

**Progress in International Reading Literacy Study
(PIRLS) 2001
Electronic Codebook Quick Guide**

1. INTRODUCTION TO THE PROGRESS IN INTERNATIONAL READING LITERACY 2001 ELECTRONIC CODEBOOK

The Progress in International Reading Literacy 2001 Electronic Codebook (ECB) software tool enables analysts to review and extract United States PIRLS 2001 data. With this ECB software tool, an analyst can perform the following actions:

- Search the names and labels of variables in the PIRLS 2001 national data files (called catalogs) to select variables for analysis;
- Create a list of variables to be extracted from the catalog (called taglists), save the list for later use, print the list as a codebook, or use a predefined taglist;
- Examine the response categories, frequencies, and percentages of responses for one or more catalog variables; and
- Automatically generate SAS, SPSS for Windows, or STATA programs to extract selected variables from the whole data set or for a subset of defined cases.

The ECB software tool only works on personal computers running a Windows-based environment (Windows 95 or higher). It will not run with other computer operating systems, such as Macintosh and Linux. The ECB software tool includes the PIRLS 2001 national data files stored as flat (ASCII) data files on a separate data compact disc.

2. THE PROGRESS IN INTERNATIONAL READING LITERACY STUDY 2001 ELECTRONIC CODEBOOK TUTORIAL

This tutorial provides specific directions for using the Progress in International Reading Literacy Study (PIRLS) 2001 Electronic Codebook (ECB) software tool. The ECB software tool allows analysts to review and extract variables stored in the PIRLS 2001 national data files. With this ECB software tool, an analyst can perform the following actions:

- Search the names and labels of variables in the PIRLS 2001 national data files (called catalogs) to select variables for analysis;
- Create a list of variables to be extracted from the catalog (called a taglist), save the list for later use, print the list as a codebook, or use a predefined variable taglist;
- Examine the response categories, frequencies, and percentages of responses for one or more catalog variables; and
- Automatically generate SAS, SPSS for Windows, or STATA programs that extract selected variables from the whole data set or for a subset of defined cases.

This ECB software tool only works on personal computers running a Microsoft Windows operating system (Windows 95 and later versions). It will not run in any other operating system environments, such as Macintosh.

Starting the ECB

1. From the desktop, click on the **Start** button, select **Programs**, and click on **Electronic Codebook**. Double clicking on the Electronic Codebook icon on the desktop will also invoke the software.
2. After the initial splash-screen appears, the Select Catalog screen will appear. Highlight the catalog (data file) to examine. Click the **OK** button.
3. The Main ECB screen now appears (Figure 1). The Variable List comprises the left part of the screen, while the Working Taglist comprises the right part of the screen. The Title Bar sits on top of the screen, while the Menu Bar sits underneath the Title Bar.

