National Data Archive on Child Abuse and Neglect

SPSS Instructions and Syntax to Restructure LONGSCAN Files from being Multiple-Records-Per-Participant to One-Record-Per-Participant

This document will assist LONGSCAN data users with restructuring (also called reshaping) data files from being multiple-records-per-participant to one-record-per-participant. This is a common task prior to merging data files, of varying structures, together to form one data file.

First, examine the data file to see if there are multiple records per participant. You can do this by conducting a "Frequencies" command on the "id" variable. If the output displays a number higher than "1" in the frequencies column for any id, then your data file is a stacked dataset.

Start the Restructure Data Wizard: From the "Data" menu at the top of the screen for SPSS, you will select "Restructure." This will start the Restructure Wizard for SPSS. You may be prompted to save your dataset before the Wizard opens.

Step One: Select "Restructure selected cases into variables" from the "Welcome to the Restructure Data Wizard" screen (see Figure 1) and then click "Next."

*Figure 1. A screenshot of step one from the Restructure Data Wizard.*
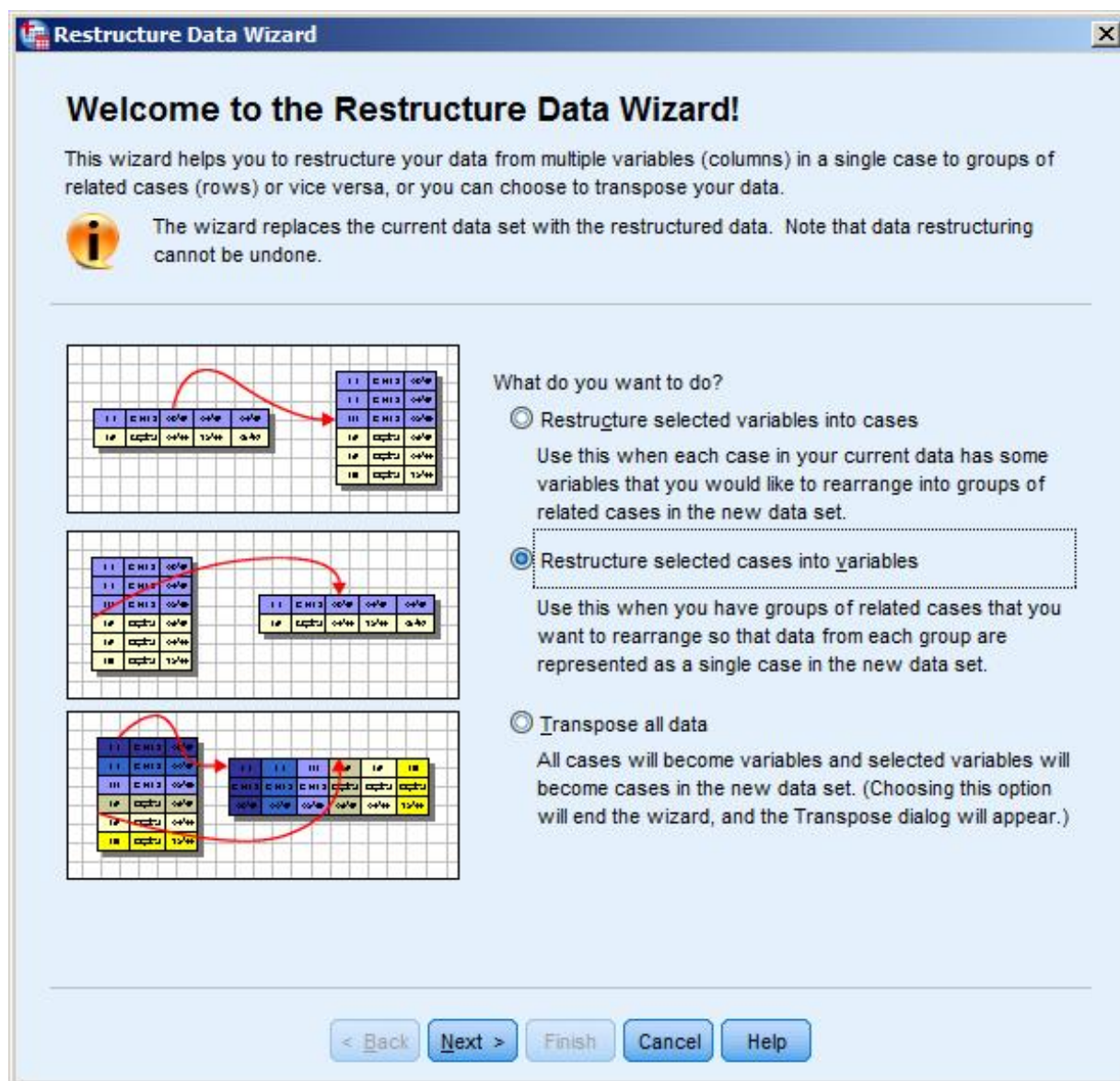
Step Two: In step two of five of the restructure wizard, you will select the variables that are the identifying variable and the index variable. The list of all variables from your active dataset appear in the box labeled "Variables in the current file." Each variable is represented by the variable label appearing first followed by the variable name enclosed in brackets. You will select and move the variable named "id" from the "Variables in the current file" list to the "Identifier Variables" box. Also, you will move the variable named "visit" from the "Variables in the current file" list to the "Index variable(s)" box (see figure 2 below). Click "Next."

