



U.S. Department of Labor  
Bureau of Labor Statistics

## Training Questions in the NLSY97

This article describes the types of information collected about the training and career preparation experiences of NLSY97 respondents. The training section of each NLSY97 survey examines the formal training experiences of respondents outside of their regular schooling. The section's main question series, asked only of respondents aged 16 and older as of the survey date, explores what kind of training youths obtain, where and when they are trained, and how the training is paid for. In addition, this article discusses the data available on school-based learning programs in which NLSY97 youths participated. Because these programs are intended to aid students in defining a career and obtaining a job, they can be considered in conjunction with training programs. The article closes with a brief comparison with training data gathered in other NLS surveys, particularly the NLSY79.

### Training programs

In general, the training section of the questionnaires gathers information on each training program with which a respondent is involved, including any programs in which he or she is currently enrolled. The information is collected in an event history format. In each round, the youth questionnaire asks respondents who have turned 16 since the last interview whether they have ever participated in any occupational training programs outside of their regular schooling. These programs are specified as courses designed to help people find a job, improve their job skills, or learn a new job. For each program, the survey then collects the following basic information:

- the type of program
- start and stop dates
- time devoted to the training
- periods of nonattendance lasting a week or longer during the training
- whether the program was completed (and if not, the reason)

In surveys after round 1, the training questions asked of respondents who had already turned 16 before a previous interview are slightly different. The survey first confirms the list of programs that were current at the time of the last interview, and then collects the details listed above for programs the respondent had attended since the last interview.

One key piece of information is the type of training program in which NLSY97 respondents participated. Answer categories for this question include the following:

- Adult Basic Education (pre-GED program)
- Apprenticeship program
- Correspondence course
- Formal company training run by employer
- GED program
- Government training
- Nursing school (LPN or RN)

If a government program paid for the training, respondents indicate the program in which they participated, such as Job Corps or Upward Bound, and the location at which the government training took place—for example, at a business school, nursing school, or community college.

An extended series of questions explores the financial arrangements for the training program. Specifically, respondents first state whether they received any allow-

ance or additional money for participating in the training. Respondents then report whether they or their families paid any fees for the training program. The survey also records whether respondents used any student financial aid or loans—and the amount of such aid or loans—to help pay for the training. Examples of financial aid or loans include Pell Grants, Stafford Loans, Supplemental Educational Opportunity Grants (SEOGs), or scholarships. Finally, respondents are asked whether an employer helped to pay for the training. If an employer did contribute, the respondent reports whether the training was required by the employer and the main reason why he or she undertook the training—for example, to get a promotion, to obtain a license or certificate, or to upgrade employee skills.

The next series of questions explores the content of the training program. Information is collected about whether any skill tests—such as reading, writing, or math—were given upon enrollment in the program. Information also is gathered about any tests or skill demonstrations used in the program to measure the respondent's progress.

Finally, respondents report the types of certificates, licenses, or degrees received from the program. Respondents are questioned about any classes taken while enrolled in the training program to improve their reading, writing, or math skills; to get their GED or high school diploma; or to learn English. The survey asks about additional activities in which the respondent was participating while involved in a training program, such as vocational training for a specific job, job search assistance classes, and on-the-job training. The skills learned in training programs also are determined.

Respondents report learning or improving computer skills, team work skills, health or safety procedures, equipment operation or repair, sales or managerial skills, and so on.

### **School-based learning programs**

In addition to the data on training, the NLSY97 interview also collects details on school-based learning programs in the schooling section of the youth questionnaire. Closely related to formal training, these programs often help students decide on a career, learn job skills, and look for a job.

In each survey, any respondent who is currently enrolled in a secondary school is asked whether the school has a day on which adults come to talk about their jobs. The survey then asks about the respondent's participation in any school-based learning programs. In round 1, the universe for these questions was respondents who reported a highest grade attended of 9th through 12th grade; in subsequent rounds, this question is addressed to all respondents currently enrolled at any level. The surveys collect information about a number of different school-based learning programs. Respondents first are asked about past or present enrollment in a career major program, and then indicate any other programs in which they have participated. Categories include a cooperative education program, internship or apprenticeship, job shadowing, mentoring, school-sponsored enterprise, and tech-prep program. Definitions of these categories are provided in the school-based learning section of the NLSY97 user's guide, available from NLS User Services.