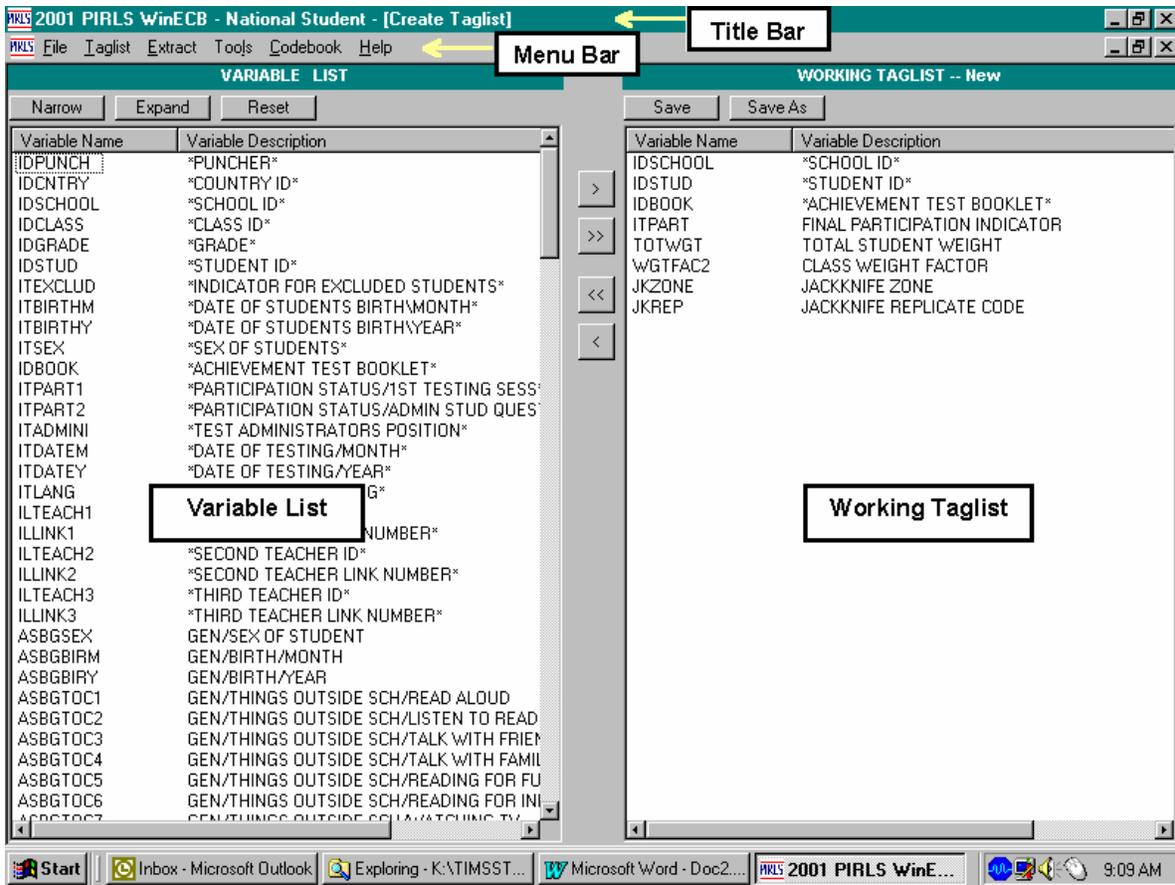


Figure 1. Main ECB Screen

Main ECB Screen

The **Variable List** appears on the left half of the Main ECB screen. It displays all variables for the catalog. Besides the variable names, the list also displays a brief description of the variables. From the variable list, an analyst can view codebook information for a variable, narrow or expand the variable list, or tag variables for extraction.

The **Working Taglist** appears on the right half of the Main ECB screen. It displays a list of variables that are currently selected (tagged) for extraction. The initial working taglist consists of a set of required variables for that catalog. Analysts may alter the working taglist by adding and dropping variables from the variable list, but they may not drop (untag) required variables. From the working taglist, an analyst can view codebook information for a tagged variable, delete variables from the taglist, use predefined taglists, view codebook information for a taglist, and save taglists.

The **Title Bar** is the horizontal bar located at the top of the Main ECB screen. Common to all Windows-based programs, it gives the name of the application.

The **Menu Bar** lies underneath the title bar. The bar contains a series of pull-down menus (e.g., File, Taglist, Extract, Tools, Codebook, and Help). Selecting items from the pull-down menus provide access to the action commands available in the ECB.

To access the menus, follow these steps:

1. Point to an item on the menu bar and click.
2. Click on a command from the drop-down list.

Keystrokes can also be used to access menu bar items. All of the menus have one letter underlined; this letter is the hot key. Pressing <ALT> and the underlined letter at the same time opens that menu. Within the menu, use the arrow keys to move to menu items. Press the up and down arrows to move to the desired command and press <ENTER>. An item that appears in dimmed print is not available for the current screen.

Variable List Functions

Viewing Codebook Information for a Variable

To view codebook information for a variable, follow these steps:

1. Scroll through the variable list until the desired variable is found.
2. Click on the variable name to highlight it. The variable name is the only active field in the window. No action will be taken by clicking on the variable description.
3. Press <ENTER> or double-click on the highlighted variable name.
4. The Variable Quick View window opens, containing the codebook information for the highlighted variable.
5. To close the Variable Quick View window, click on the **X** button in the upper right-hand corner of the window. The Main ECB screen will reappear.

Information displayed in Variable Quick View window includes the variable name and label, any question number and wording associated with the variable, the record number, position, and format of the variable on the data file, and any comment fields. The Variable Quick View window also displays values of discrete variables, ranges for continuous variables, unweighted and weighted frequencies, and unweighted and weighted percentage distributions. Note that the Electronic Codebook may not contain all this information for each catalog.

Narrowing or Expanding the Variable List

To narrow the list of variables to include only variables that contain a specific keyword, follow these steps:

1. Click the **Narrow** button under the menu bar. The Narrow Text dialog box will appear.
2. Type the keyword in the text box underneath “**Enter Narrow Text.**”
3. Click the circle beside **Variable Name**, **Variable Description**, or **Both Variable Name and Description** to specify where to search for the keyword.

4. Click the **Search** button.

To add to the list of variables any catalog variables that contain a specific keyword, follow these steps:

1. Click the **Expand** button under the menu bar. The Expand Text dialog box will appear.
2. Type the keyword in the text box underneath “**Enter Expand Text.**”
3. Click the circle beside **Variable Name**, **Variable Description**, or **Both Variable Name and Description** to specify where to search for the keyword.
4. Click the **Search** button.

Because the initial variable list contains all variables for a catalog, selecting the Expand option is not necessary unless the original variable list has been narrowed. Note that altering the variable list has no effect on the working taglist.

Clicking the **Reset** button under the menu bar refreshes the variable list to show all catalog variables.

If a keyword specified in the search to narrow or expand the variable list does not appear in any of the catalog variable name and/or descriptions, the No Matches Found message box will appear and the variable list will disappear from the screen. Simply click the **OK** button, then the **Reset** button to retrieve all the catalog variables.

Tagging Variables for Extraction

While browsing the variable list, a variable can be tagged for extraction in the following ways:

1. To tag one or more variables, highlight the variables in the variable list. Click the > button on the vertical bar that separates the variable list and working taglist.
2. To tag all of the variables at one time, click the >> button.

Working Taglist Functions

Viewing Codebook Information for a Tagged Variable

To view codebook information for a tagged variable, follow these steps:

1. Scroll through the working taglist until the desired variable is found.
2. Click on the variable name to highlight it. The variable name is the only active field in the window. No action will be taken by clicking on the variable description.
3. Press <ENTER> or double-click on the highlighted variable name.

4. The Variable Quick View window opens, containing the codebook information for the highlighted variable.
5. To close the Variable Quick View window, click on the **X** button in the upper right-hand corner of the window. The Main ECB screen will reappear.