Figure 2. A screenshot of step two of the Restructure Data Wizard.
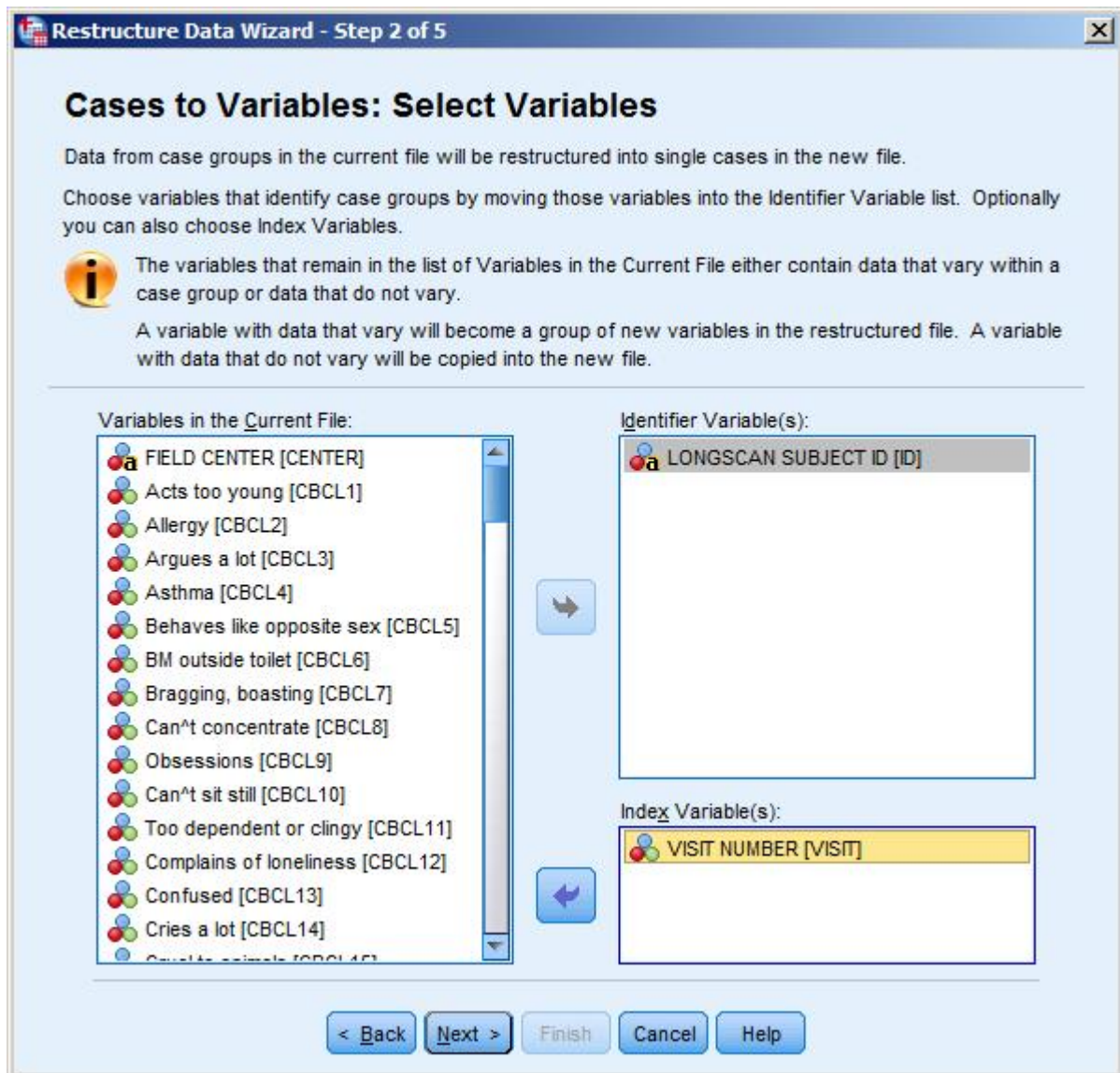
Step Three: In Step 3 of the Restructure Data Wizard, you will be asked if you want the data sorted by the identifier and index variables or if you want to use the data in the order in which they appeared prior to starting the Restructure Data Wizard. It is up to you regarding which to select. In Figure 3, I have selected "Yes, data will be sorted by the identifier and index variables." After you have made your selection, click "Next."

