Updated NLSY79 Release Includes New AFQT/ASVAB Variables

A January 2014 release of the 1979–2010 National Longitudinal Survey of Youth (NLSY79) main dataset includes new variables associated with the Armed Forces Vocational Aptitude Battery (ASVAB), a series of tests administered to NLSY79 respondents back in the summer and fall of 1980.

The ASVAB consists of a battery of 10 tests measuring knowledge and skill. A composite score derived from select sections of the battery can be used to construct an approximate and unofficial Armed Forced Qualifications Test (AFQT) score for each NLSY79 respondent. (Many researchers use the AFQT score as a measure of IQ.)

The updated NLSY79 release contains new, individual AFQT scores, including responses for each item (scored as correct or incorrect) in arithmetic reasoning (AR), mathematics knowledge (MK), word knowledge (WK), and paragraph comprehension (PC). New norms for these subsets were generated with the use of item response theory (IRT) theta scores and the ASVAB sampling weights. To norm the subtests, respondents were grouped into 4-month age intervals for each birth year. Percentiles and z-scores are available for MK, AR, PC, WK, the combined math (MK and AR) scales, and the combined verbal (PC and WK) scales.

Factor analysis suggests that PC and WK measure a single verbal factor and that MK and AR measure a single quantitative factor, explaining why each of those pairs of test results were combined.

More information about these variables and their creation can be found in the NLSY79 Codebook Supplement Appendix 24: www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/codebook-supplement/nlsy79-appendix-24-reanalysis-1980.

The updated release includes NLSY79 data from rounds 1-24 of the survey. The variables can be accessed for use through NLS Investigator (www.nlsinfo.org/investigator) by searching on “ASVAB” as the Word in Title search in the NLSY79 rounds 1-24 database.

NLSY79 Celebrates 35th Birthday

Thirty-five years ago this year, data collection began for the NLSY79. That year, 12,686 men and women born during the years 1957 through 1964 were interviewed for the first survey.
Since then, the NLSY79 cohort has been reinterviewed regularly—annually from 1979 to 1994 and biennially from 1996 to 2012—for a total of 25 times. Data for rounds 1 through 24 are now available; data for round 25 will be released in the fall of 2014. Round 24 collected data from 7,565 members of the cohort, or about 80 percent of the eligible respondents. (The original sample included supplemental samples of active military and economically disadvantaged non-Black and non-Hispanic youths; both of these groups had been dropped from the original sample by 1991. In addition, as of the 2010 survey year, 573 NLSY79 respondents were reported as deceased.) Survey respondents were ages 14 to 22 when first interviewed in 1979; in round 24 (2010), the men and women in the sample were ages 45 to 53.

The U.S. Department of Labor designed the NLSY79 cohort to replicate much of the analysis gained through the NLS of Young Women and the NLS of Young Men, both of which began in the 1960s. As with the original cohorts, the NLSY79's primary focus has been labor force behavior; but the content also includes detailed questions on education, training, income and assets, health and physical condition, marital and family characteristics, attitudes, and military experiences. In addition, the full ASVAB was administered to 94 percent of the sample respondents in 1980.

As the NLSY79 sample has matured, the health-related questions have expanded to include a baseline profile of the respondents' overall health as they turned 40 and then 50 years old. (See the HEALTH MODULE 40 & OVER and HEALTH MODULE 50 & OVER areas of interest in the NLS Investigator.) Starting in 2006, respondents also have answered questions about any preliminary retirement preparation they may have initiated, such as consulting a financial planner and calculating retirement income needed. (See the retirement area of interest.)

Interviews and assessments of the children of the NLSY79 female respondents began in 1986, providing researchers with an even richer body of information than that for men. (See the NLSY79 Child and Young Adults surveys for more information.)

To access the NLSY79 1979-2010 public data, go to NLS Investigator at www.nlsinfo.org/investigator.

New NLSY79 Childhood Health and Adversity Variables Available

Seven NLSY79 retrospective variables about childhood health and adversity (including health problems, living in households with mental health or substance abuse issues and with dysfunctional parents) are now available as a zip file for NLSY79 users. (Respondents answered questions pertaining to these variables during the 2012 survey year.) The variables have been made available ahead of the regular 1979–2012 dataset release, scheduled for fall of 2014.

