

U.S. Department of Labor
U.S. Bureau of Labor Statistics

Now Available: Two Waves of School Transcript Data

To complement data on respondents' educational experiences collected during yearly interviews, staff administering the National Longitudinal Survey of Youth 1997 (NLSY97) collected transcripts on two occasions directly from respondents' high schools. This data gathering, known as the high school transcript survey, was done in 2000 (wave 1) and in 2004 (wave 2).

Transcripts were requested for all NLSY97 youths who had provided a signed authorization for transcript release and whose interview records included contact information for at least one school since grade 9. Transcript data were received for 1,417 respondents in wave 1 and 4,815 respondents in wave 2, totaling 6,232 respondents. From these transcripts, survey staff abstracted a series of variables, including dates of each of the respondent's terms in school, courses taken, grades, and other academic indicators. These variables can help researchers understand schooling experiences and examine the school-to-work transition in greater detail.

Data from both waves of the transcript survey now are available in the public use data files under the "Transcript Survey" area of interest and with variable names beginning with "TRANS." To facilitate analysis of the transcript data, wave 2 information has been appended to the wave 1 variables. Thus, all transcript survey variables are associated with the year 1999 (where the wave 1 data appear), including a handful of variables added only for wave 2 respondents.

Data collection

In wave 1, survey staff collected hard-copy transcripts for NLSY97 respondents who

had provided signed authorization for transcript release and who were no longer enrolled in high school in spring 2000. Nonenrollment occurred when a youth either graduated from high school or dropped out of school and was at least 18 years old. In wave 2 (2004), staff collected transcript data for all remaining eligible NLSY97 youths, including those who were fielded in wave 1 but for whom no transcript had been received.

A transcript request packet was mailed to each school from which an NLSY97 youth received his or her high school diploma, or to the last school the youth reported attending in the youth interview. Survey staff abstracted a standard set of variables from the transcripts into a computer-assisted data entry (CADE) system. These data then were used to code courses into the Revised Secondary School Taxonomy (SST-R) framework and to create a standardized variable for the grade received. When available, hard-copy and online course catalogs were useful in clarifying particular types of coursework reported at a given school, as were phone calls placed by the coding staff to high school personnel.

Types of variables available

The transcript data file contains several types of variables. Variables about the schools include the variable TRANS_SCH_CAT, which indicates whether a course catalog was received from the school to aid in coding. The highest number of schools reported for any respondent is 12, so this variable is repeated 12 times. This course catalog variable also functions as the identification number of the school; the school ID number does not, however, link to any variables in the main data file.

The transcript file also includes variables that provide information about students. These are variables that are not associated with a specific term or course—for example, achievement test scores as they appear on the transcript (ACT, PSAT, SAT, SAT II, AP), information on absences and tardies, the student's school completion status, and dates of enrollment. Variables also indicate whether the student participated in programs such as gifted, bilingual, or special education.

Several variables in the transcript file provide information about the terms at the school where the respondent was enrolled. For up to 28 terms, these items report the beginning and ending dates of the term, how the school year was divided (season, entire year, or another term-based system), the academic year of the term, the respondent's grade level that term, and the number of credits earned. A variable listing the school the respondent attended during that term can be linked to the course catalog variable as previously described.

Finally, the transcript file provides details about each course appearing on a student's high school transcript. Course-specific variables include the course code from the Revised Secondary School Taxonomy (SST-R), the grade earned in the course, and the credit value of the course. Because schools use many different grading systems, the course grades were converted to a standard scale that can be compared across respondents. A variable called "Recoding Status of Grade" indicates how the "grade earned" variable for each course was created.

An important improvement in the data is the addition of a standardized credit variable, known as a Carnegie credit, for each course in the transcript record. These Carnegie credit variables, beginning with

TRANS_CRS_CARNEGIE_CREDIT, have been added for all wave 1 and wave 2 data and permit credit-based computations across schools and students. Translation of school-defined credits to Carnegie units was informed by school documentation, consultation with school staff, and application of generic rules where necessary. Variables that begin with TRANS_CRS_CARNEGIE_RECODO_STATUS indicate how the Carnegie credit was computed for each course.

More details on the variables and how they were coded can be found in appendix 11 of the *NLSY97 Codebook Supplement*, available through the Documentation link on Web Investigator: www.nlsinfo.org/web-investigator/docs.php.

