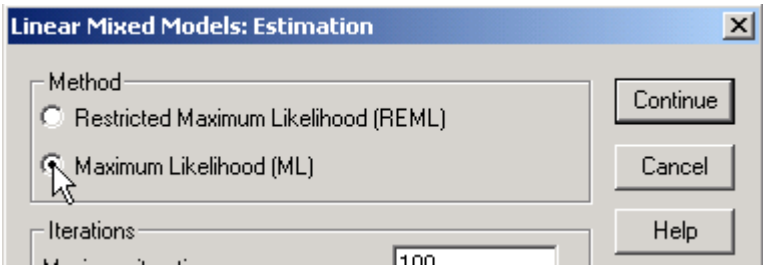
<u>Step 4c</u>: Click Continue

As stated in the article, the /RANDOM subcommand has been replaced by the /REPEATED subcommand already specified in Step 2. So the Random... **button is skipped in this analysis.**

<u>Step 4d</u>: Click the Estimation... **button.**

A selection window titled, "Linear Mixed Models: Estimation" will appear. This window allows researchers to modify the estimation default settings.

<u>Step 5a</u>: Since this analysis was conducted using maximum likelihood (ML) estimation instead of the default restricted maximum likelihood (REML; see Singer & Willett, 2003, pp. 87-90), click on the *'Maximum Likelihood (ML)'* radio button as shown below (All article analyses were performed with ML estimation).

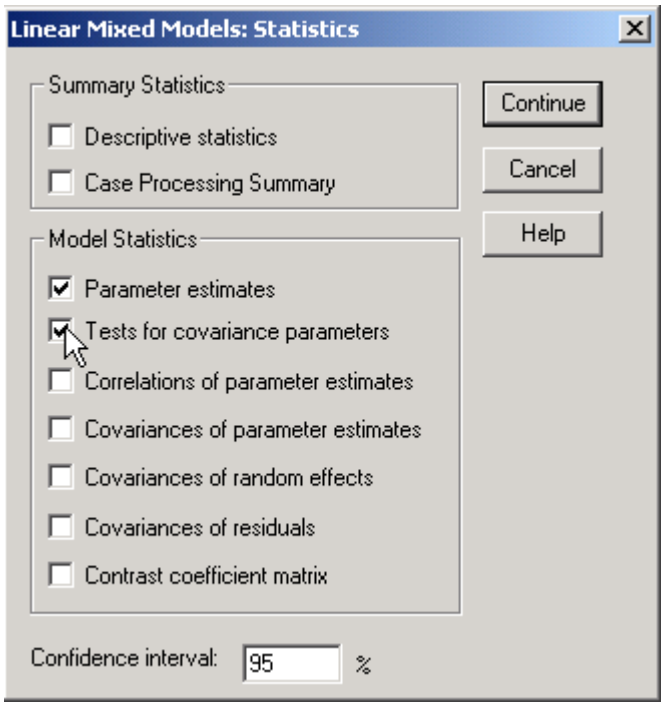


<u>Step 5b</u>: Click on the 

button.

A selection window titled, "Linear Mixed Models: Statistics" will appear. This window allows researchers to control the printed output.

<u>Step 6a</u>: Under *'Model Statistics'* click the checkboxes for *'Parameter Estimates'* and *'Tests for covariance parameters'*. These selections will print fixed effect estimates and standard errors, as well as significance tests for the variance components, respectively.



<u>Step 6b</u>: Click Continue, then click OK

Estimates of Fixed Effects[a]

						95% Confidence Interval	
Parameter	Estimate	Std. Error	df	t	Sig.	Lower Bound	Upper Bound
Intercept	164.3353	5.8127822	54.049	28.271	.000	152.6816396	175.9890136
time	27.19700	1.8467475	55.302	14.727	.000	23.4964855	30.8975068
ccovar	-.0341955	.4720700	54.049	-.072	.943	-.9806193	.9122284
time * ccovar	.4197220	.1499788	55.302	2.799	.007	.1191947	.7202493

a. Dependent Variable: y.

Estimates of Covariance Parameters[a]

						95% Confidence Interval	
Parameter		Estimate	Std. Error	Wald Z	Sig.	Lower Bound	Upper Bound
Repeated Measures	AR1 diagonal	1188.100	225.9425	5.258	.000	818.4250679	1724.753689
	AR1 rho	.8182580	.0398242	20.547	.000	.7236762	.8826694

a. Dependent Variable: y.