Deleting Variables from the Working Taglist

A variable can be deleted (untagged) from the working taglist in the following ways:

1. To untag one or more variables, highlight the variables in the working taglist. Click the < button.

Note that each catalog contains a set of required variables for extraction. If one of these required variables is selected for deletion from the working taglist, the Required Variable message box will appear. Click **OK** to clear the message box and return to the working taglist.

2. To untag all of the non-required variables at one time, click the << button.

Using Predefined Taglists

To aid analysts, one or more taglists may have been predefined for each catalog. These predefined taglists can be accessed from the **Taglist** menu on the menu bar. User-defined taglists can also be accessed from the menu.

Select the **Add** command from the Taglist menu to expand the working taglist with additional variables from a predefined or user-defined taglist. A dialog box containing all predefined and user-defined taglists associated with the catalog will appear. Click on the desired taglist name, and press the **OK** button to start the merge. Additional taglists may be added, but they must be added one taglist at a time.

Select the **Open** command from the Taglist menu to replace the current working taglist with a predefined or user-defined taglist. A prompt to save the current working taglist will appear; see the **Saving Taglists** section for more information about saving taglists. All previously tagged variables, except for the required variables, will disappear from the working taglist. A dialog box containing all predefined and user-defined taglists associated with that catalog will appear. Click on the desired taglist name, and press the **OK** button. The variables from the highlighted taglist will appear in the working taglist.

Select the **New** command from the Taglist menu to return the working taglist to its original state, eliminating all catalog variables except the required variables. If the current working taglist has been modified, a prompt to save the current working taglist will appear; see the **Saving Taglists** section for more information about saving taglists.

Select the **Delete** command from the Taglist menu to remove any user-defined taglists. A dialog box containing all user-defined taglists associated with that catalog will appear. Click on the desired taglist name, and press the **OK** button to delete the list. A message box will appear to verify permanent deletion of the taglist. Press the **Yes** button to delete the taglist, or press the **No**

button to keep the taglist. If there are no user-defined taglists associated with the catalog, an error message box will appear. Click **OK** to get rid of the message box. Note that predefined taglists can not be deleted.

Viewing Codebook Information for a Taglist

To view a codebook for the current working taglist, follow these steps:

1. Select the **Codebook** menu from the menu bar.
2. Choose the **View** command from the menu. A View Codebook screen (Figure 2) will appear.

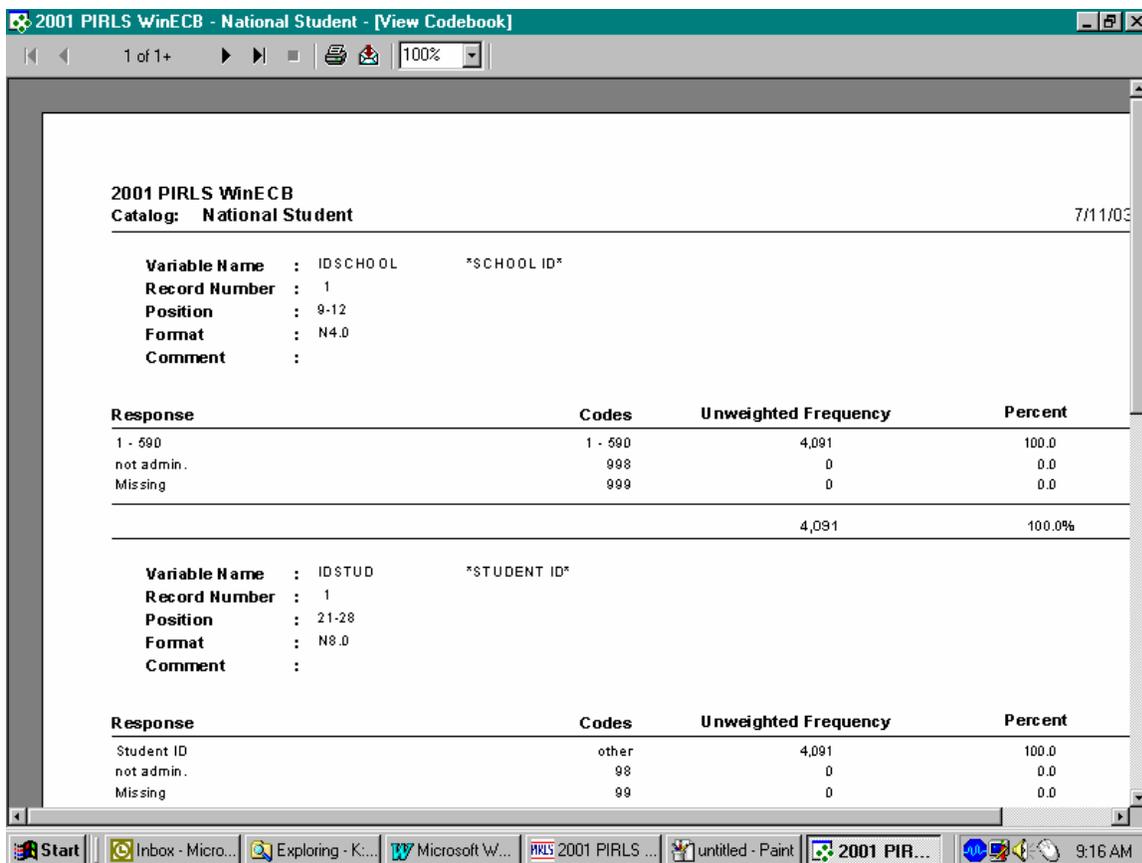


Figure 2. View Codebook Screen

The View Codebook screen has a toolbar at the top. Use the arrow icons on this toolbar to move through the codebook. Click the right arrow to move to the next codebook page, and click on the left arrow to move to the previous codebook page. Click the arrow icons containing a vertical line to move either to the first page (left arrow) or last page (right arrow) of the codebook.

