

DOCUMENTATION OF NONRESPONSE AND CONSISTENCY OF DATA CATEGORIZATION ACROSS NCES SURVEYS

Steven Fink, Mehrdad Saba, Michael Chang, Samuel Peng

Steven Fink, Synectics for Management Decisions, Inc., 3030 Clarendon Blvd #305, Arlington VA 22201

KEY WORDS: NCES survey documentation, Unit response rate, Item response rate, Data categorization

During the last two decades, interest and concern have been growing regarding nonresponse (unit and item) in federal surveys because of how this issue relates to survey data quality. This report provides a systematic review of past and ongoing research on availability and calculation of response rates (both unit and item), and uniformity of several response categories of several NCES surveys.

Unit nonresponse is vitally important to users of federal surveys. Several attempts have been made to standardize response rate definitions. For example, the Council of American Survey Organizations (CASRO) reviewed response rate definitions with the intent of trying to establish uniformity of definitions across surveys (CASRO 1982). More recently, the Subcommittee on Nonresponse, commissioned in 1991 by the Office of Management and Budget's Federal Committee on Statistical Methodology (FCSM) provided the following recommendations:

- 1) Survey staffs should compute response rates in a uniform fashion over time and document response rate components on each edition of a survey.
- 2) Survey staff for repeated surveys should monitor response rate components (e.g., refusals, not-at-homes, out-of-scopes, address not locatable, postmaster returns, etc.).
- 3) Response rate components should be published in survey reports; readers should be given definitions of response rates used, including actual counts, and commentary on how response rates affect survey data quality.
- 4) Some research on nonresponse can have real payoffs. It should be encouraged by survey managers as a way to improve the effectiveness of data collection operations.

Item nonresponse occurs when the person participates in the survey but fails to answer some of the questions. It may arise for several reasons, including lack of respondent information, refusals, and inconsistency with other responses. This last category may include an inconsistency arising from a coding or keypunching error occurring in the transfer of the response from the answer sheet to the computer data file (Kalton 1983).

This report also examines consistency in data categorization. We identify commonly used demographic variables in NCES surveys and explore question wording and response categories of nine demographic survey items.

NCES Databases

We chose a mix of 13 surveys from NCES sample populations. Among these surveys, NALS and NHES were non-school-based surveys.

Elementary / Secondary Education

- o Schools and Staffing Survey (SASS, 1990-91)
- o Teacher Follow-up Survey (TFS, 1991-92)

Postsecondary Education

- o National Household Education Survey - Adult Education Component (NHES, 1993)
- o National Postsecondary Student Aid Study (NPSAS, 1990)
- o National Survey of Postsecondary Faculty (NSOPF, 1993)
- o Recent College Graduates Study (RCG, 1991)

Educational Assessment

- o National Assessment of Educational Progress (NAEP, 1990)
- o National Adult Literacy Survey (NALS, 1992)

National Longitudinal Studies

- o Baccalaureate and Beyond Longitudinal Study (B&B, 1993-94)
- o Beginning Postsecondary Student Longitudinal Study (BPS, 1992)
- o High School and Beyond (HS&B, 1992-93)
- o National Education Longitudinal Study of 1988 (NELS:88, 1991-92)
- o National Longitudinal Study of 1972 (NLS-72, 1979-80).

Response Rate Information

Below we present technical issues affecting responses in NCES surveys.

Unit Response Rates

Unit response rate refers generally to how many survey instruments were returned/completed. Below we present responses to seven questions relevant to unit response

rates. A common definition of unit unweighted response rate is the ratio of the number of units with completed interviews (the units could be telephone numbers, households, or individuals) to the number of units sampled and eligible to complete the interview.

A. Are unweighted unit response rates calculated consistently?

Eight of the 11 NCES surveys which provided calculations in their documentation used the same basic formula to calculate unweighted response rates. However, the definitional terms and level of detail varied from one survey to another. Several surveys used different names to describe the same response rate calculation. For example, a "locating" response rate (BPS) appears similar to a "screener" response rate (NHES), and an "interview" response rate (BPS) appears similar to an "extended interview" response rate (NHES). In NAEP, the unit response rate was referred to as the participation rate (also called the cooperation rate). RCG provided figures to calculate the unweighted response rate without specifying the formula to use, and NALS presented the unweighted response rates without providing the formula. TFS did not provide unweighted response rates.