Figure 3. A screenshot of step three of the Restructure Data Wizard.
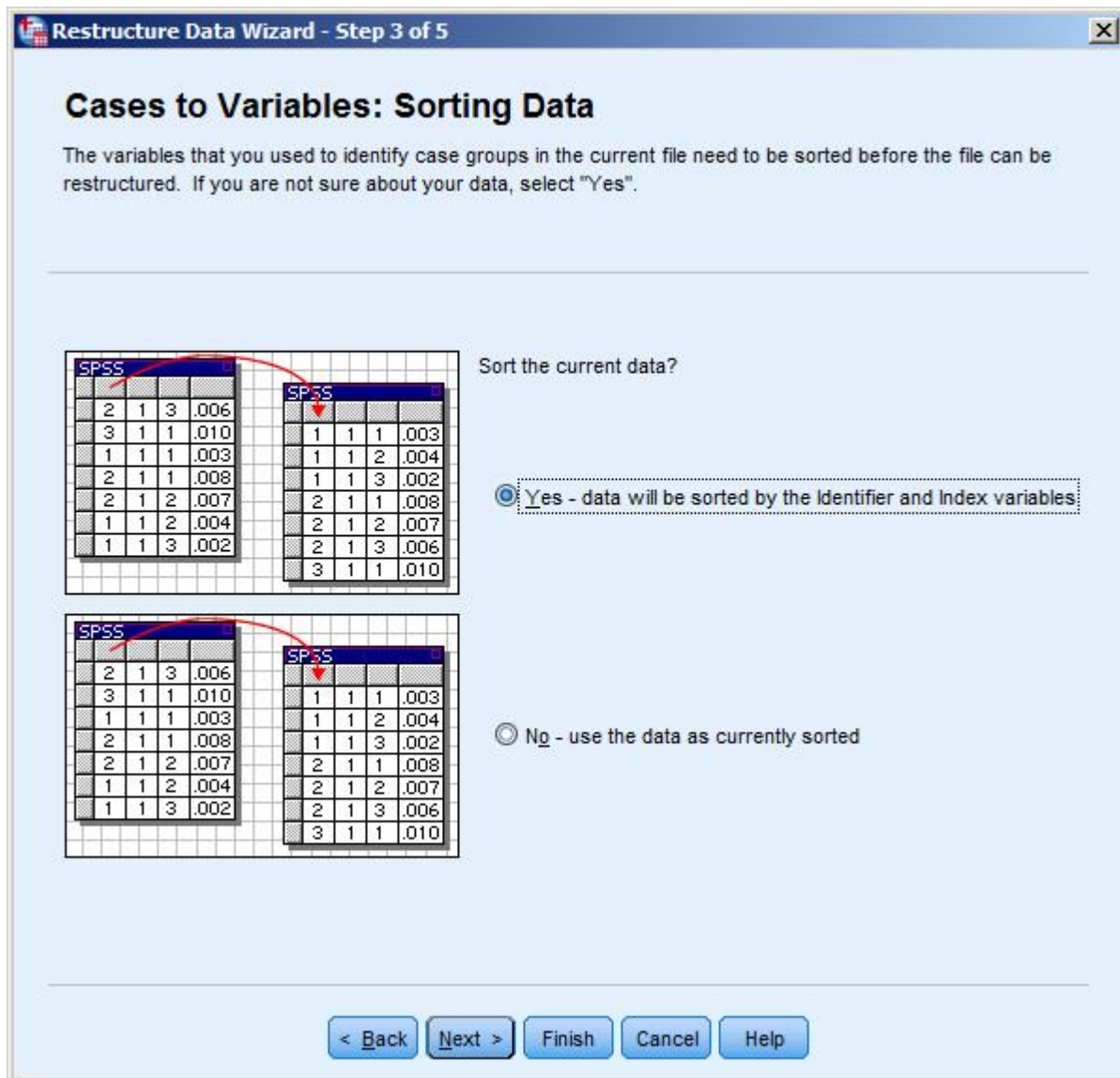
Step Four: In step 4, the Restructure Data Wizard will permit you to specify the order of the variable groups in the new data file and add a case count and index count variables to the new data file. The default selection in the "Order of new variable groups" section is "Group by original variable…" This selection works fine and will provide an order to the variables where each original variable will repeat in succession corresponding to the number of waves of data collection captured. So for example, all of the variables pertaining to question one of the CBCL will be grouped next to each other like this: CBCL1.4 CBLCL1.6 CBLC1.8. If you choose the second option called "Group by Index" then your new data file will have all of the age 4 variables grouped together followed by all of the age 6 variables grouped together and so on. It is a matter of preference as to which you choose. At this point, I will say that you can ignore the case count variable and index count variable options as they are not absolutely necessary to your