The numbers between the arrow buttons on the toolbar are page counters. If a “+” follows the page number, there are more pages to be seen in the codebook. On the last

page of the codebook, the plus sign disappears, and the number of pages in the current codebook will be displayed. The stop icon (the button with the black square on it) is active when loading tagged items for viewing codebook information, when paging through the codebook, and when exporting the codebook. The scroll bar at the far right of the screen allows you to scroll the current page of the codebook, but it does not scroll from page to page. The arrow keys on the keyboard can also be used to scroll up or down the current page.

Information displayed in the View Codebook screen includes the variable name and label, any question number and wording associated with the variable, the record number, position, and format of the variable on the data file, and any comment fields. The View Codebook screen also displays values of discrete variables, ranges for continuous variables, unweighted and weighted frequencies, and unweighted and weighted percentage distributions. Note that the Electronic Codebook may not contain all this information for each catalog.

Icons for printing and exporting are also available on the menu bar, as is a drop-down box for zooming in and out of the codebook view.

Click the Print icon (the button with the printer image) to send the displayed codebook to the default printer. The Print dialog box then will appear. Select the **Print Range** and number of **Copies** required, then click on the **OK** button to print the codebook.

Click the Export icon (the button with the envelope image) to export the displayed codebook in a format that can be read by another software application. The Export dialog box then will appear. The codebook can be exported to one of the three format choices available on the screen. Selecting Crystal Reports saves the data in a format that can be opened in Crystal Reports. The resulting file will have the extension *.rpt. Rich Text Format saves all formatting and converts it to instructions that other programs, including Microsoft compatible programs, can read and interpret. The resulting file will have the extension *.rtf. Word for Windows saves the document as a Word document; the file created will have the extension *.doc.

The only Destination option in the Export dialog box is to a disk file. Possible locations for the disk file cover floppy disk drives, read/write CD-ROM drives, hard drives, and any networked drives connected to your computer during the ECB session.

Saving Taglists

To save a new taglist, follow these steps:

1. Click on the **Taglist** menu from the menu bar, and then select the **Save As** command.
2. A dialog box will appear that lists the names of user-defined taglists associated with the catalog. To replace a user-defined taglist with the current taglist, click on the name of the user defined taglist to be replaced. To save the taglist to a new name, enter the name in the Taglist Name box.
3. Press the <ENTER> key to save the taglist with the new name.

4. If a taglist with that name exists, a verification dialog box will appear. Click the **Yes** button only for replacement of the old taglist with the new taglist. Click the **No** button to keep the original taglist. Predefined taglists cannot be replaced.

To save changes to a user defined taglist currently residing in the working taglist, click on the **Save** button above the working taglist. It will save the modified taglist under the original user defined taglist's name. Clicking on the **Taglist** menu and selecting the **Save** command will accomplish the same task.

Extracting Data

A primary function of the ECB software is that it generates SAS, SPSS for Windows, or STATA programming code necessary to extract the variables listed in the working taglist from the project's data files. The **Extract** menu on the menu bar accomplishes this task.

The **Extract** menu has three options. Selecting **SAS** will generate SAS program code to a file named by the analyst. The *.sas extension will be automatically added. Selecting **SPSS for Windows** will generate SPSS program code to a file named by the analyst, with the *.sps extension automatically added. Selecting **STATA** will generate STATA program code to two files named by the analyst, with the extensions *.dct and *.do automatically added. Analysts will need SAS, SPSS for Windows, or STATA software to run the extract programs. Reviewing an extract program before running it is recommended because it may need to be customized.

Creating an Extract Program File

To create an extract program file:

1. Click on the **Extract** menu on the menu bar.
2. Click on the statistical software package in which to create the extract program.
3. The Extract Specifications dialog box will appear. Choose the subset of catalog records that the extract program will use. See **Defining the Extract Population** for more information about subsetting catalog records.
4. The Save As dialog box will next appear. Enter a file name for the extract program. The appropriate extension (*.sas for SAS programs, *.sps for SPSS for Windows programs, or *.dct and *.do for STATA programs) will be added automatically. Click on the **Save** button.
5. If a file name that already exists is entered, a verification message box will appear. Click the **Yes** button to overwrite the existing file with the new file. Click the **No** button to prompt the input of a new name under which to save the file.
6. The Save Data File As dialog box will next appear. Specify the name for the data file that will be created. The extension (*.sd2 for SAS data sets, *.sav for SPSS for Windows data sets, and *.dta for STATA data sets) will be added automatically. If a different folder for

saved extract data files is desired, select one from the Save In drop-down menu. A new folder can be created by clicking on the icon showing a folder with a star behind it.

7. Enter a file name where the extracted data will be stored in the Save Data File As dialog box. The default name for the data file is the name of the extract file. If another name is desired, enter it in the file name text box.
8. Click on the **Save** button.