B. Are substitute schools used in the calculation of unweighted unit response rates?

Among the surveys we examined, the answer is basically "no." Survey methodology sometimes allows substitute schools to replace nonparticipating schools in a selected sample; for example, when a selected school does not respond to a survey, another school with similar characteristics was asked to fill in. However, of the 11 school-based surveys we examined, only three-- NLS-72, NELS:88, and HS&B (all National Longitudinal Studies)--used substitute schools when calculating unweighted unit response rates, and NLS-72 and HS&B also calculated what unit response rates would be without the substitute schools.

C. Are weighted unit response rates calculated consistently?

A weighted unit response, with the effects of the sampling design incorporated into the calculation gives more accurate response information than the unweighted rate. Therefore, the weighted unit response is often a better measure for deciding whether further nonresponse studies should be conducted. Among the eight surveys identified in A, four did not provide weighted response rates. The other four (NELS:88, NPSAS, SASS, and HS&B) used the same basic formula. NELS:88 also used an additional weight: student design weight or school design weight. RCG used the standard formula. NALS provided no

formula in its documentation (but rates were tabulated), and BPS and NHES utilized the sampling design to compute rates. TFS used a subsample of a previously conducted survey for its sampling frame. The weighted response rate was calculated as the product of SASS teacher list response, the SASS teacher response rate, and the TFS teacher response rate. Although NALS did not provide the formula for calculating weighted response rates, the documentation does state that the weighted response rates were calculated by applying the sampling weight to each individual to account for his/her probability of selection into the sample.

D. Are school/institution level response rates weighted by enrollment?

School or institution may be used as the sampling frame variable because small schools may have unique characteristics not associated with larger schools. Enrollment may then be used to weight the data. Only three of the school-based surveys examined--RCG, HS&B, and NELS:88--provided information on school level response rates, with only two weighting the response rate directly by school enrollment. In RCG, institution weight takes into account the sampling probability of the institution, which is proportionate to enrollment size. In HS&B, the school sampling probability was also proportional to the estimated enrollment.

E. Is there an intensive follow-up of nonrespondents? If so, were results built into the response rates?

One of the most pervasive and challenging sources of nonsampling error in estimates from sample surveys is the bias associated with nonresponse. Respondents may differ significantly from nonrespondents. Most of the time, funds are not available to conduct respondent follow-ups and convert every nonrespondent. One way to reduce bias is to take a subsample of nonrespondents and conduct an intensive follow-up to get everyone to respond. Different modes of data collection are used to encourage respondents to return their survey. However, only NELS:88 took a subsample of nonrespondents and conducted an intensive follow-up. NCES usually attains relatively high response rates and quality data. This may explain why intensive follow-up is usually not conducted.

F. Are unit response rates tabulated by the frame variables?

Frame variables such as sector and school or institution type are often used to select the samples from the populations. Tabulating unit response rates by frame variables helps to identify low and high response in certain strata. This practice can help researchers identify and perhaps improve future response in low response

strata or frames. Frame variables were used to tabulate unit response rates for all but three of the surveys we examined (NSOPF, NALS, and NLS-72). Given that most of those surveys were school-based, institution/school type was the frame variable most commonly used in the tabulations.

G. How is the overall unit response rate (covering all stages of sampling) calculated?

An overall unit response rate is only discussed for surveys using multi-stage sampling designs. Typically, an overall unit response rate for a two-stage sample survey is calculated as follows:

Overall unit response rate =
(First stage rate * Second stage rate)

The seven surveys which did calculate overall response rates all used this basic formula, although language differs. Four of the surveys we examined (NSOPF, HS&B, NALS, and NELS:88 2nd Follow-up) did not calculate an overall response rate or did not mention ways of calculating this type of response rate.

Item Response Rates

Item nonresponse has the effect of diminishing the number of observations that can be used in calculating statistics from affected data elements and thus increases sampling variances (Ingels et al. 1994). NCES standards stipulate that item response rates "are to be calculated as the ratio of the number of respondents for which an inscope response was obtained divided by the number of completed interviews for which the question was intended to be asked." Below we present responses to three questions relevant to item response rates, followed by two questions examining nonresponse research and availability of nonrespondents on data files.

H. Are unweighted item response rates calculated consistently?

SASS, BPS, HS&B, NSOPF, and B&B used the NCES standard as the means of calculating unweighted item response rate, although the exact wording varied. B&B and NSOPF defined item nonresponse. (It should be noted that the documentation for three of those surveys--BPS, HS&B, and B&B--did not explicitly identify the item response rate definition as unweighted or weighted.) A look at four surveys examine "Don't know" responses as a source of possible difference when calculating inscope responses. B&B provided separate tabulations for refusals and "don't know" responses, and presented a combined nonresponse rate integrating the two. NELS:88 used "don't know" as a

valid response to certain questions, so it did not classify "don't know" as a nonresponse. In RCG, item nonresponse included responses of "don't know," "refused," and "not ascertained." However, there were no questions where "don't know" was considered a response (Westat, Inc. 1994). Finally, for NSOPF, "don't know" was included as an item nonresponse even in cases where "don't know" was an explicit response category for the item (Abraham et al. 1994).