Respondents answered the following questions about their lives from birth to 17 years:

- Q11-RCH-HLTH-1  Consider your health when you were growing up, from birth to age 17. Would you say your health during that time was excellent, very good, good, fair, or poor?
- Q11-RCH-HLTH-2  From birth to age 17, did you ever have a hospital stay lasting at least two weeks?
- Q11-RCH-HLTH-3  From birth to age 17, were you ever confined to bed or home for four or more weeks because of a health condition?
- Q11-RCH-HLTH-4  Before age 18, did you live with anyone who was depressed, mentally ill, or suicidal?
- Q11-RCH-HLTH-5  Before age 18, did you live with
anyone who was a problem drinker or alcoholic?

- Q11-RCH-HLTH-6 Before age 18, how often did a parent or adult in your home ever hit, beat, kick or physically harm you in any way? Do not include spanking. Would you say never, once, or more than once?

- Q11-RCH-HLTH-7 Before age 18, how much parental love and affection did you receive growing up? Would you say a great deal, quite a lot, a little, or none at all?

To access these variables via a zip file, go to www.nlsinfo.org/content/cohorts/nlsy79/other-documentation/errata and click on “Release of 2012 Questions on Childhood Health and Adversity.” Note that the assigned reference numbers accompanying the variables are preliminary and will be changed with the full data release.

---

**NLS Researchers Win 2013 IZA Young Labor Economist Award**

The Institute for the Study of Labor (IZA) recently presented its 2013 IZA Young Labor Economist Award to three researchers doing NLS-related research. Martha Bailey (University of Michigan), Brad Hershbein (Upjohn Institute), and Amalia Miller (University of Virginia) received the prize for their article “The Opt-In Revolution? Contraception and the Gender Gap in Wages,” *American Economic Journal: Applied Economics* 4, 3 (July 2012): 225–254.

The prize-winning research article tested the influence of a contraceptive pill (“the Pill”) on the gender wage gap, using data from the National Longitudinal Survey of Young Women. The NLS Young Women cohort is one of the original cohorts of the NLS, which followed a sample of U.S. women born in the 1940s and 1950s. An annotated citation of the article is available in the NLS Annotated Bibliography (https://nlsinfo.org/bibliography-start). The article itself can be found on the American Economic Association’s website (http://pubs.aeaweb.org/doi/pdfplus/10.1257/app.4.3.225).

Headquartered in Bonn, Germany, IZA is a private independent economic research institute focused on the analysis of global labor markets. The IZA Young Labor Economist Award was established in 2006 to honor researchers who are under the age of 40 at the time of publication.

---

**Did You Know? NLSY97 Dataset Has Skin Tone Variable**

Researchers interested in studying discrimination based on skin color (or “colorism”) can pursue that research opportunity in the NLSY97 dataset. Interviewers of round-12 respondents recorded the color of the respondent’s skin, using a skin color card as a guide to determine a hue that most closely corresponded to the respondent’s facial coloring. The skin color gradients ranged from 1 to 10, with 1 being the lightest and 10 the darkest. If a respondent was not a round-12 (2008 survey year) participant, interviewers recorded the skin color in round 13 or 14.

The NLSY97 variable can be found in the NLS Investigator (www.nlsinfo.org/investigator) by searching the NLSY97 database by “Word in Title” is “Skin” or the question name “YIR-530.” A PDF of the color guide card can be found at the NLSInfo.org website within the “Interviewer Remarks, Characteristics, and Contacts” section of the NLSY97 Cohort area. (See paragraph on “other respondent characteristics.”)
NLS Bibliography Reaches 8,000 Citations

Users can now find more than 8,000 NLS-related research citations in the NLS Annotated Bibliography at www.nlsbibliography.org. In January 2014, the official count of NLS-related citations stood at 8,043. Category tallies at that time included the following:

- 304 chapters
- 3,400 journal articles
- 1 piece of legislation
- 110 master's theses
- 204 monographs
- 193 newspaper articles or mentions
- 1,288 conference presentations
- 1,007 Ph.D. dissertations
- 426 reports
- 4 webcasts
- 1,076 working papers
- 30 honors theses (B.A.)