Defining the Extract Population

After specifying the statistical software in which to create the extract program, the Extract Specifications window appears. This window displays various subgroups of respondents typically of interest to analysts. Selecting specific subgroups allows the extraction of information for only the population of interest to the analyst, thereby minimizing the size of the data file created. The default extract population includes all catalog records.

Defining the subgroups of respondents for extraction depends on whether the subsetting variables are either *continuous variables* or *discrete variables*. Continuous variables, as defined by the ECB program, are catalog variables with an infinite number of possible formatted values. Discrete variables are catalog variables with a finite number of possible formatted values.

Continuous variables, as displayed in the Extract Specifications window, have values in the **Low** and **High** range columns and a value of “n/a” (not applicable) in the **Select** column. To specify a range of values other than the default range:

1. Click on the low or high value to highlight that field.
2. The number in the field appears in the text box at the top of the window.
3. Click that text box and delete the number in it.
4. Type the low or high value desired.

Another method of changing the range of values:

1. Double-click on either the low or high value.
2. Type the appropriate value desired.

Discrete variables, as displayed in the Extract Specifications window, do not use a range of values to subset the data file, as designated by the “n/a” (not applicable) values in the **Low** and **High** range columns. Discrete variables instead start with a “Yes” next to each of their possible codes in the **Select** column. To exclude certain codes from the extract:

1. Double-click on a variable code in the **Select** column to change the value from “Yes” to “No;” or
2. Click on a variable code in the **Select** column, and then press the N key or the space bar to change the value from “Yes” to “No.”

The Extract Specifications window also allows for defining the drive where the extract program will find the project data files. In the CD-ROM Drive drop-down box (lower left corner of the screen), the drive in which the software will look for data is displayed. Change the drive letter if it does not correspond to the correct drive where the extract program will find the data file.

After selecting the extract criteria, click on the **OK** button to continue creating the extract program.

Reviewing the Extract Program

It is recommended to review the generated SAS, SPSS, or STATA program code before running it. Check that the line of code that defines the input data points to the data file's correct directory. Check also that the data subsetting statements are correct. Code for generating frequencies and means are written in the program but "commented out." Remove the comment symbols to obtain that output.

It is easy to correct mistakes discovered after writing out or running the extract program if the working taglist was saved before exiting the ECB. Simply restart the ECB, select the appropriate catalog, open the saved taglist, then go through the extract process again, selecting the desired subset of cases and writing out a new extract program.

Repair/Compact

The ECB program utilizes a relational database that sits behind its screens and contains the information for all the catalogs and their variables. If many taglists are created and deleted on a regular basis, then the database will contain lingering references to old taglists that are no longer needed. Repairing and compacting of files allows the ECB program to "clean house" and make the database more efficient. It also decreases the size of the database, so space is conserved.

To repair and compact files:

1. Click on the **Tools** menu on the menu bar.
2. Select **Repair and Compact Database**.
3. The file repair and compact process will begin. As it runs, the ECB program will display an hourglass figure.
4. A confirmation message box will appear when the repair and compact process ends. Click the **OK** button to close the message box.

Switching Catalogs

After completing work with one catalog, it is possible to access another catalog without quitting the ECB.

1. Click on the **File** menu on the menu bar.

2. Select the **Open Catalog** command. If the working taglist has not been saved, a Save Changes message box will appear. To save changes to the current working taglist, press the **Yes** button. Otherwise, press the **No** button. Pressing the **Cancel** button returns the Main ECB screen.
3. The Select Catalog screen will reappear. Choose another catalog with which to work.

Exiting the ECB Program

To exit the ECB program:

1. Click on the **File** menu on the menu bar.
2. Select the **Exit** command. If the working taglist has not been saved, a Save Changes message box will appear. To save changes to the current working taglist, press the **Yes** button. Otherwise, press the **No** button. Pressing the **Cancel** button returns the Main ECB screen.

Clicking on the square **X** button in the top right-hand corner of the Main ECB screen will also exit the ECB program.

A Sample ECB Session

The script below provides an example of a session with the PIRLS 2001 ECB software. In this example, a user wants to create a file of variables associated with girls' and boys' computer use in school. The actions of the user are shown below in italics.

Double-click on the PIRLS 2001 ECB icon on the computer desktop.

The PIRLS 2001 ECB splash screen flashes on the computer screen. The Select Catalog screen then appears.

Double-click on the PIRLS 2001 Student Catalog in the Select Catalog screen.

The variable list and initial working taglist associated with the PIRLS 2001 Student catalog appear on the screen. The working taglist contains the variables required for all student catalog extracts.

Click on the Narrow button in the Variable List window.

Type the string PC into the box, change the search option to Both Variable Name and Description, and click the Search button.

Several variables appear pertaining to personnel computer use. The user is specifically interested in computer use in school.

Double click on the variable ASBGPCU2 to read the full text of the item.

A Variable Quick View window for the ASBGPCU2 variable opens. The item (PC use at school), its response categories, frequencies, and percentages are displayed.

Click on the X button at the top right-hand corner of the Variable Quick View window

The user decides to add the variable to the taglist by clicking the single arrow key between the variable list and the working taglist.

The user knows that the student catalog also contains the plausible values variables for the achievement scores, but does not know the name of the variables. The user decides to search for this variable on the student catalog's variable list.

Click on the Expand button in the Variable List window.

