DOCUMENT RESUME

ED 350 325 TM 019 078

AUTHOR Mohadjer, Leyla; And Others
TITLE Proxy Reporting of Dropout Status in the NHES Field Test. National Household Education Survey Technical Report.
INSTITUTION Westat, Inc., Rockville, MD.
SPONS AGENCY National Center for Education Statistics (ED), Washington, DC.
REPORT NO NCES-92-103; NHES-TR-4
PUB DATE Jul 92
NOTE 35p.; For related documents, see TM 019 075-079.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adolescents; Adult Education; Data Collection; Dropout Research; *Dropouts; *Enrollment; Error of Measurement; *Estimation (Mathematics); *Field Tests; *National Surveys; Reliability; Research Methodology; Sampling; Statistical Bias; Young Adults
IDENTIFIERS *National Household Education Survey; *Proxy Reports

ABSTRACT

The National Household Education Survey (NHES) was conducted for the first time in 1991 as a way to collect data on the early childhood education experiences of young children and participation in adult education. Because the NHES methodology is relatively new, field tests were necessary. A large field test of approximately 15,000 households was conducted during the fall of 1989 to examine several methodological issues. This report focuses on measurement errors arising from the use of proxy respondents. During the 1989 Field Test, a knowledgeable household member was used as a source of information on the school enrollment of each sampled 14- to 21-year-old in the household. In addition, this age group was asked to report on its own school enrollment. The correspondence between the responses given by proxy respondents and the youths themselves is reviewed. Field test findings indicate that it will be possible to estimate the size of measurement error associated with using proxy respondents, but that it may be prudent to collect data from a household member and from the youths themselves to improve the reliability of dropout estimates. Fourteen tables and two figures illustrate the discussion, and an appendix contains statistics on the measurement errors for variables used to derive event and state dropout rates. (SLD)

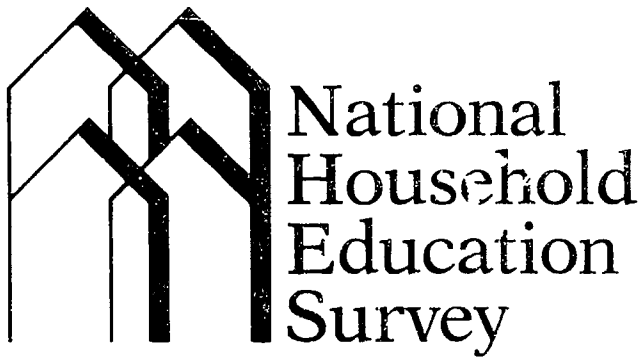
NATIONAL CENTER FOR EDUCATION STATISTICS

Technical Report July 1992

National Household Education Survey

Technical Report No. 4

Proxy Reporting of Dropout Status in the NHES Field Test



Leyla Mohadjer
J. Michael Brick
Westat, Inc.

Jerry West
National Center for Education Statistics

U.S. Department of Education
Office of Educational Research and Improvement NCES 92-103

U.S. Department of Education
Lamar Alexander
Secretary

Office of Educational Research and Improvement
Diane Ravitch
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Acting Commissioner

National Center for Education Statistics

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

July 1992

Contact:
Jerry West
(202) 219–1574

Foreword

The National Household Education Survey (NHES) represents a major new initiative of the National Center for Education Statistics (NCES). Between February and May of 1991, the NHES was fielded for the first time as a mechanism for collecting data on two different sectors of education policy interest: the early childhood education experience of young children and participation in adult education. Because the NHES methodology is relatively new and relies on some innovative approaches, a field test of the methodology was an essential first step in the development of the survey. Many of the methods of evaluated during the 1989 NHES field test were adopted for the full-scale survey.

A large field test of approximately 15,000 households was conducted during the fall of 1989. A number of methodological issues associated with collecting and analyzing data on education issues from a random digit dialing telephone survey were examined. This report is one of five that describe the 1989 NHES Field Test experience. The five reports are the first in a series of technical publications pertaining to the design and conduct of the NHES that NCES hopes to continue in the years to come. NCES believes that the reports contained in this series will provide users of the NHES data with a better understanding of the NHES methodology and that they will assist the survey design efforts of others.