Citations generally are accompanied by an abstract and list details, including which NLS cohorts were used for the analysis. Bibliographic entries often include a publisher's link to the abstract or article. Researchers are encouraged to share information about their own NLS-related research by going to the bibliography website (www.nlsbibliography.org) and providing details at the link labeled “Submit Citation.”

NLSY Terms to Know: Asterisk Tables

Within the NLSinfo.org website are cohort-specific tables commonly referred to as “asterisk tables.” These tables summarize selected variables across all survey rounds and are divided into drop-down topical sections. The tables allow researchers to ascertain quickly what
topics the various NLS datasets have to offer and how often a question is repeated. If a variable is present in a given round, the box usually is marked with an asterisk. For instance, in the NLSY97 dataset, a question about whether a respondent has health insurance coverage has been asked in every round since round 6, so users will see an asterisk in the boxes for round 6 through 15.

To access these tables, go to NLSinfo.org and click on a specific cohort; then select “Asterisk Tables.”

Additions and Corrections: Errata

A zip file is now available that contains the dates on which the NLSY79 Mother Supplements were conducted in the 1994–2000 survey years. These interview dates were inadvertently omitted from the 1986–2010 NLSY79 Child/Young Adult data release and will be included in the next public release. The Mother Supplement (MS) date-of-interview data are provided in the formats .dct, .dat, and .csv. Statements are included that enable the user to read these files in STATA, SPSS, or SAS, respectively. The comma-delimited file (MS_Dates.csv) has a header row with the names of the variables, so it may be opened directly by spreadsheet or similar programs. The zip file link containing the data and documentation for these variables can be found at the Children and Young Adult cohort area on the NLSinfo.org website at www.nlsinfo.org/content/cohorts/nlsy79-children/other-documentation/errata.

Note: The interview dates on this raw data file have not been edited or cleaned and appear as they were recorded by interviewers on paper questionnaires in the field. Because, in the early survey rounds, the Child and Mother Supplements were nearly always administered on the same date, interviewers did not always record the interview date on both surveys. Users who wish an accurate child age at the date of the child’s interview should use the CSAGE and MSAGE variables provided in the CHILD BACKGROUND Area of Interest on the public NLSY79 Child/Young Adult data file in NLS Investigator (www.nlsinfo.org/investigator). Users who wish to identify a child’s interview status in a particular survey round should rely on the child sampling weights (values greater than 0) for each survey year.

Looking Back: CAPI Interviewing Began 21 Years Ago

Happy 21st birthday to CAPI! Twenty-one years have passed since Computer Assisted Personal Interviews (CAPI) became part of the regular routine of the NLS interviewing process. Prior to the mid-1990s, NLS interviewers used the traditional paper and pencil data collection method to conduct interviews. In 1993, the NLSY79 field efforts switched to a CAPI approach, using a survey instrument loaded onto a laptop computer. (That switch was preceded by mode effect experiments in 1989 and 1990, in which respondents were randomly assigned either paper or CAPI instruments.) The NLS Young Women and Mature Women surveys (the female components of the original cohort surveys) followed suit and began using CAPI instruments in 1995. The NLSY97 main surveys were administered using CAPI instruments from the first interview (1997) on.

Although computer-administered surveys are considered standard practice today, the introduction of CAPI was groundbreaking technology in the 1990s. CAPI allowed for more complex questionnaire design.

- More complex questionnaire programming. The questionnaires could be more easily tailored for individual participants by making it more feasible for branching or skipping within the survey instrument.
• Bounded interviewing. In CAPI, some responses (such as dates) are bounded or limited to specific ranges, so that when a participant gives an answer outside those ranges, he or she is questioned about the response and prompted to give a more accurate one.

• Fewer skip-pattern errors. Skip patterns are the paths each participant takes through the instrument. Not every question is answered by every respondent. (For instance, some sections are answered only by females; other sections are answered only by those who report a spouse or partner, etc.) In pre-CAPI days, the interviewer needed to hand check previous responses or an information sheet to determine whether a respondent navigated through or skipped past a particular section. Interviewer skip-pattern errors allowed some participants to answer questions that did not apply to them, while other respondents were skipped out of sections that they were supposed to answer. With CAPI programming, these skip-pattern errors are reduced considerably.