Respondents describe the characteristics of the most recent program in more detail, providing information on the type of program, the number of days or weeks spent at a worksite, and the number of hours per day or week spent at the worksite. If the respondent was paid for participating in the program, he or she is asked to give the rate of pay. Other questions determine whether the respondent took any classes at the worksite, whether the employer wrote an evaluation of the respondent, and other program details.

### **Comparison to other NLS surveys**

Researchers may be interested in comparing training experiences of NLSY97 respondents with those of other NLS cohorts.

Information on training programs has been collected from NLSY79 respondents in each survey. Although surveys through 1986 focused on government training programs, the questions in later surveys are very similar to those in the NLSY97 questionnaires. Children of the NLSY79 who were aged 15 and older were asked comparable questions in 1994, 1996, and 1998. Finally, most surveys of the original cohorts have asked respondents about training both on and off the job, including information about the type of program, its duration, and any degrees, licenses, or certificates received.

Users should consult the *NLS Handbook* or the cohort-specific user's guide for more information about exact questions and universes. These documentation items are available from NLS User Services; see the back of this newsletter for contact information. □

## **Health Data in the NLS Mature and Young Women Surveys and the NLSY79**

The NLS surveys contain a number of questions about respondents' health. The early surveys of the NLS Mature and Young Women cohorts and the NLSY79 focused mainly on the ways in which health affected employment, with questions about whether respondents were able to perform certain types of jobs due to their health. In more recent interviews, the health sections have been expanded to collect more general information about both physical health and psychological well-being. This article provides a general overview of the health data collected from the mature and young women cohorts and the NLSY79, focusing on recent additions to the health modules.

**Mature and young women.** The health data collected from these two cohorts cover both physical and psychological well-being. Specific topics include the effect of health on employment, general physical health, and limitations on daily activities due to the respondent's health. Recent surveys have expanded the health data with additional questions on health problems and conditions. Finally, psychological well-being has been addressed in several surveys.

Because health may affect labor supply,

most interviews of the women's cohorts include questions about whether the respondent's health limited her work activity or prevented her from working altogether. Four mature women surveys and the 1978 interview of young women asked whether any of the respondent's health problems were the result of an accidental injury and whether her most serious injury occurred on the job. Respondents also were asked at multiple survey points to identify one or more specific workplace characteristics that they would have trouble working around because of their health (for example, places that were hot, damp, or that had fumes or noise).

The women's surveys also have included more general questions about physical health. To report their health status, respondents compare their personal health to that of other women their age. If they consider their health to have changed over a specific period, respondents state whether it has improved or declined. Self-reported height and weight of respondents also are available in the data collected in recent years. In 1991 and 1993, young women were questioned as to the number of times they were in the hospital during the past year and how often they were treated at clinics or by physicians over the past 5 years.

Limitations on daily activities are an important gauge of health status. Beginning in the 1970s, respondents periodically were asked if their health problems prevented them from performing a predetermined set of activities, such as walking, using stairs, stooping or crouching, and so forth. Respondents then reported whether they experienced certain health-related problems, such as pain, tiring easily, aches or swelling, fainting spells or dizziness, and anxiety or depression. In several surveys, mature women also stated whether they were able to go outdoors, use public transportation, and perform personal care without help.

Because of their increasing age and shifting health concerns, both the mature and young women have answered expanded health questions in more recent surveys. Several interviews of each cohort collected information on specific diseases that limited the respondent's work. The 1991 survey of young women included an extensive list of medical conditions, ranging from heart problems to allergies to

osteoporosis, and asked respondents to state whether they had ever experienced each condition. In 1997 and 1999, a series of questions asked all mature and young women respondents about high blood pressure, cancer, and heart disease, including limited information about treatment.