I. Are weighted item response rates calculated consistently?

Only three surveys, NELS:88, RCG, and TFS, defined weighted item response rates. All used the standard definition, although exact wording varied. Considering unweighted and weighted item response rates together, all eight surveys which provided definitions used the NCES standard definition.

J. Are item response rates tabulated by subgroups?

Presentation of item response rate information varied considerably. At one end of the spectrum, RCG and NSOPF presented item response rates for all questions. At the other end, NAEP, NALS, and NLS-72 did not tabulate any item response rates. NHES and NPSAS are the two surveys which used subgroups in their presentations on item response rates. The tabulated subgroups on NHES were participation items, course or activity items, and sociodemographic items. NPSAS used four subgroups: student characteristics, enrollment variables, costs, and aid eligibility variables. The other surveys took one of two approaches. B&B and HS&B simply presented item response rates for a selected number of items. The rest--SASS, TFS, BPS, and NELS:88--presented information only on those items which exceeded a designated response rate (or nonresponse rate) threshold.

K. Is there any research dealing with nonresponse rates; e.g., adjustment, incentives, etc.?

We identified research done on three surveys--SASS, NSOPF, and NHES. For SASS, there were several reports (often in the form of memos or articles) examining characteristics of nonrespondents. NSOPF included an experimental design to examine the effect incentives and prompts can have on nonresponse rates. For NHES, there were internal memos and a report examining telephone undercoverage. One reason there may be so little research on nonresponse in NCES surveys is that response rates are generally high. As the following table shows, the majority of unit response rates exceed 80 percent (see Table 1).

Table 1: Unit Response Rates, by NCES Survey

Survey Name	Unweighted (%)	Weighted (%)
Elementary/Secondary Education		
Schools and Staffing Survey (SASS)		
School Administrator (public)	96.9	96.7
School Administrator (private)	91.1	90.1
TDS (public) ¹	93.7	93.5
TDS (private)	84.8	83.9
School (public)	95.0	95.3
School (private)	85.1	83.9
Teacher (public) ²	91.5	90.3
Teacher (private) ²	83.1	83.6
Teacher Follow-up Survey(TFS)		
Current (public)	not avail.	97.4
Current (private)	not avail.	92.4
Former (public)	not avail.	96.2
Former (private)	not avail.	94.1
Postsecondary Education		
National Household Education Survey (NHES)	82.1 ³	not avail.
National Postsecondary Student Aid Study (NPSAS)		
Institutions	95	89
Students	77	76
National Survey of Postsecondary Faculty(NSOPF)	not avail.	not avail.
Recent College Graduates Study (RCG)	83.1	83.2
Educational Assessment		
National Assessment of Educational Progress (NAEP)		
School	86.0	not avail.
Student	87.4	not avail.
National Adult Literacy Survey (NALS)	89.1	not avail.
National Longitudinal Studies		
Baccalaureate and Beyond Longitudinal Study (B&B)	85.4	not avail.
Beginning Postsecondary Student Longitudinal Study (BPS)	96.1	not avail.
High School and Beyond (HS&B)	not avail.	86.1
National Education Longitudinal Survey of 1988 ((NELS:88)	92.5	91.5
National Longitudinal Study of 1972 (NLS-72) Target Sample (4th Follow-up)	89.3	not avail.

¹ Combined School and TDS ³ Using Business office method

² Percent of eligible teachers in sample responding

L. Has any information on nonrespondents been included on the data file?

Eight out of the 13 NCES surveys examined include information on nonrespondents. For five out of those eight (SASS, TFS, BPS, HS&B, and B&B) however, this information was contained only on the restricted-use data file. Only two surveys (RCG and NELN:88) contain information on nonrespondents on the public-use data file.

Analysis of Response Categories

Researchers using more than one NCES database soon discover that there is minimal uniformity in demographic data collected: either the question wording or the response categories differ. We have identified nine common demographic survey items, five representing institutional characteristics and four representing individual characteristics.

Institutional variables

A. Sector

Twelve of the 13 NCES surveys collected data about the school sector (public, private, etc.). Unlike other variables described in this chapter, school sector was often not directly asked to respondents, but was a sampling frame variable. School sector was asked on five of the NCES surveys examined (NELN:88, RCG, HS&B, B&B, and NHES).

