NATIONAL CENTER FOR EDUCATION STATISTICS

User's Manual September 1994

NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988

SECOND FOLLOW-UP: STUDENT COMPONENT DATA FILE USER'S MANUAL

U.S. Department of Education
Office of Educational Research and Improvement NCES 94-374

NATIONAL CENTER FOR EDUCATION STATISTICS

User's Manual **September 1994**

NATIONAL EDUCATION LONGITUDINAL STUDY OF 1988

SECOND FOLLOW-UP: STUDENT COMPONENT DATA FILE USER'S MANUAL

Steven J. Ingels
Kathryn L. Dowd
John D. Baldridge
James L. Stipe
Virginia H. Bartot
Martin R. Frankel

National Opinion Research Center (**NORC**)
at the University of Chicago

Peggy Quinn
Project Officer
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement **NCES 94-374**

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon Robinson
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner of Education Statistics

National Center for Education Statistics

"The purpose of the Center shall be to collect, and analyze, and disseminate statistics and other data related to education in the United Sates and in other nations."--Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e-1).

September 1994

Contact:
Peggy Quinn
(202) 219-1743

Foreword

This manual has been produced to familiarize data users with the procedures followed for data collection and processing of the second follow-up student component of the National Education Longitudinal Study of 1988 (NELS:88). A corollary objective is to provide the necessary documentation for use of the data file.

Use of the data set does not require the analyst to be a sophisticated statistician or computer programmer. Most social scientists and policy analysts should find the data set organized and equipped in a manner that facilitates straightforward production of statistical summaries and analyses. This manual provides extensive documentation of the content of the data file and how to use it. **Chapter VII and Appendix I, in particular, contain essential information that allows the user to immediately proceed with minimal startup cost. A careful reading of Chapter VII and Appendix I will help users to avoid common mistakes that result in costly computer job failures or incorrect results.**

The rest of the manual provides a wide range of information on the design and conduct of the National Education Longitudinal Study of 1988 (NELS:88). Chapter I begins with an overview and history of NCES's National Education Longitudinal Studies program and the various studies that it comprises. Chapter II contains a general description of the data collection instruments used in the NELS:88 second follow-up.

The sample design and weighting procedures used in the second follow-up study are documented in Chapter III, as well as standard errors and design effects, non-sampling measurement errors, and problematic variables.

Data collection procedures, schedules, and results are presented in Chapter IV. Chapter V describes data control and preparation activities such as monitoring receipt of questionnaires, editing, and data retrieval. Chapter VI describes data processing activities including machine editing and construction of the cleaned data tape. Finally, Chapter VII describes the organization and contents of the data file and provides important suggestions for using it.

The appendices contain a list of other NCES NELS:88 publications; guidelines for Statistical Analysis System (SAS) users; the second follow-up student questionnaire; the record layout for the student questionnaire; specifications for the composite variables; the content areas of the second follow-up components; a glossary of project terms; a discussion of conducting cross-cohort trend analyses of students; and a codebook for the student questionnaire data.

In addition to the study described in this manual, a number of supplemental NELS:88 components are also described in Appendix A.

Earlier NCES longitudinal studies that may be of interest to NELS:88 users are described in Appendix B including the following: the High School and Beyond (HS&B) base year files; merged HS&B first, second, third, and fourth follow-up files; related HS&B files; and assorted files related to the National Longitudinal Study of the High School Class of 1972 (NLS-72).

A Note on Data Use and Confidentiality

The NELS:88 second follow-up data files are released in accordance with the provisions of the General Education Provisions Act (GEPA) [20-USC 122e 1] and the Carl D. Perkins Vocational Education Act. The GEPA assures privacy by ensuring that respondents will never be individually identified.

The National Center for Education Statistics (NCES) is responsible under the Privacy Act and Public Law 100-297 for protecting the confidentiality of individually identifiable respondents, and is releasing this data set to be used for statistical purposes only. Record matching or deductive disclosure by any user is prohibited.