The first report in this series, *Overview of the National Household Education Survey Field Test*, describes the design of the field test and the outcomes of the field test data collection activities. It reports on the response rates obtained, both unit and item, and the burden associated with survey participation. Each of the next four reports in the series focuses on a specific issue that was examined in the 1989 NHES field test.

The second report, *Telephone Undercoverage Bias of 14- to 21-Year-Olds and 3- to 5-Year-Olds*, analyzes data from the Current Population Survey to identify the extent of telephone coverage for two distinct populations of interest and the bias associated with this type of undercoverage for estimates of school dropouts and early childhood education program participation. Methods for adjusting survey estimates to partially reduce this bias are developed and evaluated.

The third report, *Multiplicity Sampling for Dropouts in the NHES Field Test*, examines a technique that was used to increase the coverage of 14- to 21-year-olds and to capture more dropouts in the sample. The report describes the effectiveness of the multiplicity sample in achieving these goals.

The fourth report, *Proxy Reporting of Dropout Status in the NHES Field Test*, focuses on measurement errors arising from the use of proxy respondents. During the 1989 Field Test, a knowledgeable household member was used as a source of information on the school enrollment of each sampled 14- to 21-year-old in the household. In addition, 14- to 21-year-olds were asked to report on their own school enrollment. The report describes the correspondence between the responses given by proxy respondents with those provided by the youths themselves.

The fifth report, *Effectiveness of Oversampling Blacks and Hispanics in the NHES Field Test*, describes the approach used to increase the number of black and Hispanic households/youth in the sample. During the field test, an approach that uses demographic information at the telephone exchange level to develop sampling strata was used to oversample black and Hispanic households. The report examines the yield of the field test sample design versus that which would have been expected without oversampling. The effects of oversampling on the precision of survey estimates are reported.

Paul Planchon
Associate Commissioner
Elementary and Secondary Statistics Division

Jeffrey A. Owings
Chief
Longitudinal and Household Studies Branch

Acknowledgments

Many individuals made significant contributions to the 1989 National Household Education Survey (NHES) Field Test and to the preparation of this report. The authors gratefully acknowledge their efforts. The 1989 Field Test was conducted by Westat Incorporated under contract with the U.S. Department of Education, National Center for Education Statistics (NCES). Lance N. Hodes was the corporate officer at Westat with oversight responsibility for the NHES Field Test. Mary Collins served as the associate project director for the NHES Field Test and was a major contributor to all phases of the study. In her role as survey manager, Carin A. Celebuski led the development of the Field Test dropout component questionnaires. Jacque Wernimont and David R. Martin, both Westat senior systems analysts, played major roles during the design and implementation of the survey and creating the database used in the analyses found in this report.

Critical technical review of this report was provided by NCES staff Michael Cohen, Bob Burton, Marilyn McMillen, and Jeffrey A. Owings, Branch Chief, Longitudinal and Household Studies Branch. Seymour Sudman, Survey Research Laboratory, University of Illinois also reviewed the report. The authors wish to thank each of these individuals for their careful reading of this report and for their comments and suggestions.

Table of Contents

List of Tables

List of Figures

Introduction

During the fall of 1989, the Field Test of the National Household Education Survey (NHES) was conducted by the National Center for Education Statistics (NCES) to explore the feasibility of collecting education data by telephone from a sample of persons in their households. The NHES is the first major attempt by NCES to go beyond its traditional surveys, which rely upon school-based data collection systems and are typically conducted by mail or in-person data collection methods.

A household survey has the potential to provide the types of data needed to study current issues in education, particularly those which can not be easily addressed through a school-based survey. Such issues include dropping out of school, adult and continuing education, preschool education, the status of former teachers, and home-based education. Consequently, the NHES methodology may greatly enhance the scope of issues covered by the data collection activities of NCES.