Type the string PLAUSIBLE VALUE into the box, change the search option to Both Variable Name and Description, and click the Search button.

The ECB now searches all variable names and labels for the string PLAUSIBLE VALUE. Three sets of plausible value variables now appear in the variable list. There are sets for the information (ASRINF01-ASRINF05) and literary (ASRLIT01-ASRLIT05) subscales and the combined scale (ASRREAD01-ASRREAD05). Note that the narrowed variable list does not affect the working taglist.

The user is interested in the combined plausible values and adds these to the working tag list.

Highlight the variables ASRREAD01-ASRREAD05 in the Variable List window by pressing the Ctrl key and clicking on the variable name.

Click on the right arrow button (>) on the bar between the Variable List and Working Taglist windows. The five plausible value variables have been added to the working taglist.

Finally, the user searches for a variable measuring the sex of the student.

Click on the Expand button in the Variable List window.

Type SEX into the box, change the search option to Both Variable Name and Description, and click the Search button.

The ECB now searches all variable names and labels for SEX. Two variables – ITSEX and ASBGSEX are added to the variable list.

To decide which variable to use, the user decides to look at their codebook information.

Double-click the ITSEX variable name in the variable list.

A Variable Quick View window for the ITSEX variable opens. The item (sex of students), its response categories, frequencies, and percentages are displayed. This variable is taken from the student sampling form and is not student reported. The user decides to look at the other variable ASBGSEX.

Click on the X button at the top right-hand corner of the Variable Quick View window

Double-click the ASBGSEX variable name in the variable list.

This variable is the student report of their sex. The user decides to use ITSEX because it has no missing data.

Click on the X button at the top right-hand corner of the Variable Quick View window.

Highlight the ITSEX variable in the Variable List window by clicking on the variable name.

Click on the right arrow button (>) on the bar between the Variable List and Working Taglist windows.

The working taglist now includes the tagged ITSEX variable.

The user decides not to add any other variables to the working taglist. This set of variables will compose the extract file. Since this set of variables may be needed in future analyses, the user saves the current working taglist to a user-defined taglist called MYLIST.

Click on the Taglist menu on the menu bar, and click on the Save As option.

Type the new name MYLIST into the Taglist Name box and click the Save button.

With the taglist saved, the user now wants to create an extract program from the taglist. The program will contain SAS statements that will extract only the currently tagged variables. The user wants to save these SAS program statements in a file named MYEXAMPL.SAS.

Click on the Extract menu on the menu bar, and click on the SAS option.

The Extract Specifications window will appear. It contains a variable that can be used to subset the student file – student participation (IDPART). This window is required by the software and it is suggested that the user simply bypass this screen by clicking OK. To allow for full flexibility, modifications to variables should be conducted in the extract program.

Click on the OK button at the bottom of the window.

The Save As dialog box now appears.

Type the new name MYEXAMPL into the File name box, select the directory where the file should be saved, then click the Save button.

The SAS extract program will have the name MYEXAMPL.SAS in the specified directory. The Save Data File As dialog box now appears.

Click the Save button in the Save Data File As dialog box.

The SAS extract program will now create a SAS database called MYEXAMPL.SD2 that will store the variables the user selected in the working taglist.

With the SAS extract program written and saved, the user can now quit the ECB and run the SAS statistical software package. Using the program statements in MYEXAMPL.SAS, the software

will extract a subset file of records and variables from the PIRLS 2001 national student data file. The user can also revise the SAS programming code to perform additional statistical analyses. *To quit the ECB, click the File menu on the menu bar then click on the Exit option.*

3. ELECTRONIC CODEBOOK (ECB) TROUBLESHOOTING GUIDE

What hardware and software requirements are needed to run the ECB software?

The Electronic Codebook (ECB) software is designed to run under Windows 95 and Windows 98. The ECB requires approximately 18 MB (megabytes or million bytes) of available disk space on the hard drive. If 18 MB are not available, delete files from the hard drive to make space for the ECB. After being installed, the ECB can create and save personal taglists. These taglists will require minimal disk space. The ECB can also create SAS, SPSS for Windows, and STATA extract programs; these extract programs will require little disk space.

How do I uninstall/reinstall the ECB software from my computer?

Depending on the configuration of the computer it may be necessary to remove the software by finding the ECB setup executable in a subdirectory on the local disc. This will be located in **C:\Program Files\InstallShield Installation Information**. Within this subdirectory there will be a folder containing configuration settings and additional files as well as the setup executable. If the user has installed multiple ECBs there will be one folder for each ECB. The user can determine which folder is assigned to the proper ECB by double clicking the **Setup.exe** file. Once this is determined, the user can remove the ECB by following the prompts. This process may automatically restart the computer. After the computer has been restarted the ECB may be reinstalled from the CD-ROM.

Why are the ECB window screens so hard to read? Why are the windows too small (large)?

The ECB program fits best visually on screens set to a desktop area of 800 x 600 pixels. It will still work on other screen settings, but it may not make the best use of the available screen space. To check and/or set your desktop area, follow these steps:

1. Click on the **Start** button.
2. Select **Settings, Control Panel**, and then **Display**.
3. Select the **Settings** tab.
4. Set the **Desktop Area** to 800 x 600 with the Desktop Area sidebar.

What do I need to run the extract programs generated by the ECB software?