To ensure that the confidentiality provisions contained in PL 100-297 and the Privacy Act have been fully implemented, procedures commonly applied for disclosure avoidance in other Government-sponsored surveys were used in preparing the data file associated with this manual. These include suppressing, abridging, and recoding identifiable variables. Every effort has been made to provide the maximum research information that is consistent with reasonable confidentiality protection. Deleted, abridged, and/or recoded variables appear with an explanatory footnote in the codebook attached to each user's manual.

Acknowledgements

A study such as this is built first and foremost upon the students, dropouts, teachers, school administrators, and parents who have so generously provided its basic data. We are grateful for their cooperation. We also thank the considerable numbers of school personnel who have assisted in the implementation of NELS:88.

We wish to acknowledge the role of a number of other individuals in the realization of the aims of this study. Donald Rock and Judith Pollack of Educational Testing Service served as task leaders for cognitive test development. Miriam Clarke provided counsel on management issues in the main study. Leslie Scott contributed significantly to the conceptualization and development of file specifications and composite variables for the components of the study.

We are also grateful to the members of NCES staff in the Longitudinal and Household Studies Branch who worked closely with us on this project. Jeffrey Owings, chief of the Longitudinal and Household Studies Branch; Peggy Quinn, project officer for the second follow-up; as well as other branch staff--Ralph Lee, Shi-Chang Wu, and Jerry West--who contributed to various aspects of this study. Bob Burton of the Statistical Standards and Methodology Division supplied statistical advice and review.

Three individuals in other agencies have worked particularly hard and effectively to help realize and extend the potential of NELS:88. Larry Suter of the National Science Foundation, Dick Berry (formerly of the National Science Foundation), and Carmen Simich-Dudgeon (formerly of the Office of Bilingual Education and Minority Languages Affairs (OBEMLA) of the U.S. Department of Education). We are grateful for their efforts.

In addition, we would like to express our appreciation to the members of what began in the base year as our National Advisory Panel, and became in 1989 the NELS:88 Technical Review Panel. The panelists--Jerald G. Bachman, Gordon Ensign, Lyle V. Jones, Nancy Karweit, Richard J. Murnane, Patricia Shell, Marshall S. Smith, and John Stiglmeier--provided wise counsel on many difficult issues of design, instrumentation and implementation. As consultants to the second follow-up, Aaron Pallas, Joan Talbert, Leigh Burstein, Anthony Bryk, and Senta Raizen also contributed importantly to the design and ultimate success of the study.

Steven J. Ingels was overall NELS:88 second follow-up project director. Lisa Thalji was associate project director responsible for securing school cooperation and locating NELS:88 cohort members. Katy Dowd was associate project director responsible for the student component during data collection. Laura Reed and Virginia Bartot were the data processing managers, and Martin R. Frankel was the task leader for sampling and statistics.

The authors also wish to acknowledge those who contributed to the production of this manual. Kenneth A. Rasinski performed the confidentiality disclosure analysis for the NELS:88 Second Follow-Up. Additionally, Doug Barge, Michael Ma, Gloria Rauens, Supriti Sehra, Shiow-Ling Tsai-Ma, and Hsiuling Young provided a great deal of their time and expertise to produce the statistics reported throughout the manual. Our appreciation is also extended to Karen Sutherlin and Cynthia Mathews for their patience and thoroughness in the production of the manuscript. Finally, we would like to thank the National Opinion Research Center field and telephone center interviewers and supervisors who with such energy and determination collected the NELS:88 data.