Since the NHES data collection methods were untested for education surveys, the Field Test was developed to evaluate the use of this approach. Two topics of broad policy interest were included in the Field Test: the early childhood education characteristics of 3- to 5-year-olds, and the educational status of 14- to 21-year-olds with a special focus on youth who dropped out of school before completing high school. By including both of these study areas in the Field Test, the ability to use the NHES to study multiple, complex topics, employing different sampling requirements and respondent rules could be evaluated.

Westat, Inc., under contract with NCES, conducted all of the Field Test interviews using computer-assisted telephone interviewing (CATI) methods. The use of CATI methods made sampling respondents for interviews easy and nearly invisible to the telephone respondent, an important benefit when several persons may be sampled in a household. CATI also directed the interviewers through complex skip patterns and provided the opportunity to incorporate edit checks to help resolve inconsistencies in the data while the respondents were still on the telephone. Another major advantage of the use of CATI was that data analysis could begin soon after data collection ended, because data entry and many of the edit checks were done during the interview.

The sampling scheme used in the Field Test was a variant of the Mitofsky-Waksberg random digit dial (RDD) procedure[1] in which every residential telephone number has the same chance of being drawn into the sample. Because of the need for more precise estimates of blacks and Hispanics, special sampling methods were used to increase the sample size for these persons. The design for the Field Test was essentially the same as planned for a full-scale NHES study, except the overall sample size was smaller.

The sample resulted in collecting data from 15,037 households representing all civilian, noninstitutionalized persons in the 50 states and the District of Columbia. Although only persons living in telephone households could be sampled for the Field Test, adjustments were made in the weights so that the estimates of persons living in both telephone and nontelephone households could be produced.

Respondents in sampled households were asked a series of screening questions. This interview, called the Screener, was used to enumerate all the members of the household, determine the eligibility of each person in the household for the early childhood education (3- to 5-year-olds) and youth (14- to 21-year-olds) studies, and obtain some data on the characteristics of the household. A total of 4,374 households had at least one person enumerated in the Screener who was eligible for an extended interview. The response rate to the Screener was 79 percent.

The early childhood education interview was conducted with the parent or guardian who knew the most about each sampled 3- to 5-year-old child's care and education. Accordingly, this interview was called the Parent Interview. Of the 1,551 children identified in the Screener, parents completed interviews for 1,530 children, a completion rate of 99 percent.

If the household contained any 14- to 21-year-olds, then a Household Respondent Interview (HRI) was attempted for each of these members. The HRI was used to determine the current and previous educational status of the youth; this interview could be completed by any adult household member who knew about the educational activities of the youth, including self-reports by the youth. Of the 4,441 youths identified in the Screener, HRIs were completed for 4,313 youths, for a 97 percent completion rate. As part of a special methodological study of multiplicity sampling, mothers in a subsample of the households were asked to complete

the HRI for their 14- to 21-year-old children who did not live in their household. These youths are included in the numbers stated above.

A Youth Interview (YI) was then attempted for a subsample of the 14- to 21-year-olds in the household. All the youths who were not currently enrolled in school and did not have a high school diploma or equivalent (as reported in the HRI), and a sample of all other youths, were targeted for the YI. The interview contained more detailed items on the educational experiences of the youth that could only be answered by the youth. Of the 1,863 youths sampled, 1,604 completed the YI, a completion rate of 86 percent. These numbers include a sample of 133 youths (with 67 completed YI) who did not live in the sampled households, but were included through the multiplicity sample when their mothers completed the HRI.

This report describes the correspondence between the responses of knowledgeable household respondents and the youths themselves, one of several methodological studies undertaken in the Field Test. The Field Test is described in greater detail in another report entitled *Overview Report on the 1989 National Household Education Survey Field Test*, the first in a series of reports on the Field Test. The Overview Report describes the sample design, the data collection methods and instruments, the response rates, and other salient aspects of the collection and analysis process for the Field Test.