FAQs

Frequently Asked Questions

The NLS staff encourages researchers to contact NLS User Services (usersvc@chrr.osu.edu) with questions and problems encountered while accessing and using NLS data or documentation. Every effort is made to answer these inquiries. Following are some recent questions and answers that may be of interest to NLS users:

1. Does the NLSY97 dataset provide any information (name or location, for instance) about the colleges attended by NLSY97 respondents?

   The Integrated Postsecondary Education Data System (IPEDS) assigns identifying code numbers for colleges, universities, and technical and vocational postsecondary institutions in the United States. These IPEDS numbers (for the colleges attended, as well as the colleges applied to) are part of the NLSY97 geocode dataset, a restricted-use dataset that also contains a variety of statistics about the counties where respondents have lived. For more information on how to apply for obtaining the geocode data CD, go to [www.bls.gov/nls/nlsgeo97.htm](http://www.bls.gov/nls/nlsgeo97.htm).

2. I’m interested in the number of employees (if any) an NLSY79 respondent has supervised at work. Is this number available?

   These questions about supervising were asked in the 1988, 1989, and 1990 survey years for the respondent’s current or most recent job (this job is known as the “CPS job,” because the questions designed to identify it replicate questions from the Current Population Survey) and for all jobs in 1996 and 1998. The 1996 and 1998 questions about supervising were more detailed. In addition to providing the number of employees they supervised, respondents answered questions about the gender breakdown of those supervised, how much responsibility the respondents had in determining the pay, promotions, and job tasks for their supervisees, and how closely they monitored or supervised these employees. Respondents also reported if they themselves had a supervisor and, if so, the gender of that supervisor and the number of people the respondent’s supervisor had supervised.

3. In the NLS Investigator, most NLSY79 reference numbers start with the letter R (for instance: R230320) but some begin with other letters. Is there a significance in the different letters?

   Reference numbers identify individual variables. Back in the first days of the NLSY79, all reference numbers began with R, but as data collection continued through the years and each of the collected variables needed reference numbers, the R number possibilities ran out and reference number assignment was expanded to include other letters. In addition to R, they can now begin with E, G, H, T, or W (and a single A number
4. I am using the NLSY79 dataset and NLSY79 Child and Young Adult dataset to predict child educational outcomes. For part of my analysis, I need the child’s father’s highest level of education. If the father is living in the household, what is the best variable to use for his level of education?

Using the Child/Young Adult dataset in NLS Investigator (www.nlsinfo.org/investigator), first select DADHMyyyy (Does father of child (living in HH) live in this household?) to identify whether the father is in the household. Then, within the same survey round, use HGCPTRyyyy (highest grade completed by partner in household of mother) and HGCSPSyyyy (highest grade completed by spouse in household of mother) to identify the father’s highest grade completed. These variables are based on mothers’ reports from the NLSY79 and have been “flipped” to be child-based variables.

5. While looking at income variables in the NLSY97 dataset, I ran across three respondents who reported a huge amount of negative income on CV_INCOME_GROSS_YR. Are these legitimate cases, or should I be assuming that the numbers are data errors?

All three of these cases are legitimate. The large negative-income amounts came from negative farm/business income—a situation that is not uncommon.

6. Are there ever cases where one respondent appears in multiple NLS surveys, such as in the NLSY79, the NLSY97, the NLS Young Women/Mature Women, or the NLS Young Men Survey?

Each of the samples for these surveys were drawn independently, and no respondent belongs to more than one cohort survey.

7. My research interest is in rural adolescents and adults. For both the NLSY97 and the NLSY79, is there a way of pulling out just those respondents who are living in rural areas?

In the NLSY79, there is a constructed variable named “URBAN-RURAL.” (See T31101.00 in the 2010 survey year.) In the NLSY97, the created variable name is “CV_URBAN-RURAL.” (See T66633.00 in 2011.) These variables assign the respondent an urban or rural classification based on where the respondent lives at the time of each interview. For details on how the classifications are assigned, see Appendix 6 (NLSY79) and Appendix 4 (NLSY97) under the two cohorts’ “Other Documentation” links at the NLSinfo.org website.