Table of Contents

Foreword i

A Note on Data Use and Confidentiality iii

Acknowledgements iv

I. Introduction 1

1.1 Organization of the Data User's Manuals 1

1.2 Overview 2

1.2.1 NCES's National Education Longitudinal Studies Program 2

1.2.2 The National Longitudinal Study of the 1970s: NLS-72 2

1.2.3 High School and Beyond of the 1980s: HS&B 3

1.3 The National Education Longitudinal Study of 1988: Overview 6

1.3.1 NELS:88 Study Objectives 7

1.3.2 Base Year Study and Sample Design 9

1.3.3 First Follow-Up Core Study and Sample Design 12

1.3.4 Second Follow-Up Core Study and Sample Design 13

1.3.5 Second Follow-Up Design Enhancements 15

1.4 NELS:88 Sponsors 16

1.4.1 Sample Supplements and Augmentations 16

1.4.2 Instrument Supplements 16

1.5 NELS:88 Data and Documentation 17

1.5.1 Base Year Data Tapes and Documentation 18

1.5.2 First Follow-Up Data Files and Documentation 18

1.5.3 Second Follow-Up Electronic Codebook on CD-ROM and Documentation 20

II. Data Collection Instruments 21

2.1 Instrument Development 21

2.2 Survey Instruments and Content Coverage 22

2.2.1 Student Questionnaire and Cognitive Tests . . . 22

2.2.2 Dropout Questionnaire 24

2.2.3 New Student Supplement 26

2.2.4 Early Graduate Supplement 26

2.2.5 Adapting Questionnaires for Telephone Administration 26

III. Sample Design and Implementation; Survey Error Assessment 28

3.1 NELS:88 Sample Design 28

3.1.1 Base Year Sample Design 28

3.1.2 First Follow-Up Sample Design 29

3.1.3 Second Follow-Up Sample Design 34

3.2 Calculation of Weights 42

3.2.1 Calculation of Base Year Sample Weights 42

3.2.2 Calculation of First Follow-Up Sample Weights . 44

3.2.3 Calculation of Second Follow-Up Weights 48

3.3 Standard Errors and Design Effects 55

3.3.1 Base Year Standard Errors and Design Effects . . 56

3.3.2 First Follow-Up Standard Errors and Design Effects 85

3.3.3 Second Follow-Up Standard Errors and Design Effects 88

3.3.4 Design Effects and Approximate Standard Errors . 90

3.4 Additional Sources of Nonobservational Error . . 92

3.4.1 Biases Caused by Undercoverage of Special Populations 92

3.4.2 Unit and Item Nonresponse 97

3.4.3 Observational Error: The Quality of Responses . 109

IV. Data Collection 111

4.1 Base Year Data Collection 111

4.1.1 Base Year Pre-Data Collection Activities 111

4.1.2 Base Year Cohort Data Collection Activities . . 112

4.1.3 Base Year Data Collection Results 113

4.2 First Follow-Up Data Collection 113

4.2.1 First Follow-Up Pre-Data Collection Activities . 114

4.2.2 First Follow-Up Cohort Data Collection Activities 116

4.2.3 First Follow-Up Dropout Survey 117

4.2.4 First Follow-Up Survey of Base Year Ineligible Students 119

4.2.5 First Follow-Up 1990 and 1988-90 Panel Data Collection Results 121

4.3 Second Follow-Up Data Collection 121

4.3.1 Second Follow-Up Pre-Data Collection Activities 123

4.3.2 Second Follow-Up Cohort Data Collection Activities 126

4.3.3 Second Follow-Up Student Survey and Cognitive Tests 126

4.3.4 Second Follow-Up Dropout Survey 130

4.3.5 School Effectiveness Study 131

4.3.6 Followback Study of Excluded Students (FSES) . . 133

4.3.7 Second Follow-Up Data Collection Results 134

V. Data Control and Preparation 143

5.1 On-Site Editing and Retrieval 143

5.2 Monitoring and Receipt Control 143

5.3 In-House Editing and Coding 143

5.4 Data Capture and Archival Storage 144

VI. Data Processing of the Student Questionnaires . . . 145

6.1 Machine Editing 145

6.2 Data File Preparation 147

6.3 CD-ROM Electronic Codebook 148

VII. Guide to the Data Files, Documentation and CD-ROM Electronic Codebooks 149

7.1 Basics for Analyses: Second Follow-Up Questionnaire and Sample Indicators, Statistical Weights and Use of Statistical Packages 153

