README.PDF for the ELS:2002 CD-ROM and Electronic Codebook (ECB) (NCES 2004-404)

1. What files are on the ELS:2002 CD-ROM?
The ELS:2002 CD-ROM contains 3 folders and 2 files in the Root folder that are described below:

• The Root folder includes this document (README.PDF), HELP.PDF (see item #4 below), and folders *ECBW, Report, and Tag*
• Folder *ECBW* includes two data files (student level and school level), documentation for data files, and an installation program (setup.exe) for the ELS:2002 ECB.
• Folder *Report* includes an electronic copy of the ELS:2002 Base-Year Data File User's Manual.
• Folder *Tag* includes two "tag" files that provide the user with tags of critical variable (e.g., IDs, stratification variables, and design weights) which can be imported into an ECB session and used as the basis for producing SPSS or SAS card files.

2. Installation instructions for the ELS:2002 (E2P) electronic codebook (ECB)
•Insert the ELS:2002 CD into the CD-ROM drive.
•Click the Windows START menu button, and select RUN.
•Type: **D:\ECBW\Setup.exe** (If your CD-ROM drive is not D, enter the appropriate drive letter).
•Click on **OK** to run the setup program, and follow the directions on your screen.

3. What is an electronic codebook (ECB)?
The ECB system is primarily an electronic version of a fully documented survey codebook. For example, the ELS:2002 ECB allows the data user to browse through all instrument items and composite variables in the ELS:2002 data files, to search variable and value labels for key words related to particular research questions, to review the actual wording of these items along with notes and other pertinent information related to them, to examine the definitions and programs used to develop derived variables, and importantly, to create syntax files that import data from the selected variables. The ECB also provides an electronic display of the distribution of counts and percentages for each variable in the data set for the ELS:2002 Base Year interview in 2002.

Analysts can use the ECB to select or "tag" variables of interest, print hardcopy codebooks that display the description and distributions of the tagged variables, and generate SAS and SPSS program code (including variable and value labels). Note that the ECB does not analyze data, but only generates the program code to import the data into statistical analysis software packages (e.g. SAS and SPSS).

The public-use E2P ECB contains data for 15,362 sample members who participated in the base year student survey. The restricted-use E2R ECB contains data for 15,525 ELS:2002 sample members: data for the 15,362 sample members included in the public-use ECB; plus data for 163 sample members who were unable to take part in the survey due to disability or limited English proficiency, for whom contextual data is included.

NOTE: ECBs are defined by three characters, with the first character representing the study name (N = NELS:88; E = ELS:2002), the second indicating the year or wave of the study (2 = 2002 for ELS:2002), and the last character describing the type of information contained in the ECB (e.g., "P" for public-use or "R" for restricted-use data). The ELS:2002 public-use ECB is named E2P, where the "E" indicates the Education Longitudinal Study of 2002, the "2" represents the 2002 data collection, and the "P" signifies that the data is public and thus available for use without a license.

4. Getting started with the ELS:2002 ECB
Data users new to electronic codebooks are encouraged to read and understand the documentation for these products to ensure that these powerful analysis tools are effectively used. HELP.PDF, included in the root folder, includes useful guidance on the ECB, its system requirements, and operation. Appendix A of the base-year data file user's manual (Introduction to the Electronic Codebook) includes useful guidance about analyzing ELS:2002 information. This document is included in the Reports folder on this CD. NCES' ELS:2002 website (http://nces.ed.gov/surveys/els2002/) is also an excellent resource for new users. This site includes downloadable versions of many of the study's important technical and methodological reports.

5. Description of Data Files on ECB E2P
The ELS:2002 ECB (E2P) includes two megafiles: STUDENT and SCHOOL.

STUDENT – Student-level information, including data from the base-year student questionnaire, the student assessments in reading and mathematics, the teacher questionnaire, and the parent questionnaire are represented. Weights, participation flags and status indicators, and composite variables (also called constructed variables, derived variables, or created variables) are located at the beginning of the file, followed by the questionnaire variables. STU_ID is the unique identifier and there is one case per student respondent.

SCHOOL – School-level information, including data from the base year school administrator questionnaire, library media center questionnaire, and facilities checklist. SCH_ID is the unique identifier and there is one case per school. Weights, participation flags and status indicators, and composite variables are located at the beginning of the file, followed by the questionnaire/checklist variables.

Student and school files can be merged by SCH_ID. When merged, school level data is replicated for each student associated with that school.
Please Note: For one school (SCH_ID=3423), none of the sampled students were eligible. This school will have data for school level variables but no data for student level variables.

6. ELS:2002 Tag Files
The *Tag* folder in the CD-ROM includes two files. Tag files can be imported into the ECB, which will automatically select or "tag" items. The tag files are available to guide users in selecting items critical to the analyses of student-level or school-level data. The items automatically selected when importing the tag files are as follows:

Student tag file (d:\tag\student.tag)
- STU_ID – Student ID, the unique identifier for students on the Student data file
- SCH_ID – School ID, the unique identifier for schools on the School data file
- STRAT_ID – the analysis stratum to be used for computing Taylor Series variance estimates
- PSU – the analysis primary sampling unit to be used for computing Taylor Series variance estimates; the combination of STRAT_ID and PSU maps to a unique school
- BYSTUWT – Student final design weight for all base year responding students

School tag file (d:\tag\school.tag)
- SCH_ID – School ID, the unique identifier for schools on the School data file

- STRAT_ID – the analysis stratum to be used for computing Taylor Series variance estimates
- PSU – the analysis primary sampling unit to be used for computing Taylor Series variance estimates; the combination of STRAT_ID and PSU maps to a unique school
- BYSCHWT – School final design weight, generalizes to all regular high schools with 10th grades in the fifty states and the District of Columbia in the 2001-2002 school year

7. NCES contacts
The National Center for Education Statistics, U.S. Department of Education, strives to make all of its products available in a variety of formats and in language that is appropriate to a variety of audiences. If you have any comments, suggestions, or questions about this data product or any other NCES product, please direct your comments to:

> National Center for Education Statistics
> Institute of Education Sciences
> U.S. Department of Education
> 1990 K Street NW
> Washington, DC 20006

For additional information and questions about ELS:2002 and NCES' education longitudinal studies program, please link to the ELS:2002 web page:

http://nces.ed.gov/surveys/els2002/

Or contact:
> Jeffrey Owings
> (202) 502–7423
> Jeffrey.Owings@ed.gov