Additionally, two new series of questions were added in 1995 in response to the aging of the women's cohorts. Asked of both cohorts in 1995–99, the menopause series gathered current ovulation status, details of when and why a respondent's period stopped (if applicable), whether she had surgery to remove her ovaries or uterus, and details on symptom treatment. Information on respondents' driving activities was collected in 1995–99 for mature women and in 1995 only for young women. This series asked whether the respondent had ever driven, had driven anytime in the past year, the number of miles driven, whether she drove after dark, and when she stopped driving (if applicable).

Finally, a number of surveys have included questions that permit researchers to gauge the respondent's general mental health. In many years, respondents were asked to classify themselves as generally happy or unhappy with their lives. The 1989 survey of mature women and the 1993 survey of young women administered the full 20-item Center for Epidemiological Studies Depression (CES-D) scale. The 1995–99 surveys of both cohorts contained a reduced set of 7 items from the original 20-item CES-D scale, which measures symptoms of depression and discriminates between clinically depressed individuals and others. In several surveys, the mature and young women responded to a measure of internal-external locus of control, which reveals whether an individual perceives events as being under personal control or as unrelated to one's own behavior. Finally, in 1981, mature women respondents were administered a 10-item Affect Balance Scale that asks about positive and negative feelings in the weeks before the interview.

**NLSY79.** In each year, the NLSY79 surveys have included questions about health limitations on respondents' ability to work and some basic health characteristics. Beginning in 1998, two new health modules focused on the amount and kind of physical activity and exercise by all respondents and on the health of respondents at age 40.

The effect that the respondents' health has on the amount or kind of work they can do has been included to varying degrees in every survey. For respondents whose work was limited in some way by their health, the 1979–82 surveys gathered details on up to three limiting conditions. This series was shortened in 1983–87 and 1991 to ask about only the existence and duration of any work-limiting condition. Besides this basic data on work-limiting conditions, questions on work-related injuries or illnesses were included in NLSY79 surveys beginning in 1988.

Data on some basic health characteristics also have been gathered from NLSY79 respondents over the years. Height and weight have been reported numerous times; because it is more likely to fluctuate, weight has been reported in more years than has height. In addition, data on hair and eye colors were collected in 1985, and information on left- or right-handedness was obtained in 1993.

The NLSY79 survey included two additional health modules beginning in 1998 because, up to that point, the surveys had not collected information on a condition unless it was severe enough to limit the respondent's work. Questions on physical activity and exercise by all respondents and on the health of respondents at age 40 will help researchers identify connections between health and labor market activity before these respondents reach retirement.

Asked of all respondents, the first health module added in 1998 explores the level of physical activity and exercise that NLSY79 respondents typically experience at work and in daily life. It begins by addressing how often physical effort is required at the respondent's job and then asks the respondent to be more specific about the frequency and the tasks involved. The questions then shift to other daily activities involving physical effort, such as walking, dancing, or gardening. Respondents also report how often they engage in activities or sports (aerobics, swimming, biking, and so forth) and heavy housework such as scrubbing floors or washing windows.

The second new health module establishes a baseline health profile of respondents as they reach early middle age. This series of questions, asked only once of each respondent, is addressed to NLSY79 respondents during the first interview after their 40th birthday. The first part of this se-

ries is a seven-item version of the CES-D scale, which measures symptoms of depression in the past week. In addition to appearing in the 1998 health module addressed to respondents aged 40 and older, the full 20-item CES-D scale was addressed to all respondents in 1992, while the shortened 7-item version also was included for all respondents in 1994.

The 40-and-older module next questioned respondents about their last visit to a health professional. Collected are the dates of their last physical exam and last visit for any reason, as well as information on those who never visit a doctor. To improve researchers' understanding of hereditary health problems, the health and life status of the biological parents of respondents are gathered in this section. For deceased parents, the survey asks about cause of death, age at death, and which major health problems they had.