Obtaining school transcript data

Researchers can find NLSY97 transcript data on the most recently released NLS data set, which can be accessed through the Web site at www.bls.gov/nls. Choose the “NLS public-use files” link under “Obtaining NLS Data.” Then click on the link to the “NLS Product Availability Center,” which leads to Web Investigator. Extraction of variables and generation of tables require a free user account. □

Corrections Made to Schooling Event History

Researchers who used NLSY97 schooling event history data prior to the round 7 event history release (October 2005) should take note. A data quality review after round 6 found several grade progression inconsistencies (spurious repeated and skipped grades) that needed to be corrected. These corrections have been made and are reflected in the round 7 and round 8 main releases.

Here are the main problems and the steps taken to correct them:

- Using the grade reported in the initial 1997 survey and the date of high school graduation, survey staff identified all respondents who moved from the grade reported in 1997 to high school graduation in the expected amount of time. If a respondent’s graduation date suggests that the respondent followed a normal school progression—that is, completing

one grade per school year—the event history program flagged the respondent and imposed a normal progression on the event history variables.

- Many respondents were enrolled in college courses while still in high school. Event history arrays contain only a single grade attended for a given time period, and the original event history program was written to give college courses precedence over high school. For example, if a high school junior also took a freshman-level college class during his or her first semester, the program assigned a grade of “13” (first year in college) for that semester. If the student then finished 11th grade but did not take any college classes during the second semester, the data incorrectly showed that the student had jumped backward from year 13 to year 11. This anomaly resulted in extra promotions and regressions. The event history program has been rewritten to prioritize high school over college, removing these spurious grade changes.
- Some respondents provided a high school graduation date but then reported additional secondary school enrollment after that date. Survey staff decided to exclude post-graduation secondary school enrollment from the event histories, although this information is preserved in the raw data for researchers who might be interested in additional training received after graduation.
- Some respondents reported initial enrollment at a school but apparently did not understand that they should report each grade attended at that school in a separate loop within the “Schooling” section. As a result, some respondents appeared to remain in one grade for a long time (particularly if they had missed one or more interviews), then seemed to jump ahead several grades. If, for example, a respondent appeared to be in 9th grade for 3 years and then jumped ahead to 12th grade, the most likely reason was that he or she did not understand the schooling questions and actually did progress normally through 10th and 11th grade. The event history program flagged these respondents and adjusted their schooling history to follow a normal grade progression.
- In some cases, respondents appeared to

jump backward and then forward across multiple grades. For example, some respondents were listed as attending 9th grade, then 1st grade, then 11th grade. The most likely explanation for this pattern is a data entry error in which a zero was dropped from “10th.” Jumps in a normal school progression that appeared to be caused by a missing digit in a two-digit grade were corrected.

- When asked what grade they had first attended at a given school, some respondents instead reported the first grade *offered* at that school, causing a jump backward across grades and forward again the next year. Hand edits were made to adjust the event histories for these respondents to a normal grade progression.

The six changes described above significantly reduced the number of abnormal grade progressions found in the event history (SCH_GRADE_PROGRESS) variables. Approximately three-quarters of the problem cases were fixed. Although it is possible that errors remain, the survey staff feels that most of the remaining atypical grade changes reflect actual atypical progressions or demotions. Researchers with questions about schooling event history data may contact NLS User Services. □

Spotlight: Measures of Mental Health in the NLS

Researchers looking for psychological variables will be happy to learn that, for selected years, some of the NLS cohorts offer measures of mental health based on standardized psychological scales. Mental health measures can be used to help researchers understand why some individuals have difficulty holding down jobs or completing schooling and training. Here are three scales—the Center for Epidemiological Studies–Depression (CES-D) Scale, Bradburn’s Affect Balance Scale, and Rotter’s Internal/External Control Scale—and the cohorts that contain these scales. In some years, abbreviated scales were included. Partial scales yield similar results to full scales, although the number ranges are different. These variables are found in the main data files, accessible through Web Investigator at www.bls.gov/nls.

CES-D Scale

The CES-D Scale measures symptoms of depression. Developed in the 1970s, the full 20-item CES-D Scale covers affective, psychological, and somatic symptoms.

Questions. Individuals are asked how often certain statements such as “I was bothered by things that didn’t usually bother me” or “My sleep was restless” have been true during the past week. Response categories range from “rarely or none of the time” to “all of the time.”

Cohorts with the scale. The full CES-D Scale was administered to NLSY79 respondents in 1992. A seven-item subset was given to the NLSY79 cohort in 1994. In 2002, as part of an extended health module, NLSY79 respondents age 40 and older were given a reduced set of nine CES-D Scale items.

The seven-item CES-D subset used in the NLSY79 cohort in 1994 has been administered in all young adult surveys since 1994. Because the administration pattern has varied from round to round, based on the date of the last interview, users are encouraged to consult the *NLSY79 Child and Young Adult User’s Guide* for specifics.