success in restructuring the data file (see Figure 4). After you have made your selection, click "Next."

Figure 4. A screenshot of step four of the Restructure Data wizard.
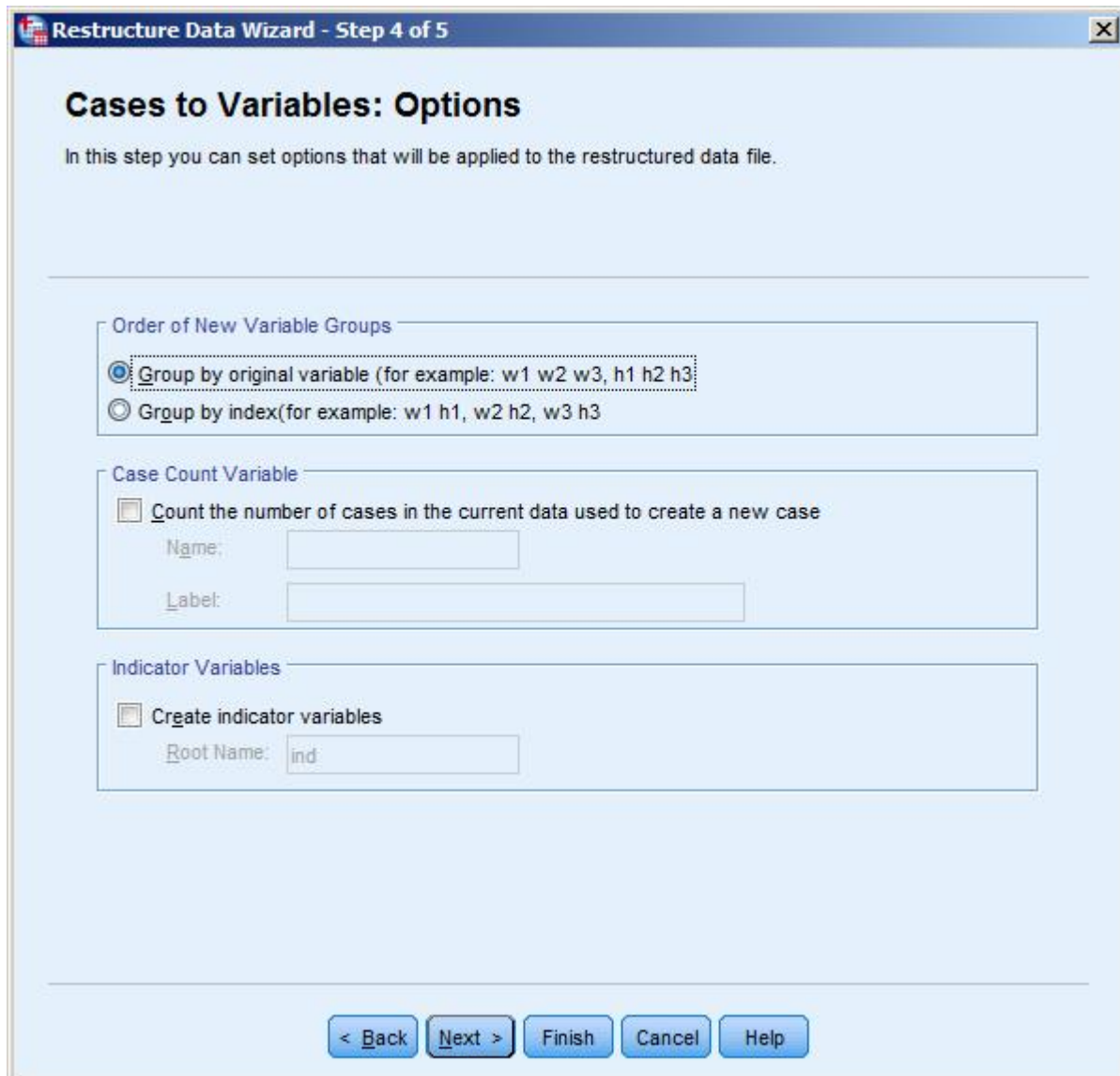
Lastly, the Restructure Data Wizard will ask you if you want to restructure the data now or paste the syntax generated by the Wizard into a syntax window to run later. If you want to move forward with the restructure, you can just select "Restructure data now" and then click "Finish" (see Figure 5). If you think you will need to go through the process of restructuring another data file, then it would be good to select "Paste the syntax generated by the wizard into a syntax window" option. By doing this, you can reuse the syntax over and over again without having to re-run the wizard. After you have made your selection, click "Finish."