The SF-12 scale, reproduced in the next section of the 40-and-older health module, is a 12-question survey designed by John Ware of the New England Medical Center Hospital. A shortened version of the SF-36 Health Survey, it provides a measure of respondents' mental and physical health, regardless of their inclination to use formal health services. (More information about the SF-12 can be found at <http://www.outcomes-trust.org/> or by contacting NLS User Services.)

The last section of the module gathers extensive data concerning any major or minor health conditions that the respondent has. More information on the date diagnosed and other related details are collected if the respondent reports certain major conditions such as cancer, hypertension, diabetes, and so forth. Next, respondents state whether they have any of the less serious health problems listed; these range from asthma and back problems to anemia and osteoporosis. Finally, data regarding any injuries and sight and hearing problems are gathered.

**Children of the NLSY79.** A large amount of health data is available for the NLSY79 children and their mothers. Prenatal care, birth and postnatal care (including infant feeding), and the child's health history all are addressed in these data. Health questions continue to be asked as the children move into the young adult sample. These health topics were thoroughly discussed in

issue #96 (1999) of the *NLS News*. Interested users should contact NLS User Services for a copy of this newsletter or for child-specific documentation.

The above paragraphs provide a general overview of the health data collected from mature women, young women, and NLSY79 respondents. For more information about the years in which specific questions were asked, researchers should consult the cohort-specific user's guides, available from NLS User Services. Variables described in this article are discussed in the health section and the attitudes and expectations section of the guides. □

### The H. Gregg Lewis Prize

In June of 2000, the *Journal of Labor Economics* awarded the H. Gregg Lewis prize to Derek Neal for his article "The Complexity of Job Mobility among Young Men" (April 1999). The Lewis prize was established by the *Journal of Labor Economics* in honor of one of the founders of modern labor economics, H. Gregg Lewis (1914–92). A continuing reminder of Lewis's strong interest in students and in new developments in labor economics, the prize is awarded biennially to the author of the best article in the journal during the previous 2 years.

In his article, Neal examines the patterns of job changes among male respondents in the NLSY79. His research develops a model in which workers conduct a two-stage search: in the first stage, workers search for a good career match and, in the second stage, they search for a good employer match. From his model, Neal arrives at several testable implications. One implication is that complex job changes (that is, those involving a change in careers) should occur early in a worker's life, and simple job changes (those involving a change in employer within the same career) will tend to occur later. A second implication is that complex job changes should never follow simple job changes. These implications are then tested with work history data from the National Longitudinal Survey of Youth 1979 (NLSY79).

Critical to Neal's ability to test his model are data that keep track of the timing of all job changes, as well as information on whether the move involved a change in career, where a career change is

defined as a change in industry and occupation. The NLSY79 work history data file is uniquely suited to analyses like Neal's because the file provides researchers with a week-by-week longitudinal record for each NLSY79 respondent. The work history data track the vast majority of jobs held by NLSY79 respondents and record key information about each job, including the industry and occupation of the job.

As discussed in the article, Neal encountered problems in the work history data that involved missing industry and occupation codes as well as false changes in these codes. Neal used data from 1979-92 and, during these survey years, respondents were simply asked to describe the industry and occupation of each job, and survey staff coded those descriptions using the Census Bureau's coding frame. As noted in the article, variations in coding sometimes resulted in apparent changes in industry or occupation when no change had actually occurred. The problems that Neal encountered in the NLSY79 have been greatly reduced in data for 1994 and subsequent years, due to the adoption of computer assisted interviewing. Since 1994, industry and occupation coding has been done with dependent interviewing so that, if a given job had already been reported at a previous interview, the respondent is first asked whether the industry and occupation has changed. If not, the code from the previous interview is assigned, reducing spurious changes over time.