The subject of this report is the measurement errors arising from the use of proxy respondents in surveys. In many surveys, such as the Current Population Survey (CPS), one person typically responds to questions for all members of the household. The responses of the household respondent are then used to classify the individuals by characteristics (e.g., whether they are enrolled in school). The responses of the youths themselves may be different from the household respondent's responses for a number of reasons. The purpose of this analysis is to determine the impact on the estimates of dropout rates that might arise if proxy respondents are permitted in the NHES. Since the youths could have also served as the household respondent (if they were at least 18 years old), the analysis investigates the impact of this reporting structure.

Most of the previous research in this area covers proxy respondents used for health-related statistics, major or minor events of life (e.g., car accidents), and statistics on expenditure data. The effects of proxy respondents on measurement errors for estimates of education characteristics, and dropout estimates in particular, have not been systematically examined.

Sample Design and Data Sources

One aspect of the sample design for the Field Test that affected the precision of the computed measurement error statistics was the unequal probabilities of selection used to sample youths for the extended interviews. Those youths who were identified as potential dropouts by the household respondent were sampled with certainty for the extended interview. Those youths identified as not-dropouts were subsampled at a rate of 1/4.7 for the extended interview. Some of the latter youths reported themselves as being dropouts. The unequal probabilities of selection for different youths in the sample distort the sample distribution of dropouts and not-dropouts as reported by household respondents or youths themselves.

As a result of this and other features of the sample design, the sample distribution for the YI did not resemble the distribution of the population and, using the observed sample sizes to compute measures of reporting reliability, provides biased estimates. To account for the differential probabilities and the fact that the sample distribution was not the same as the population distribution (i.e., the sample was not self-weighting), weighted data were used to compute the reliability of reporting measures. The weighted data adjust for the unequal probabilities of selection.

As noted earlier, up to two interviews were conducted for each 14- to 21-year-old after the initial household screening was completed. The HRI collected detailed information on the educational status of each 14- to 21-year-old. The HRI was used as the source of information to

1) produce national estimates of dropout rates for 14- to 21-year-olds using different dropout definitions,
2) identify dropouts for extended interviews, and
3) sample non-dropouts for extended interviews.

A set of identical items was asked in the HRI and the YI. These items were repeated because it was important to understand the relationship of the responses of the household respondent to those of the youth. If the household respondent and the youth reported essentially identical data for the most important characteristics, such as school enrollment and attendance, then it would be possible to estimate dropout rates without contacting the youth in future surveys. The reliability of the household respondents also has implications for studies that screen at the household level in order to identify dropouts for more intense analysis. In these respects, the reliability of the HRI responses is a critical feature for future surveys.

In addition to educational items which paralleled those in the HRI, the Youth Interview collected information on the educational history of the youths, including suspensions, grades, retentions in grades, etc., and other potential correlates of dropping out of school, including events such as marriage and births.
The analyses summarized in this report are based only on persons for whom there was a complete HRI and a complete YI and excludes data from the "out-of-household" youths. The total sample size for this group is 1,537. The out-of-household youths were excluded because of the small numbers of completed interviews for this group (67) and because to include them would deflect attention from the central analytic issue. Another important reason for excluding them, is the fact that they have very different characteristics as compared to the youths who live at home.

The out-of-household youths are, of themselves, an interesting and policy-relevant group. Preliminary analysis showed that the correspondence between the youth's and HRI's responses were smaller for this group. However, due to the small sample size, this evidence is very limited.

Methods Used for Comparison

In this section several statistics that examine various aspects of reporting in the HRI and YI are introduced and applied to estimates of event dropouts and status dropouts. A status dropout is defined as a 14- to 21-year-old who was not enrolled in school in October of the current year and did not have a high school diploma or equivalent. Event dropouts are defined as the subset of status dropouts who were enrolled in school in October of the previous year. In other words, a status dropout is someone who is not currently enrolled and does not have a diploma or equivalent, and an event dropout is a dropout who left school within the last year.