7.1.1 Questionnaire/Sample Flags Included on Magnetic Tape and ECB Releases 153

7.1.2 Packaged Statistical Programs 165

7.2 Content and Organization of the Data Files . . . 165

7.2.1 Identification Codes 167

7.2.2 The Student Survey Instruments 167

7.2.3 Composite Variables 170

7.3 Guide to the NELS:88 Codebooks 173

7.3.1 Hardcopy Codebooks in NELS:88 Data User's Manuals 174

7.3.2 The NELS:88 Electronic Codebook System (ECB) . . 178

Appendices

Appendix A: NELS:88 Sources of Contextual Data: Parent, Teacher, School Administrator, Transcript, and Course Offerings Components

Appendix B: NELS:88-Related Data Files Available from the National Center for Education Statistics

Appendix C: National Center for Education Statistics, Longitudinal and Household Studies Branch, NELS:88 Publications

Appendix D: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors: Analytical Implications of Design Differences Between the Studies

Appendix E: NELS:88 Second Follow-Up Item Overlap with NELS:88 Base Year and First Follow-Up; and with HS&B and NLS-72

Appendix F: NELS:88 Second Follow-Up Followback Study of Excluded Students: Screeners

Appendix G: Public Use Record Layout for NELS:88 Second Follow-Up Data Tape

Appendix H: NELS:88 Student Data Weights, Flags, and Composite Variables (BY, F1, and F2)

Appendix I: Guidelines for Using SAS with NELS:88 Second Follow-Up Student Data

Appendix J: NELS:88 Second Follow-Up Student Questionnaire Codebook

Appendix K: NELS:88 Second Follow-Up Student Questionnaire and New Student Supplement

Appendix L: Critical Items, Abbreviated Questionnaire Items, and Refusal Conversion Items from the Second Follow-Up Student and Dropout Questionnaires and New Student Supplement

Appendix M: NELS:88 Second Follow-Up Content Areas: Student, Dropout and School Components

Appendix N: NELS:88 Second Follow-Up Content Areas: Student, Teacher and Parent Components

Appendix O: Glossary of NELS:88 Terms

Appendix P: Selected Measures of NELS:88 Base Year Data Quality

Appendix Q: NELS:88 Third Follow-Up (1994) Questionnaire

Appendix I

Guidelines for Using SAS with

NELS:88 Second Follow-Up Student Data

Guidelines for using SAS with NELS:88 Second Follow-Up Student Data

The files provided for public release tapes include SAS cards and SAS system files for the NELS:88 second follow-up. The SAS system file for each survey wave includes:

1) Questionnaire data
2) Flags, Weights and Composites

Users who plan to analyze NELS:88 data on personal computers can seek counsel in the Guide to the NELS:88 ECB/CD-ROM. The sections that follow pertain primarily to mainframe applications.

In the points below, methods to contain potential difficulties that may be encountered when using large data files with SAS are discussed.

1. Use the '(KEEP=...)' and '(DROP =...)' options in the 'SET' statement and/or in the 'DATA' statement when creating working data files so that unwanted variables are not included in the files. The '(KEEP =...)' option does not reorder the variables in the new data set.

   The files are large and the SAS cards associated with all of the variables within a file require a great deal of memory. Eliminating unwanted variables and the cards associated with them will reduce the amount of memory necessary to run jobs.

2. Some of the label statements given in the student and dropout SAS card files may need to be eliminated because of SAS system limitations present at many computer installations.

3. ***The large*** number of VALUE statements in the PROC FORMAT section of the student and dropout SAS cards require that a special DD statement be placed just after the // EXEC SAS statement to increase the capacity of the format library during a SAS run:

   //LIBRARY DD SPACE =(TRK,(25,25,60))

   Since this may not be possible at some computer installations, it may be necessary to delete some VALUE statements.