The article by Neal is an example of the type of research that can be done using the NLSY79 work history file. Researchers who are interested in using these data should contact NLS User Services by e-mail at [usersvc@postoffice.chrr.ohio-state.edu](mailto:usersvc@postoffice.chrr.ohio-state.edu) or at the address and phone number on the back of this newsletter. □

### Frequently Asked Questions

NLS User Services encourages researchers to contact them with questions and problems they have encountered while accessing and using NLS data and/or documentation. Every effort is made to answer these inquiries. Some recently asked questions that may be of general interest to NLS users are listed below with their answers. All of these questions refer to NLSY97.

*Q1: In round 2, there are two variables reporting whether female respondents have ever been pregnant. In R21850., 143 respondents answered yes to the question "Have you ever been pregnant?" In R21851., 190 respondents answered yes to the question "Have you been pregnant since the last interview?" Could you explain why these two variables do not seem to agree?*

**A1:** To understand these questions you will have to examine the skip patterns. In the first round, questions about pregnancy and sexual intercourse were addressed only to respondents aged 15 and older as of December 31, 1996. In round 2, these respondents were asked only about pregnancies since the last interview (R21851.). Some respondents who were not old enough to be asked about pregnancy in round 1 reached age 15 by December 31, 1997, and became eligible for these questions in round 2. These respondents answered the question about ever being pregnant (R21850.).

You should be aware that a similar skip pattern occurs for questions about substance use. Although there was no age restriction, some round 1 respondents refused to report whether they had ever used marijuana, smoked a cigarette, or drunk alcohol. In round 2, these respondents were again asked whether they had ever used these substances, while respondents who answered the questions in round 1 were asked only about use since the last interview.

*Q2: I am working with NLSY97 round 2 employment data. In the questionnaire, there are questions at the beginning of the section (for example, YEMP-305, YEMP-307, YEMP-1050) that do not have any reference numbers. I cannot find these questions on the CD-ROM. Why are these questions missing?*

**A2:** The questions at the beginning of the employment section are used to construct a roster that lists all of the respondent's employers. The rest of the section uses the roster as the basis for all of the other questions. For example, questions about the industry, occupation, fringe benefits, and so forth, of employer #01 refer to the employer listed first on the roster. However, this might not be the first employer that the youth reported in the interview. If we released those initial questions, it would be difficult for users to

match the information gathered later in the section with the correct employer from the beginning of the section. Instead, all of the information from these questions is included on the roster, which is much easier to use.

More information about the employer roster, its construction, and the data that it contains can be found in the section on employment in the *NLSY97 User's Guide* (available from NLS User Services).

*Q3: I have noticed that the cross-sectional sampling weight (question name CS\_SAMPLING\_WEIGHT) is zero for some respondents. I thought that all respondents would have a positive value for sampling weight. Could you please explain why these zeros occur?*

A3: The NLSY97 cohort is actually made up of two samples. The first represents a cross-section of all U.S. residents in the sample age range at the time the respondents were selected. The second, called the oversample or supplemental sample, contains additional respondents who were identified as black or Hispanic. This oversample enables researchers to analyze data for these minority groups.

If you are using the cross-section sample only, the cross-sectional sampling weight is appropriate. This weight variable provides researchers with information about the probability that a youth would be selected for the cross-sectional sample. However, if you are using the full sample, use the full sample weight variable (SAMPLING\_WEIGHT). This variable indicates the probability of selection for the total sample of respondents, including the oversample. Most research uses the weight variable for the total combined sample so that data from all respondents are included.

The sampling process and the construction of the sampling weights are described in chapter 2 of the *NLSY97 User's Guide*.

*Q4: Would you please tell me what "date" means in the variable CV\_HS\_DIPLOMA in the NLSY97 data? Is it the total months of the youth's age?*

A4: Variables like CV\_HS\_DIPLOMA are calculated using a continuous month scheme. In this system, month number 1 is January 1980, the month in which the oldest NLSY97 respondents were born. Month number 2 is February 1980, month number

3 is March 1980, and so on. The continuous month variables report the number of the month in which an event occurred, counting from January 1980. For example, a value of 240 associated with receipt of a diploma would mean that the respondent received the diploma in December 1999.