The full CES-D Scale was administered to the mature women cohort in 1989 and again in 2003. The 1995-2001 mature women surveys contain a reduced set of seven CES-D items; these variables can be found in the “Health” and “Retirement and Pension” sections.

In 1993 and in 2003, the full CES-D Scale was administered to the young women respondents. Like the surveys administered to the mature women cohort, the 1995-2001 young women surveys contained only seven items from the original scale.

For the older men, the full CES-D Scale was asked in 1990.

Affect Balance Scale

The 10-item Affect Balance Scale, a measure of positive and negative feelings, was developed by Norman Bradburn. This measure of positive and negative feelings in the weeks before the interview provides an indication of the respondent’s general psychological well-being.

Questions. The items probe, for example,

whether the respondent felt particularly excited or interested in something in the past few weeks, whether he or she was so restless he or she could not sit long in a chair, and whether he or she felt that things were going his or her way.

Cohorts with the scale. The complete 10-item Affect Balance Scale was administered to the mature women respondents in 1981 (see variables R05226-R05235). The same scale also was given to the older men cohort in 1981, 1983, and 1990. (As an example, see the 1981 variables R04865-R04874.)

Rotter’s Internal-External Locus of Control Scale

The Locus of Control Scale, a 23-item questionnaire developed by J.B. Rotter, measures generalized expectancies for internal versus external control of reinforcement. Internal control refers to the perception of events as being under personal control; external control involves events being perceived as unrelated to one’s own behavior.

Questions. Respondents are presented a series of paired statements such as “What happens to me is my own doing” and “Sometimes I feel that I don’t have enough control over the direction my life is taking” and asked to select the statement closer to their opinion.

Cohorts with the scale. The NLSY79 cohort answered questions on a four-item abbreviated version of Rotter’s scale in 1979. In 1992, NLSY79 respondents also were asked items from the Pearlin Mastery Scale, which is a similar measure of internal versus external control. The Pearlin scale also has been administered to the young adult cohort in all young adult survey years.

The mature women cohort answered questions on an 11-item abbreviated version of Rotter’s scale in 1969, 1972, and 1977; the 2001 survey included a modified 4-item scale that allowed respondents to qualify how much closer (slightly closer or much closer) a chosen statement fit them.

At six survey points, the young women gave responses to a measure of internal-external control. In 1970, 1973, and 1978, an 11-item abbreviated version of Rotter’s

scale was used. In 1983 and 1988, this scale was further reduced to only four items. The 2001 survey included a modified four-item scale.

Abbreviated versions of the Rotter Locus of Control Scale were asked of the older men cohort in 1969, 1971, 1976, and 1981. Abbreviated versions were administered to the young men cohort in 1968, 1971, and 1976.

Other mental health measures

Additional questions measuring psychological health can be found throughout the NLS cohort sets. (For example, the 2000 and 2002 NLSY97 data include the five-item Mental Health Index III developed by the Rand Medical Outcomes Study.) See the “Attitudes” or “Health” sections of the cohorts’ user’s guides for specifics. User’s guides can be accessed online at www.bls.gov/nls through the Documentation link or through Web Investigator. □

Frequently Asked Questions

The NLS staff encourages researchers to contact NLS User Services with questions and problems encountered while accessing and using NLS data and/or documentation. Every effort is made to answer these inquiries. Some recent questions and answers that may be of general interest to NLS users are listed below.

Q1. Are the young men and young women original cohorts asked any “job risk” questions similar to those asked in the NLSY79? I’m referring to the questions in which the respondents are asked if they would accept a job for which there was a 50-50 chance of cutting income by a certain percent or increasing income by a certain percent.

A1. Similar variables (though not worded exactly the same) exist in both the young women and young men cohorts for several of the survey years. Search on the word “hypothetical” in Web Investigator, then go to those questionnaires and get the actual wording of the questions. In addition, the young women and young men cohorts were asked a somewhat related set of questions about jobs, which can be found by using “commitment” as a search word.

Q2. What is item latency? I'm wondering what this term means in the context of the ASVAB test given to many of the NLSY97 respondents.

A2. Item latency variables, identified by titles beginning with "IL," essentially measure how long an item or question appeared on the respondent's computer screen. Latencies primarily are used to detect aberrant responses and can help identify individuals who did not take the tests or questionnaire seriously. They are one of the pieces of information that can be used to "clean" the data for constructing norms. The two exceptions are for the Numerical Operations and Coding Speed subtests, in which latencies are used to compute scores. Both of those tests measure speed and accuracy of responding when performing routine tasks.