B. Region

Four of the 13 NCES surveys examined did not provide a region designator (NSOPF, RCG, BPS, and B&B). The remaining nine surveys may be divided into five categorization schemes: Four surveys--SASS, TFS, NALS, NHES--used the FIPS (Federal Information Processing Standards) categorization: Northeast, Midwest, South, and West. NAEP, NELN:88, and NLS-72 also provided four categories, but used slightly different categories. One of the two region categories provided on HS&B also provides four categories. NAEP used Northeast, Southeast, Central, and West, however, the part of Virginia that is included in the Washington, DC metropolitan statistical area (MSA) is included in the Northeast region, while the remainder of the state is included in the Southeast region. NELN:88, NLS-72, and NHES use Northeast, North Central, South, and West.

C. Urbanicity/locale

Six surveys provided documentation on urbanicity/locale (NLS-72, NELN:88, NAEP, TFS, HS&B, and NHES). Three of these surveys (NLS-72, TFS, and HS&B) presented very similar categories: a rural or farming community,

not a suburb of a larger city, a medium-sized city (50,000 to 100,000 people), a suburb of a medium-sized city, a large city (100,000 to 500,000 people), a suburb of a large city, a very large city (over 500,000 people), and a suburb of a very large city. TFS also included Indian reservation and military base, while HS&B only included military base. NELS:88 included only three categories: urban, suburban, and rural, developed from a composite variable created directly from QED (Quality Education Data), using the FIPS designator, utilized by the U.S. Census. NAEP collapses an urbanicity/locale variable into three categories: urban, suburban and rural. It also provides more detailed categories based on 1980 Census information. These categories included: rural, disadvantaged urban, advantaged urban, big city, fringe, medium city, and small place. Since 1990, SASS has replaced the self-reported community type with a 7-category scheme determined by the ZIP Code of the school and matched to the Census community size for that ZIP Code (Johnson, 1989).

D. School level

School level identifies whether the school is primary, secondary, or a combination of the two. (This analysis is not applicable to postsecondary schools.) Only three of the 13 NCES surveys examined (NELS:88, NHES, and SASS) provided such designation and all use different categories. NELS:88 does not provide school level exactly, but classifies the type of school by the grades spanned, which were collapsed into seven categories, using school data first. NHES classifies by lowest grade (prekindergarten to 11th) and highest grade (3rd to 12th). SASS provided four choices: elementary (if the school has only grades below 8th grade), middle school/junior high, secondary (if the school has grades between 7th and 12th, and combined elementary and secondary (if the school has any other combination of grades).

E. School/Institution size

Six surveys provided information on school/institution size: no two surveys used the same categories. NLS-72 indicated school size by enrollment of seniors--less than 400 or greater than 400. NELS:88 provided a composite variable, categorizing the entire school enrollment as reported by the school. These values were 1-199, 200-399, 400-599, 600-799, 800-999, 1000-1199, and 1200+. On the public school questionnaire, SASS asked for the total number of students enrolled in grades K-12 or comparable ungraded levels. RCG had three categories: less than 1,500, 1,500 to 5,999, and 6,000 or more. NPSAS set its categories at less than 1,000, 1,000 - 2,499, 2,500 -

4,999, 5,000 - 9,999, 10,000 - 19,999, and 20,000 or more. NHES defined school size as under 300, 300 - 599, 600 - 999, and 1,000 or more.

Individual characteristics

F. Race/ethnicity

All 13 NCES surveys inquired about respondents' race; however, differences were found in categories from one survey to another. The first difference is the order of the race response categories. Some surveys begin with a minority response category such as black, American Indian, Asian, etc. (NLS-72, NSOPF, NELS:88, and HS&B), while others begin the response categories with white. Six surveys (NLS-72, NAEP, RCG, BPS, HS&B, and NHES) provide an other race category, while the remaining surveys do not. Race categories also varied by whether a Hispanic item was provided. RCG and NAEP combine race and Hispanic origin, e.g., white, non-Hispanic. Seven surveys ask for race information, followed by asking about Hispanic origin. On SASS, TDS (Teacher Demand and Shortage) and the School Survey include Hispanic origin as part of the race item, while the Administrator Survey and Teacher Survey ask this item separately. HS&B provides Hispanic as a type of race, not distinguishing white, black or other race. Only NLS-72 does not include a Hispanic designator.