To find out how old the youth was at the time the diploma was received, you can use two other variables. First, subtract the respondent's age in months as of the interview date (R25534, in round 2) from the month in which the respondent was interviewed (R25682, in round 2) to find the number of the month in which the respondent was born. Then, subtract the birth month number from the number of the month in which the diploma was received to find the youth's age at that time.

Appendix 7 in the *NLSY97 Codebook Supplement*, available from NLS User Services, describes the continuous month system in more detail, and lists month numbers with the corresponding dates. □

### Completed NLS Research

The following is a listing of recent research based on data from the NLS cohorts that have not appeared in their current form in a previous issue of the *NLS News*. See the online *NLS Annotated Bibliography* at <http://www.nlsbibliography.org> for a comprehensive listing.

Barron, John M.; Ewing, Bradley T.; and Waddell, Glen R. "The Effects of High School Athletic Participation on Education and Labor Market Outcomes." *The Review of Economics and Statistics* 82,3 (August 2000): 409-421. [NLSY79]

Cawley, John. "Body Weight and Women's Labor Market Outcomes." NBER Working Paper No. 7841, National Bureau of Economic Research, August 2000. [NLSY79]

Currie, Janet and Thomas, Duncan. "Does Head Start Help Hispanic Children?" *Journal of Public Economics* 74,2 (November 1999): 235-262. [NLSY79, NLSY79 Children]

Ganzach, Yoav. "Parents' Education, Cognitive Ability, Educational Expectations and Educational Attainment: Interactive Effects." *British Journal of Educational Psy-*

*chology* 70 (September 2000): 419-441. [NLSY79]

Heckman, James and Vytlacil, Edward. "Identifying the Role of Cognitive Ability in Explaining the Level of and Change in the Return to Schooling." NBER Working Paper No. 7820, National Bureau of Economic Research, August 2000. [NLSY79]

Hockaday, Cathy; Crase, Sedahlia Jasper; Shelley, Mack C., II; and Stockdale, Dahlia F. "A Prospective Study of Adolescent Pregnancy." *Journal of Adolescence* 23,4 (August 2000): 423-438. [NLSY79]

Klawitter, Marieka; Plotnick, Robert D.; and Edwards, Mark E. "Determinants of Initial Entry onto Welfare by Young Women." *Journal of Policy Analysis and Management* 19,4 (Fall 2000): 527-546. [NLSY79]

Manski, Charles F. and Pepper, John V. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68,4 (July 2000): 997-1010. [NLSY79]

Morrison, Donna Ruane and Ritualo, Amy. "Routes to Children's Economic Recovery after Divorce: Are Cohabitation and Remarriage Equivalent?" *American Sociological Review* 65,4 (August 2000): 560-580. [Children of the NLSY79]

Raley, R. Kelly; Harris, Kathleen Mullan; and Rindfuss, Ronald R. "The Quality and Comparability of Child Care Data in U.S. Surveys." *Social Science Research* 29,3 (September 2000): 356-381. [NLSY79]

Reagan, Patricia B. and Olsen, Randall J. "You Can Go Home Again: Evidence from Longitudinal Data." *Demography* 37,3 (August 2000): 339-350. [NLSY79]

Smith, Judith R.; Brooks-Gunn, Jeanne; Klebanov, Pamela K.; and Lee, Kyunghee. "Welfare and Work: Complementary Strategies for Low-Income Women?" *Journal of Marriage and the Family* 62,3 (August 2000): 808-821. [Children of the NLSY79]

Williams, Donald R. "Consequences of Self-Employment for Women and Men in the United States." *Labour Economics* 7,5 (September 2000): 665-687. [Young Women, NLSY79]

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

U.S. DEPARTMENT OF LABOR  
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 Contact Information

NLS News is published quarterly by the 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: <http://stats.bls.gov/nlshome.htm>

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 Editors: Mary Joyce and Donna S. Rothstein  
Joyce\_M@bls.gov  
Rothstein\_D@bls.gov

NLS Program Director and Media Contact: Michael W. Horrigan  
Horrigan\_M@bls.gov  
(202) 691-7386