Running the extract programs will require the appropriate personal computer-based statistical software package (SAS, SPSS for Windows, STATA) and the project data files from which the program can extract data. The ECB does not create a SAS, SPSS, or STATA data file. It simply prepares the statements that can be used with these statistical software packages to create those files.

Why can I not delete some variables from the working taglist?

A catalog may contain one or more "required variables." Required variables are deemed necessary for any analysis of that catalog's data, and they will be automatically included in all extracts. If a required variable is selected for removal from the working taglist, an error message box will appear. Click the **OK** button to remove the message box.

Why can I not combine variables from two or more catalogs in a working taglist?

By its design, the ECB program examines only one catalog at a time. Altering the current program to allow combinations of variables from two or more catalogs introduces severe errors in both the codebooks and extract programs. Only if the catalogs have a one-to-one relationship, i.e. the record in position *n* of each catalog refers to the same entity, would the frequencies and percentages of the merged variable codebook be both accurate and meaningful. Extract programs with merged variables could contain data input errors, formatting errors, and frequency generation errors. There are no plans at this time to allow merging of two or more catalogs in future versions of the ECB software.

Creating a database with variables from two or more catalogs can be done with the current ECB software, although it requires some user-written programming code. First, for each catalog, generate an extract program that includes only those variables needed in the merged database. Second, run these extract programs on the appropriate statistical software package. Finally, write and run an additional program that merges the resulting databases into a single database.

How many user defined taglists can you save for each catalog?

Each catalog can have an unlimited number of saved user defined taglists, dependent only upon the memory limitations of the hard drive. If many taglists are created and deleted on a regular basis, then periodically repair and compact the Microsoft Access database the ECB program uses for all the catalogs and their variables. Taglist creation and deletion will lead to the database containing lingering references to old taglists that are no longer needed. Repairing and compacting files allows the ECB program to "clean house" and make the database more efficient. It also decreases the size of the database, so space is conserved.

For certain variables, why does the codebook list erroneous or nonsensical values in the Codes column?

Printing limitations may cause what looks to be erroneous or nonsensical data printed in the Codes column of the codebook. The ECB program uses the Crystal Reports report generation software to produce its codebooks. The codebook's printed format follows a layout designed in the Crystal Reports software. In the current codebook layout, the Codes column lists the actual value - or range of values - of a variable associated with the response category, which is listed in the Response column. Data for this column comes from the Format Value field in the Format Master table of the Microsoft Access database used by the ECB program. The Codes column has a maximum length of 12 characters. If the value(s) listed in the Format Value field take up more than 12 characters, then the Codes column only displays the last 12 characters of the field. It is possible, but not likely, that similar printing errors could occur in other fields and columns listed in the codebook. Future versions of the ECB software may contain a revised Crystal Reports codebook layout that increases the size of the Codes column.

Why can I not do cross tabulations of two or more catalog variables in a codebook?

The ECB does not actually calculate either the frequencies or percentages of variables listed in a codebook. Those values, both unweighted and weighted, are stored in the Frequency Master table within the project's Microsoft Access database read by the ECB program. Once a variable is tagged for a codebook, the ECB program searches the Access database tables, including the Frequency Master table, to collect the codebook data for that variable. It then plugs the data into a Crystal Reports report format to produce the codebook.

To produce cross tabulations of two or more catalog variables, start by adding those variables to the working taglist. Generate an extract program with those variables. Before running the extract program, modify the code that produces the frequencies to include the cross tabulations. If several sets of cross tabulations are required, it would be a wise idea to save a working taglist that includes all of the variables listed in the cross tabulations for any future analyses.

Why limit the choices of codebook export files?

An exported codebook can be saved in one of the three following formats: Crystal Reports, Rich Text Format, and Word for Windows. For other formats (like Excel, WordPerfect, etc.), save the exported codebook with Rich Text Format. Rich Text Format saves all format commands in the codebook and converts it to instructions that other programs, including Microsoft compatible programs, can read and interpret. Future versions of the ECB software may include additional options for codebook export files.

Can I subset a catalog on variables not listed on the Extract Specifications screen?

Not while generating an extract program. The ECB program utilizes a Microsoft Access database to store information about each catalog and its variables. Within the database, the Variable Master table lists which variables can be used to subset each catalog. If one of those variables appears on the working taglist, then it will appear on the Extract Specifications screen. If the variable in question is not marked as a subsetting variable in the Variable Master table, then that variable will never appear on the Extract Specifications screen.

It is possible to extract a subset of catalog records based on a variable not found on the Extract Specifications screen. Create an extract program from a working list that includes the subset variable. Before running the extract program, modify the programming code to include that variable in the subsetting commands. Use the values listed in the Codes column of that variable from the codebook to determine how to subset the catalog.

Why do the ECB-generated extract programs not run properly?

It is recommended to review the generated SAS, SPSS, or STATA program code before running it. The programs assume that the project data files used by the extraction program will be found on a CD-ROM loaded in the CD-ROM drive. If the project data files are in another location (hard drive, network drive, etc.), change the code that specifies where the project data file is

located. At the top of the SAS program, edit the line that defines the macro variable ROOT. An example of that SAS line of code follows:

```
%LET ROOT = q:\data\ ;
```

At the top of the SPSS program, edit the line that assigns the handle FHAND to the input data set. An example of that SPSS line of code follows:

```
FILE HANDLE FHAND /NAME='q:\data\school.dat' /LRECL=91.
```

At the top of the STATA dictionary program (*.dct), edit the dictionary header line. An example of that STATA header line follows:

```
DICTIONARY USING "q:\data\school.dat" {
```

For each case, replace the directory listed in the specified line of code ("q:\data\" in the examples above) with the correct data directory.