4. When working with large files, it may be necessary to override the default work space with the following DD statement:

   //WORK DD UNIT=SYSCR,SPACE=(CYL, (40,40))

   Place the //WORK DD statement just after the // EXEC SAS statement (or after the //LIBRARY DD statement, if that is included as well).

5. ***The*** formats given in the PROC FORMAT step here are not permanently associated with each variable. Whenever they are needed for a procedure, it is necessary to include them in this PROC FORMAT step before the procedure(s) that will use them. The following example will help to illustrate this point.

Suppose you were interested in assessing the association between fathers' educational aspirations and a son's versus a daughter's educational expectations. That is, overall do students' expectations reflect their father's aspirations, and might such an association vary by sex? To do this you might construct a three-way crosstab.

In the following example PROC FORMAT is used first to make a temporary library of formats (sets of value labels). Then PROC FREQ is used to access the second follow-up student SAS system file and to create a three-way crosstab. The FORMAT statement in PROC FREQ links each variable in the crosstab to the appropriate set of value labels stored in the temporary format library.

```
// EXEC SAS
//LIBRARY DD SPACE =(TRK,(25,25,60))
//WORK DD UNIT=SYSCR,SPACE=(TRK,(1000,1000))
//IN1 DD DSN=ACT.PUBL.F2ST. SASLIB,DISP=SHR
//SYSIN DD *

OPTIONS DQUOTE;

PROC FORMAT;
VALUE SC42V

        00 = "DOES NOT APPLY"
        01 = "LESS THAN HS"
        02 = "HS ONLY"
        03 = "LESS 2YRS/SCHL"
        04 = "2YRS MORE/SCHL"
        05 = "TRADE SCHL DGREE"
        06 = "LESS 2YRS CLLEGE"
        07 = "MORE 2YRS CLLEGE"
        08 = "FINISH COLLEGE"
        09 = "MASTER'S DEGREE"
        10 = "PH. D., M. D., OTHER"
        11 = "DON'T KNOW"
        96 = "MULT RESPONSE"
        97 = "REFUSED"
        98 = "MISSING"
        99 = "LEGITIMATE SKIP"

VALUE SC44V
        01="LESS THAN HS"
        02 = "HS ONLY"
        03 = "LESS 2YRS/SCHL"
        04 = "MORE 2YRS/SCHL"
        05 = "TRADE SCHL DGREE"
        06 = "LESS 2YRS CLLEGE"
        07 = "MORE 2YRS CLLEGE"
        08 = "FINISH COLLEGE"
        09="MASTER'S OR EQU"
```

```
        10 = "PH. D., M. D., OTHER"
        11 = "DON'T KNOW"
        96 = "MULT RESPONSE"
        97 = "REFUSED"
        98 = "MISSING"
        99 = "LEGITIMATE SKIP"
 ;

VALUE MC1V
        1 = "MALE"
        2 = "FEMALE"
;

PROC FREQ DATA=IN1.F2STUDNT;
FORMAT
F2S42A   SC42V.
F2S43    SC44V.
F2SEX    MC1V.
;

TABLES  F2SEX*F2S42A*F2S43;
TITLE "DAD'S AND R'S EDUCATIONAL EXPECTATIONS BY R GENDER";
```

At the end of each SAS card file, there is a frequency procedure which contains FORMAT statements for every variable for which there is a format. These FORMAT statements can be used in any SAS procedure. However, if there are a large number of format links, they must be divided into several format statements to work. (Using about 90 format links in the format statement proved successful on the University of Chicago mainframe.)

6. Whenever variables are needed from several student level files (i.e., second follow-up student and base year student), the files may be merged by STU_ID using SAS MERGE statements. A simple one line MERGE statement will put variables from separate files together in a single record for analysis.

The following second follow-up and base year example may help to illustrate the merge statement. Suppose you wanted to examine how the educational expectations of respondents who are still enrolled (in twelfth grade in 1992) and who completed a student questionnaires have changed since the eighth grade. To do this you might construct a three-way crosstab.