The statistics we have computed for this study are the set of statistics developed for assessing response reliability based upon re-interview data. The statistics include the gross difference rate, the net difference rate, the percentage of cases with identical responses, and the index of inconsistency. The gross difference rate measures the proportion of cases that had different responses in the two surveys—the HRI and the YI. The net difference rate measures the bias after the offsetting misclassifications have been taken into account. The index of inconsistency is a less familiar statistic. For qualitative data, the index of inconsistency is a measure of the impact of misclassification error on the estimates.[2]

These statistics have typically been applied to the situation in which two interviews are conducted with the same respondent. In this situation the statistics measure the consistency of reporting by the respondent. In the Field Test, the statistics are based on different respondents. Therefore, as used here, the statistics are estimating the consistency of reporting over time and the differences between respondents in the same measures.

Traditionally, these statistics are computed for samples that are almost self-weighting and are, therefore, based on the number of sample cases reported as having the characteristic in the original survey and in the re-interview. However, as mentioned previously, the NHES Field Test design did not produce a self-weighting sample of youths. To account for the differential probabilities of selection, the statistics for reporting reliability are computed using the weighted data.

The following table shows the general format of the possible reporting outcomes by youths and household respondents.

From tables formatted in this fashion it is possible to estimate several characteristics relevant to the consistency of the reporting between household respondents and youth themselves. For example, the off diagonal cells estimate the number of persons whose dropout status was reported differently in the HRI and the YI.

The following provides the definitions of the statistics computed in this report. Note that the youth's

Table 1.—Illustrative reporting outcomes by youths and household respondents

Household Respondent Interview	Youth Interview		
	Estimated number of dropouts	Estimated number of not dropouts	Total
Estimated number of dropouts	A'	C'	A' + C'
Estimated number of not dropouts	B'	D'	B' + D'
Total	A'+B'	C'+D'	N'=A'+B'+C'+D'

responses are taken as the truth or "standard," and HRI responses are compared with the standard. Since the choice of the YI as the standard is somewhat arbitrary, the implications of this choice on each statistic are briefly discussed below and in greater detail later in this section.

Gross Difference Rate

The gross difference rate is equal to the estimated percentage of cases reported as dropouts by youths but reported as not dropouts by household respondents, plus the estimated percent of cases reported as not dropouts by youths but as dropouts by household respondents. That is, the gross difference is the estimated total number of cases misclassified as dropouts or not dropouts by household respondents. The gross difference rate is the ratio of the gross difference divided by the estimated total number of youths, or

$$G = \frac{B' + C'}{N'} \times 100$$

Note that the value of G would be the same as above if the HRI was used as the standard rather than the YI.

Net Difference Rate

The net difference for dropout statistics is the difference between the estimated total number of dropouts as reported by youths and the estimated total number of dropouts as reported by household respondents. That is, (A' + B') - (A' + C') = B' - C'. Note that the gross difference includes differences in both directions, partly or substantially offsetting. The net difference is the non-offsetting part of the gross difference. The net difference rate is the ratio of the net difference to the estimated total number of youths, or

$$E = \frac{B' - C'}{N'} x100$$

Note that the choice of which interview is used as the standard only affects the sign of E.

Percent Identically Reported

The percentage identically reported as dropouts is equal to the ratio of the estimated number of youths reported as dropouts by both household respondents and youths over the total number of dropouts reported by youths,

$$R = \frac{A'}{A' + B'} \times 100$$

Since the assumption is that the youth responses are the standard, this index provides a measure of the stability of household respondent responses relative to the standard. Since R uses A'+B' as the base, the standard chosen affects the value of this statistic. However, the change in the value of R is minimal

(about 3 to 4 percent) if the HRI is taken to be the standard rather than the YI.

Index of Inconsistency

The index of inconsistency is equal to

$$I = \frac{G}{2PX} \times 100 = \frac{B' + C'}{2N'PQ} \times 100$$

where

$$P = \frac{A' + C'}{N'}, \text{ and}$$

$$Q = 1 - P.$$

In this formula, G/2 can be viewed as a measure of the probability of being misclassified for two interviews, under certain conditions[3], and PQ as the total random variance, including both response and sampling error. Therefore, I is measure of the impact of the misclassification error on the estimate. It is worth noting that the denominator for I could be unbiasedly estimated from other statistics .