When reviewing the extract programs, check also that any statements subsetting the data are correct. Code for generating frequencies and means are written in the program but "commented out." Remove the comment symbols to obtain that output.

4. ELECTRONIC CODEBOOK (ECB) GLOSSARY

Cancel A command button found in several windows and screens. Closes the current screen or window.

Catalog An ECB data file, usually corresponding with a project data file. Each catalog has a variable list and one or more predefined taglists.

Codebook (1) A detailed variable listing that can be either viewed or printed. Information displayed in a codebook includes variable name, variable label, any question number and wording associated with the variable, the record number, position, and format of the variable on the data file, and any comment fields. It also displays response categories, unweighted and weighted frequencies, and unweighted and weighted percentage distributions. (2) A menu listed on the menu bar that allows users to view and print codebooks.

Continuous Variable A catalog variable with a infinite number of possible values. Within a codebook or extract file, the response categories of continuous variables are represented by ranges of values.

Discrete Variable A catalog variable with a finite number of possible values. Within a codebook or extract file, the response categories of discrete variables are represented either by single values or ranges of values.

ECB Electronic Codebook. A software package that allows analysts a simple computerized interface to examine variables within a project's data files (called catalogs).

Expand A command button in the Variable List window. Adds to the current variable list any variables for the catalog whose name and/or description contains a specified keyword.

Export A command button in the View Codebook screen. Creates a copy of the codebook as either a Crystal Reports file, Microsoft Word file, or a rich text format (RTF) text file.

Extract (1) A subset of records and/or variables within a catalog. (2) A menu listed on the menu bar that allows users to create a statistical program (SAS, SPSS, STATA) that will pull out a defined subset of records and variables from a catalog.

Extract Program A software program that creates an extract from a catalog. The ECB software package can create an extract program for the SAS, SPSS, or STATA statistical software packages.

Extract Specifications A blueprint of what variables and/or records to choose for an extract. The Working Taglist defines which variables will appear in the extract, while the Extract Specifications screen limits the number of records in an extract based on selected response categories within subsetting variables.

File A menu listed on the menu bar that allows users to open a catalog, define the printer settings, or exit the ECB software.

Frequency The number of times the value(s) of a variable appears within a catalog.

Help A menu listed on the menu bar that calls up the ECB software's electronic help file.

Menu Bar The bar that lies underneath the title bar on the main ECB screen. The menu bar contains a series of pull-down menus (e.g., File, Taglist, Extract, Tools, Codebook, and Help). Selecting items from the pull-down menus provide access to the action commands available in the ECB.

Narrow A command button in the Variable List window. Keeps only those variables in the current variable list whose name and/or description contains a specified keyword.

Percentage A relative measure of how often the value(s) of a variable appears within a catalog as compared to all values of that variable.

Predefined Taglist A saved list of catalog variables that came with the original ECB software package. It was created in an earlier ECB session, or was created and saved earlier in the current ECB session.

Print A command button in the View Codebook screen. Sends a copy of the codebook to a printer connected to the computer running the ECB software.

Repair/Compact An option under the Tools menu on the menu bar. Eliminates lingering references to deleted taglists within a catalog, which also decreases the size of the database.

Reset A command button in the Variable List window. Restores all of the variables within a catalog to the current variable list.

Response Category A possible value, or set of possible values, for a catalog variable, as defined by the ECB software. For continuous variables, which have an infinite set of possible values, a category will list a range of values. For discrete variables, which have a finite set of possible values, a category most likely will represent a single value, but it can also represent a range of values.

SAS One of three software statistical packages in which the ECB software creates extract programs. A SAS extract program will have an *.sas extension.

Save A common button in the Working Taglist window. Stores the working taglist as a user defined taglist.

SPSS One of three software statistical packages in which the ECB software creates extract programs. An SPSS for Windows extract program will have an *.sps extension.

STATA One of three statistical software packages in which the ECB software creates extract programs. A set of STATA extract programs will have *.dct and *.do extensions.

Tagged Variable A variable on either the variable list or the working taglist that has been "tagged," i.e. either highlighted or clicked.

Taglist (1) A list of data file variables. Taglists can be one of three types: predefined taglists (saved lists of variables that came with the ECB), user defined taglists (saved lists of variables created by users), and working taglists (the current list of variables defined on the main ECB screen). (2) A menu listed on the menu bar that allows a user to manipulate the current working taglist.

Title Bar The horizontal bar located at the top of the main ECB screen. It gives the name of the application (ECB for Windows) and the catalog in use.

User Defined Taglist A saved list of catalog variables created by a user in an earlier session, or created and saved earlier in the current ECB session.

Variable List The window that forms the left half of the main ECB screen. It shows all of the variables listed within a catalog, as manipulated by the Narrow, Expand, and Reset command buttons.

Weighted Frequency The number of times the value(s) of a variable appears within a catalog, as defined by the weights assigned to the data file records.

Weighted Percentage A relative measure of how often the value(s) of a variable appears within a catalog as compared to all values of that variable, as defined by the weights assigned to the data file records.

Working Taglist The window that forms the right half of the main ECB screen. It shows the current catalog variables selected for creating either a codebook or an extract pr