8. In the NLSY97, there seem to be a few cases where roster-based information about biological children does not quite match up with the constructed variables about the total number of resident biological children the respondent has. Could you explain why this is?

In each round, there are a small number of children whose reported relationship changes from “step” to “bio” or “adoptive,” and vice versa. The most common scenarios for this situation appear to be the following:

- The respondent adopts a previously reported stepchild.
- A respondent who initially thought that a child was his biological child later learns that the child is not his biological child (or perhaps the respondent chose to report the child as biological until his relationship with the child’s biological mother ended).
- After a marriage, a respondent who initially reported a child as “step” starts reporting the child as biological.

When the child’s relationship to the respondent changes—whether in a plausible or an implausible way—there is no attempt on the part of NLS staff to go back to earlier rounds and change the data. For example, a child recorded as a biological child in round 9 could be recorded as a stepchild in round 10.

9. Which NLSY97 variables note whether the job is self-employed?
Variables with the prefix YEMP_SELFEMP indicate self-employment. These variables first appeared in 2000. Prior to that time, if the respondent indicated that he or she was self-employed, the respondent was routed to the freelance section of the survey. In 2000, respondents born between 1980 and 1982 and who were self-employed were asked questions about their job and were not routed to the freelance questions, whereas the younger respondents still received the freelance questions. In 2001, those self-employed who were born between 1980 and 1983 got the self-employed questions. By 2002, there was no age restriction, and from that point on, all respondents who were self-employed got the self-employment questions.

### 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](http://www.nlsbibliography.org) for a comprehensive listing of NLS-related research.)


Houle, Jason N. “Disparities in Debt: Parents’ Socioeconomic Resources and Young Adult Student Loan Debt.” *Sociology of Education* 87,1 (January 2014): 53-69. [NLSY97]


attainment of youth.” Economics of Education Review 37 (December 2013): 165-182. [NLSY97]


Malhotra, Rahul, Truls Østbye, Crystal M. Riley, and Eric A Finkelstein. “Young adult weight trajectories through midlife by body mass category.” Obesity 21,9 (September 2013): 1923-1934. [NLSY79]


Moilanen, Kristin L. and Yuh-Ling Shen. “Mastery in Middle Adolescence: The Contributions of Socioeconomic Status, Maternal Mastery and Supportive-Involved Mothering.” Journal of Youth and Adolescence 43, 2 (February 2014): 298-310. [Children of the NLSY79, NLSY79 Young Adult]


Ricketts, Comfort F., Jon P. Rezek, and Randall C. Campbell,


Stablein, Timothy and Allison A. Appleton, “A Longitudinal Examination of Adolescent and Young Adult Homeless Experience, Life Course Transitions, and Health.” Emerging Adulthood 1,4 (December 2013): 305-313. [Children of the NLSY79, NLSY79 Young Adult]

Tewksbury, Richard, George E. Higgins, and David Patrick Connor. “Number of Sexual Partners and Social Disorganization: A Developmental Trajectory Approach.” Deviant Behavior 34,12 (December 2013): 1020-1034. [NLSY79 Young Adult]


NLS Contact Information

*NLS News* is a quarterly newsletter that provides information about changes in labor market activities and other significant events that occur in the lives of several groups of men and women over time. Also, it includes information about new data releases, completed NLS research, and other information of interest to researchers. *NLS News* is available at [http://www.bls.gov/nls/nlsnews.htm](http://www.bls.gov/nls/nlsnews.htm).

**NLS User Services:**
Center for Human Resource Research  
NLS documentation, data, and data updates  
The Ohio State University  
921 Chatham Lane, Suite 100  
Columbus, OH 43221-2418  
usersvc@chrr.osu.edu  
(614) 442-7366 / (614) 442-7329 FAX  
www.nlsinfo.org

**NLS Program Office:**
National Longitudinal Surveys  
2 Massachusetts Avenue, NE  
Room 4945  
Washington, DC 20212-0001  
NLS_INFO@bls.gov  
(202) 691-7405 / (202) 691-6425 FAX  
www.bls.gov/nls

**NLS Program Director and Media Contact:**
Charles R. Pierret  
Pierret_C@bls.gov  
(202) 691-7519

**NLS News Editor:**
Donna S. Rothstein  
Rothstein_D@bls.gov