For categorical data, the index of inconsistency measures the impact of misclassification errors on the total variance of an observation. I is not a direct measure of misclassification. More details on the description of the index of inconsistency for qualitative and quantitative variables are given in another Census Bureau publication[4].

In the findings of this report, the above methodology is applied to the sample selected for the NHES Field Test. Estimates of the number of dropouts based upon reports of the household respondents are compared to the estimates based upon the reports of youths. The relationship between the direction of errors for event and status dropouts provides further insight into the nature of the error. To further analyze the effects of possible correlates, comparisons are also given by age and race/ethnicity of the youths.

Related Issues

There are several factors that should be recognized when drawing conclusions based on the results of this study. Youths had a chance to be their own household respondents. Out of the total 1,537 extended youth interviews, 477 youths, or about 30 percent were also the household respondents. Consequently, the reliability of responses coming from household respondents is higher than in a design in which youths do not have any chance of being the household respondents. This approach is appropriate when the goal is to examine the expected reliability of responses from a sample design in which the youths can report for themselves.

Another objective of the analysis might be to assess how well proxy respondents report for youth. This goal can be met by excluding the youth who reported for themselves. Table 7 of the findings section shows the increase in the estimated measurement errors when the 477 self-responding youths are excluded from the analysis.

The next factor that needs to be considered is the amount of survey resources required to obtain responses from youths themselves. The increased cost of collecting data from the youths themselves should be compared with the possible shifts in the sample estimates based on data reported by household respondents. A related issue is the biases arising from higher nonresponse associated with hard-to-collect samples. The completion rate associated with the YI in the NHES Field Test was about 85 percent. To the extent that youth nonresponse is related to dropout status, the results of the YI may also include some unknown shift from the true values.

A third factor to be considered is the relation between the social desirability of the item in question (dropping out of school) and the respondents. As noted earlier, the assumption we have made for this study is that the youths' responses are the standard or the truth when compared to the responses of the household respondents. Since dropping out of school is not a desirable action, it is not clear which respondent will supply more reliable data. Some researchers[5] argue that items that are considered socially undesirable should be asked of "proxy" respondents rather than "self" respondents. However, "dropping out of school" may be more difficult for parents to admit than for the youths themselves. For this reason, the YI is considered to be the standard in this analyses even though in some cases the HRI will undoubtedly be more accurate. Clearly, this is a complicated issue that requires further investigation.

Findings

This section describes the results of applying the methodology outlined above to the data collected in the Field Test. A series of tables is used to evaluate the measurement errors arising from using household respondents to collect information about event and status dropouts. It should be noted that the sample sizes of dropouts in some of the categories in these tables were small, especially for event dropouts. The statistics derived for categories with small sample sizes should, therefore, be looked upon as rough estimates.

Tables 2 and 3 show the number of event and status dropouts estimated from data reported in the HRI and the Youth Interview. Note that the analysis in this report excludes the "out-of-household" youths; thus the estimated total number of youths provided in tables 2 and 3 underestimates the actual numbers in the population (this applies also to the remainder of tables showing estimated population numbers). There may be some differential effects due to the elimination of the out-of-household youths since they have different characteristics than the in-household youths. These differences are one of the primary reason for eliminating the out-of-household youths from this analysis.

Tables 2 and 3 are simply the cross classification of dropouts and not dropouts as reported by HRI and Youth Interviews. The estimates of status dropouts from the HRI are smaller than those from the YI, but this relationship is reversed for event dropouts. An examination of the tabled distributions of the estimates and the unweighted sample counts reveals how important it is to use weighted estimates in computing measures of reliability from these data. For example, the average weight for the HRI-reported event dropouts in table 2 is only half of the weight for those classified as not event dropouts.

Table 4 shows the estimated measures of reporting reliability for event and status dropouts. The percentage identically reported for status dropouts is equal to 86 percent, which is about 12 percent more than for the event dropouts. The gross difference rate and the net difference rate are about the same for both event and status dropouts. The index of inconsistency for event dropouts is about twice the one for status dropouts.