G. Socioeconomic status

Surveys rarely ask respondents to provide their socioeconomic status (SES). Instead, this variable was constructed by combining various sociological and economic data. Only two surveys (NLS-72 and NELS:88) provided a specific SES composite variable on the data file. For NLS-72, SES was derived from an equally weighted linear composite of father's education, mother's education, father's occupation, family income, and household items (such as newspaper, dictionary, encyclopedia, etc.) from the first follow-up and/or base year student questionnaire. NELS:88 used the same composite variables; however, mother's occupation was used, rather than household items. The remaining surveys do not contain an SES composite.

H. Degree

All NCES surveys examined inquired about respondents' level of education/degree. However, a variety of different questions and response categories were used to gather them. In general, we may group survey responses into three major categories: responses with detailed lower degree levels, responses with detailed higher degree levels, and those with broad categories. Surveys with detailed lower degree levels include NPSAS, BPS, NALS, and NHES. Surveys with

include NPSAS, BPS, NALS, and NHES. Surveys with detailed higher response categories include NLS-72 and B&B. NSOPF used seven detailed categories utilizing not only the names of various degrees, but mentioning words such as equivalent or certificate as completing one's degree. TFS used six categories: associate's degree, bachelor's degree, master's degree, doctorate, education specialist or professional diploma, or professional degree. Three surveys used broad categories. NELS:88 offered three: less than a bachelor's, bachelor's, and master's. RCG provided three categories: bachelor's, master's, or some other degree. NAEP provided four categories (among parents education): did not finish high school, graduated high school, some college, graduated college, or don't know. SASS asked about degree types earned on two of its surveys, the Teacher Survey and the Administrator Survey.

I. Respondents' Age Group

ALL NCES surveys inquired about age, but few provided age groupings. Only NAEP provided for age groupings for children, specifically, students who were either in the fourth grade or 9 years old; students who were either in the eighth grade or 13 years old; and students who were either in the twelfth grade or 17 years old. On the TFS survey, the restricted use file provided actual ages; however, the public release file provided four categories: Under 30, 30 to 39, 40-49, and 50 and above. On RCG, actual ages are provided for respondents and categories are provided for newly qualified teacher of: 23 or younger, 24 to 25, and 26 or older. All other surveys inquired about respondents' exact year of birth or actual ages so that researchers may combine specific ages and convert them to age groupings.

Conclusions

This paper examined two major topics: consistency of response rates information/calculation and consistency of response categories. Most NCES surveys provided detailed information on unit and item response rates and defined these consistently across surveys. The amount of documentation on the intensive follow-up of nonrespondents was minimal. Some of the response categories showed large variation across surveys, such as those used for urbanicity and race/ethnicity. Different questionnaire wording (some of which also had different response categories), were also prevalent especially for those used for degree and race/ethnicity.

Recommendations and suggestions

Several additional studies could be explored to further

elaborate on information provided in this report: 1) more efforts are needed to examine the impact of response rates on baseline statistics related to two major issues: what bias is generated by differential nonresponse rates on estimates of school resources and student outcomes across geographic or socioeconomic categories? How much bias can be measured or adjusted, if differential response rates are found? 2) the most recent surveys could be considered for the nonresponse issues since response rates change over time due to different reasons. Techniques for calculating response rates may change over time, too. Higher nonresponse rates might be due to the mode of administration or economic status of respondents. These issues could be addressed in the further studies. 3) Additional response categories may be examined, such as Likert scales (3- 5- or 7-point, response categories from low to high or vice versa, etc.).

References

Abraham, S., Suter, N., Spencer, B., Johnson, R., Zahs, D., Myers, S. and Zimble, L. (1994), 1992-92 National Study of Postsecondary Faculty Field Test Report. Technical Report NCES 93-390. U.S. Department of Education, Office of Educational Research and Improvement. Washington, DC: National Center for Education Statistics.

Council of American Survey Organizations (CASRO) (1982), "On the Definitions of Response Rates." Port Jefferson, New York.

Ingels, S., Dowd, K., Baldrige, J., Stipe, J., Bartot, V. and Frankel, M. (1994), "Second Follow-up: Student Component Data File User's Manual, Volume I." NCES 94-374. U.S. Department of Education, Office of Educational Research and Improvement. Washington, DC: National Center for Education Statistics.

Johnson, F., (1989), "Assigning Type of Locale Codes to the 1987-1988 Public School Universe," Technical Report: Assigning Type of Locale.

Kalton G. (1983). Compensating for Missing Survey Data, Survey Research Center, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan.

Westat, Inc. (1994), "1991 Survey of Recent College Graduates: Methodology Report." Washington, DC: National Center for Education Statistics.