In the following example PROC FORMAT is used to make a temporary library of formats. Next the second follow-up student system file, and the base year system file are merged. Then, PROC FREQ is used to create a three-way crosstab. When merging two large files, it is helpful to use separate "KEEP =" statements for each file being combined.

```
// EXEC SAS
//LIBRARY DD SPACE=(TRK,(25,25,60))
//WORK DD UNIT=SYSCR,SPACE=(TRK,(1000,1000))
//IN1 DD DSN=ACT.PUBL.F2ST.SASLIB,DISP=SHR
//IN2 DD DSN=ACT.PUBL.BYST.SASLIB,DISP=SHR
```

```
//SYSIN DD *

OPTIONS DQUOTE;

PROC FORMAT;
VALUE SC44V
        01 = "LESS THAN HS"
        02 = "HS ONLY"
        03 = "LESS 2YRS/SCHL"
        04 = "MORE 2YRS/SCHL"
        05 = "TRADE SCHL DGREE"
        06 = "LESS 2YRS CLLEGE"
        07 = "MORE 2YRS CLLEGE"
        08 = "FINISH COLLEGE"
        09 = "MASTER'S OR EQU"
        10 = "PH. D., M. D., OTHER"
        11 = "DON'T KNOW"
        96 = "MULT RESPONSE"
        97 = "REFUSED"
        98 = "MISSING"
        99 = "LEGITIMATE SKIP"
,
VALUE FC3V
        0 = "DID NOT COMPLETE"
        1 = "STDNT QUEX CMPLT"
        2 = "DRP QUEX CMPLT"
,
VALUE FBYS45V
        01 = "WON'T FINISH H.S"
        02 = "WILL FINISH H.S"
        03 = "VOC,TRD,BUS AFTR H.S"
        04 = "WILL ATTEND COLLEGE"
        05 = "WILL FINISH COLLEGE"
        06 = "HIGHER SCH AFTR COLL"
        96 = "MULTIPLE RESPONSE"
        97 = "REFUSAL"
        98 = "MISSING"
        99 = "LEGITIMATE SKIP"
;
```

```
DATA COMBINE;
MERGE IN1.F2STUDNT (KEEP =STU_ID F2S43 F2QFLG) IN2.BYSTUDNT
(KEEP =STU_ID BYS45); BY STU_ID;

PROC FREQ;
FORMAT
F2S43   SC44V.
F2QFLG  FC3V.
BYS45   FBYS45V.
;
TABLES F2QFLG * F2S43 * BYS45;
TITLE "EDUCATIONAL EXPECTATIONS SINCE GRADE 8";
```

7. For very large files, the user may encounter problems when sorting. Various options may be added to the //EXEC SAS card to circumvent these problems. A suggested option is given below (consult the SAS manual for descriptions of these options):

```
// EXEC SAS,OPTIONS='NODYNALLOC',REGION=1280K,SORT=30
```

8. It is suggested that the user include the LENGTH statement when creating new variables, in order to save space and computer memory.

9. For many tabulations, PROC TABULATE produces the most readable output. The SAS user may use the format statements (provided) for classification variables to produce the row values of tabulated tables.

10. Output from SAS can be downloaded to personal computers for production of final reports. NCES has available a program for taking into account the sample design when computing standard errors. The program, known as CTAB, is a Taylor series-based routine that uses an ASCII file to compute standard errors for cross-classifications. The program also produces labeled tabular output suitable for use in publications. CTAB is available for use on microcomputers, and can be obtained through NCES.

11. Use the NCES- and NORC-defined composite and classification variables whenever possible to simplify programming. These classification variables were carefully constructed and, for many of them, sources of data from outside the student questionnaire were merged into the student data to construct the composites.

12. SAS and SPSS-X system files can now be converted at many computer installations. Contact your own facility to obtain the information necessary to create an SPSS-X file from SAS and vice versa.

13. There is a peculiarity with version 6.06 of SAS. The symbol "%" will not be printed in a variable label if the label is the first character to be printed on the page.