Q3. Is there a household structure variable or set of variables in the NLSY79 where I could find without too much difficulty the number of people living in a household at any one time, and the relationships of those people to the respondent?

A3. Created variables for family size can be found in the NLSY79 "Key Variables" area of interest. These variables give the number of persons in the household related to the respondent by blood or marriage, including step/adopted children. However, to see the relationship to the respondent for all persons living in the household, one must cycle through relationship variables for all members of the household using the variables in the "Household Record" area of interest.

Q4. What is the total number of respondents in the 2004 young adult survey? I see different numbers in the "Common Key Variables" and the "Interviewer Remarks" sections.

A4. The 2004 young adult survey has 5,024 completed interviews. As long as the questionnaire is completed through the "Employment" section, it is considered a completed interview. There were eight cases in which the interview was stopped after the "Employment" section and before the "Interviewer Remarks" section. Because these cases did not complete the "Interviewer Remarks" section, they were assigned -7 for these variables (negative values differentiate invalid from valid responses).

To find out how many young adult respondents were interviewed in any given year, users can use the young adult sampling weight. In general, a case weight greater than zero indicates a case is completed for that survey round. A zero weight implies noncompletion. □

Completed NLS Research

The following is a listing of recent research based on data from the NLS cohorts that has not appeared in its current form in a previous issue of the *NLS News*. See the *NLS Annotated Bibliography* at www.nlsbibliography.org for a comprehensive listing of NLS-related research.

Aughinbaugh, Alison Aileen; Pierret, Charles R.; and Rothstein, Donna S. "The Impact of Family Structure Transitions on Youth Achievement: Evidence from the Children of the NLSY79." *Demography* 42, 3 (August 2005): 447-68. [Children of the NLSY79, NLSY79]

Caputo, Richard K. "Religiousness and Adolescent Behaviors: A Comparison of Boys and Girls." *Journal of Religion and Spirituality in Social Work* 24, 3 (2005): 39-67. [NLSY97]

Carneiro, Pedro M.; Heckman, James J.; and Masterov, Dimitriy V. "Labor Market Discrimination and Racial Differences in Premarket Factors." *Journal of Law and Economics* 48 (April 2005). [Children of the NLSY79, NLSY79]

Classen, Timothy and Hokayem, Charles. "Childhood Influences of Youth Obesity." *Economics and Human Biology* 3, 2 (July 2005): 165-87. [Children of the NLSY79, NLSY79]

Deary, Ian J.; Der, Geoff; and Shenkin, Susan D. "Does Mother's IQ Explain the Association Between Birth Weight and Cognitive Ability in Childhood?" *Intelligence* 33, 5 (September 2005): 445-54. [Children of the NLSY79, NLSY79]

Domina, Thurston. "Leveling the Home Advantage: Assessing the Effectiveness of Parental Involvement in Elementary School." *Sociology of Education* 78, 3 (July

2005): 233-49. [Children of the NLSY79]

Finger, Reginald; Thelen, Tonya; Vessey, John T.; Mohn, Joanna K.; and Mann, Joshua R. "Association of Virginity at Age 18 with Educational, Economic, Social, and Health Outcomes in Middle Adulthood." *Adolescent and Family Health* 3, 4 (2005): 164-70. [NLSY79]

Gius, Mark. "An Estimate of the Effects of Age, Taxes, and Other Socioeconomic Variables on the Alcoholic Beverage Demand of Young Adults." *Social Science Journal* 42, 1 (January 2005): 13-24. [NLSY79]

Green, Carole and Ferber, Marianne A. "Do Detailed Work Histories Help to Explain Gender and Race/Ethnic Wage Differentials?" *Review of Social Economy* 63, 1 (March 2005): 55-85. [NLSY79]

Hotz, V. Joseph; McElroy, Susan Williams; and Sanders, Seth G. "Teenage Childbearing and Its Life Cycle Consequences." *Journal of Human Resources* 40, 3 (Summer 2005): 683-715. [NLSY79]

Howell-Moroney, Michael Edward. "The Geography of Opportunity and Unemployment: An Integrated Model of Residential Segregation and Spatial Mismatch." *Journal of Urban Affairs* 27, 4 (October 2005): 353-377. [NLSY79]

Huebner, Beth Marie. "The Effect of Incarceration on Marriage and Work Over the Life Course." *Justice Quarterly* 22, 3 (September 2005): 281-303. [NLSY79]