Figure 5. A screenshot of the final step of the Restructure Data Wizard.
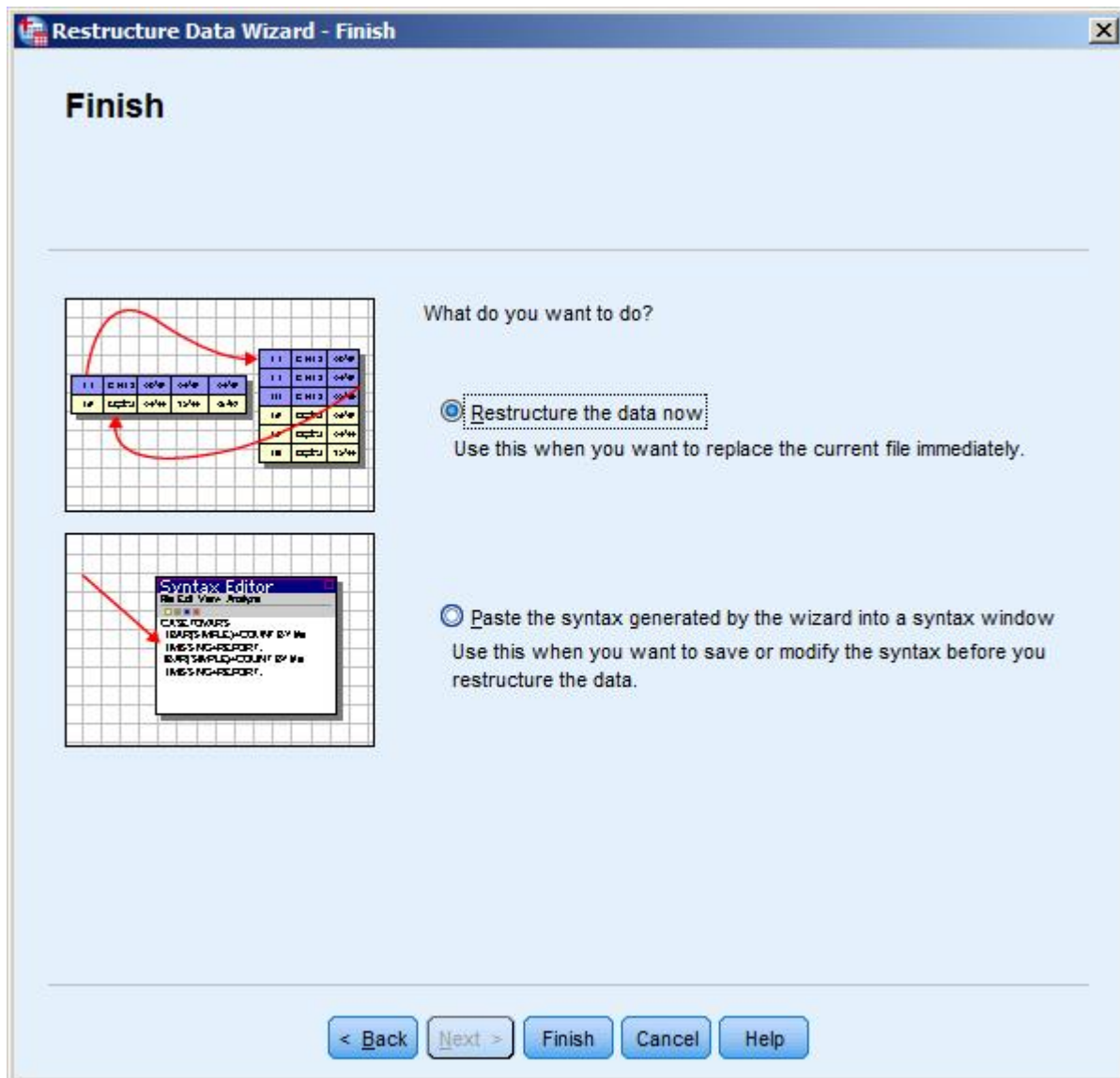
The new data file is now oriented as one-record-per-participant. Each participant's age related observation (age 8, 10, 12, etc.) is now stored as a separate variable. Here is how to read the new variable names: oldvariable.index   In the case of LONGSCAN, your Index variable is the age of administration, so for example, CBCL1 which is question one from the Child Behavior Checklist becomes, CBCL1.12 which represents question one from the CBCL at the age 12 administration of the measure.

Below, is the SPSS syntax for determining the structure of the data file and then performing the restructure to convert cases to variables or in other words change the file from multiple-records-per-respondent to one-record-per-respondent.

*note: Secondary analysts are cautioned against including variables related to gender and race/ethnicity in the FIXED variable list, as they can change across the waves of data collection.

FREQUENCIES id.

CASESTOVARS

/ID = id

/FIXED = center

/INDEX = visit.

EXECUTE.

This document was prepared by Research Analyst, Holly M. Larrabee, MSHP. Email questions to NDACANsupport@cornell.edu.