Huang, Chien-Chung. "Pregnancy Intention from Men's Perspectives: Does Child Support Enforcement Matter?" *Perspectives on Sexual and Reproductive Health* 37, 3 (September 2005): 119-24. [NLSY79]

Kanazawa, Satoshi. "The Myth of Racial Discrimination in Pay in the United States." *Managerial and Decision Economics* 26, 5 (July-August 2005): 285-94. [NLSY79]

Laws, Jerry. "Working Too Early?" *Occupational Health and Safety* 74, 7 (July 2005): 4. [NLSY79]

MacMillan, Ross and Copher, Ronda. "Families in the Life Course: Interdepen-

gency of Roles, Role Configurations, and Pathways.” *Journal of Marriage and Family* 67, 4 (November 2005): 858-79. [NLSY79]

Mazumder, Bhashkar. “Fortunate Sons: Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data.” *Review of Economics and Statistics* 87, 2 (May 2005): 235-55. [NLSY79]

Persico, Nicola; Postlewaite, Andrew; and Silverman, Daniel Susman. “The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height.” *Journal of Political Economy* 112, 5 (October 2004): 1019-53. [NLSY79]

Reid, Lori L. and Padavic, Irene. “Employment Exits and the Race Gap in Young

Women’s Employment.” *Social Science Quarterly* 86, 5 (December 2005): 1242-60. [NLSY79]

Schmitz, Mark F. “Effects of Childhood Foster Care and Adoption on Adulthood Childbearing.” *Children and Youth Services Review* 27, 1 (January 2005): 85-98. [NLSY79]

Tomaskovic-Devey, Donald; Thomas, Melvin; and Johnson, Kecia Renee. “Race and the Accumulation of Human Capital across the Career: A Theoretical Model and Fixed-Effects Application.” *American Journal of Sociology* 111, 1 (July 2005): 58-89. [NLSY79]

Turner, Michael G.; Piquero, Alex R.; and Pratt, Travis C. “School Context as a Source

of Self-Control.” *Journal of Criminal Justice* 33, 4 (July-August 2005): 327-39. [NLSY79]

Weden, Margaret M. and Zabin, Laurie S. “Gender and Ethnic Differences in the Co-occurrence of Adolescent Risk Behaviors.” *Ethnicity and Health* 10, 3 (August 2005): 213-25. [NLSY97]

Williams, Donald R. “Youth Self Employment: Its Nature and Consequences.” *Small Business Economics* 23, 4 (November 2004): 323-36. [NLSY79] □

Are You Working With NLS Data?

If you are, we are interested in your work!

- Have you received funding to sponsor a project using NLS data?
- Are you working on a paper that uses NLS data?
- Have you published a recent paper using NLS data?

If you have received funding on a project, are working on a paper, or published a recent paper that uses NLS data, please contact: NLS User Services, Center for Human Resource Research, The Ohio State University, 921 Chatham Lane, Suite 100, Columbus, OH 43221; (614) 442-7366; e-mail: usersvc@postoffice.chrr.ohio-state.edu. Or use our online submission form—just go to www.nlsbibliography.org and click on “Submit Citation.” □

FIRST CLASS MAIL
Postage and Fees Paid
U.S. Department of Labor
Permit No. G-738

U.S. Bureau of Labor Statistics
Postal Square Building, Rm. 2850
2 Massachusetts Ave., NE.
Washington, DC 20212-0001
Official Business
Penalty for Private Use, \$300
Address Service Requested

NLS News



NLS Contact Information

NLS News is published quarterly by the U.S. Bureau of Labor Statistics. It is distributed both nationwide and abroad without charge to researchers using NLS data, as well as to other interested persons.

NLS User Services:	Center for Human Resource Research The Ohio State University 921 Chatham Lane, Suite 100 Columbus, Ohio 43221-2418 usersvc@postoffice.chrr.ohio-state.edu (614) 442-7366 (614) 442-7329 (Fax)	NLS documentation, data, and data updates:	usersvc@postoffice.chrr.ohio- state.edu
		NLS Web site:	www.bls.gov/nls
		BLS-NLS publications:	Jain_Rita@bls.gov (202) 691-7405
NLS Program Office:	National Longitudinal Surveys 2 Massachusetts Avenue, NE. Room 4945 Washington, DC 20212-0001 Attention: Rita Jain NLS_INFO@bls.gov (202) 691-7405 (202) 691-6425 (Fax)	NLS News Editor:	Donna S. Rothstein Rothstein_D@bls.gov
		NLS Program Director and Media Contact:	Charles R. Pierret Pierret_C@bls.gov (202) 691-7519