As stated earlier, out of the total of 1,537 extended (in-household) youth interviews, about 30 percent (477 youths) were also the household respondents. Tables 5 and 6 show the number of event and status dropouts estimated from the HRI and the Youth interviews when the 477 self-reporting youths are excluded from the sample. Table 7 provides estimated measures of reporting reliability associated with self-reporting youths excluded from the sample.

Figure 1 shows that the percentage identically reported is improved by about 12 percent for event dropouts and by about 8 percent for status dropouts when self-responding youths are included in the sample. The index of inconsistency and the gross difference rate are also improved, but the net difference rate does not change with the inclusion of the self-responding youths in the sample.

A comparison of proxy and youth reporting of dropout status by the relationship of the proxy respondent to the youth is shown in table 8. About 75 percent of the parent proxy respondents were mothers of the youths. As shown in the table, the sample sizes associated with different types of proxy respondents were too small to support any conclusions on sizes of the measurement errors by various types of proxy respondents.

In order to appreciate how the findings from this study compare to the results from other surveys, some statistics from Census Bureau surveys are provided. Since the usual applications of these measures are re-interview programs, there are limitations of the comparisons as noted in the previous section. The major difference is that all of the comparisons cited below are based on re-interviews with the same respondent while the respondents typically are different in the NHES.

The Census Bureau reported[6] on the percent of persons who reported identical ethnic origins in interviews spaced 1 year apart in the Current Population Survey. The percentage reporting identically was 65 percent overall. The percentage reporting identically varied from 32 percent to 96 percent across ethnic origins. The ethnic origins used in this study were relatively detailed.

Brooks and Bailar[7] presented estimates of the net difference rate and index of inconsistency for employment statistics. The net difference rate for those working full time in nonagricultural varied from -.36 to -.99. The index of inconsistency for this estimate varied from 13.2 to 17.1.

Table 2.--Estimates of event dropouts as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not-event dropouts	Event dropouts
Survey Estimates (in thousands)			
Total	26,335	25,743	592
Not-event dropouts	25,724	25,572	153
Event dropouts	610	171	439
Sample Sizes			
Total	1,537	1,480	57
Not-event dropouts	1,469	1,461	8
Event dropouts	68	19	49

Source: 1989 National Household Education Survey Field Test

Table 3.—Estimates of status dropouts as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not-status dropouts	Status dropouts
Survey Estimates (in thousands)			
Total	26,335	24,171	2,163
Not-status dropouts	24,271	23,975	296
Status dropouts	2,063	196	1,867
Sample Sizes			
Total	1,537	1,305	232
Not-status dropouts	1,297	1,297	18
Status dropouts	240	26	214

Source: 1989 National Household Education Survey Field Test

Table 4.—Measures of reporting reliability for event and status dropout statistics

	Percent[1] of 14- to 21-year-olds	Gross difference rate	Net difference rate	Percent identically reported	Index of inconsistency
Event dropouts					
Estimate	2.4%	1.2%	-0.1%	74.2%	27.2%
Standard error	0.3%	0.4%	0.4%	9.5%	7.2%
Status dropouts					
Estimate	8.3%	1.9%	0.4%	86.3%	13.0%
Standard error	0.6%	0.5%	0.4%	3.7%	2.7%

[1]Based on Youth Interview reports.

Source: 1989 National Household Education Survey Field Test

Table 5.—Estimates of event dropouts as reported by household respondents and youths excluding youths who were their own household respondents

Household respondent interview	Youth interviews		
	Total	Not-event dropouts	Event dropouts
Survey Estimates (in thousands)			
Total	20,970	20,568	402
Not-event dropouts	20,549	20,397	153
Event dropouts	421	171	249
Sample Sizes			
Total	1,060	1,023	37
Not-event dropouts	1,012	1,004	8
Event dropouts	48	19	29

Source: 1989 National Household Education Survey Field Test

Table 6.--Estimates of status dropouts as reported by household respondents and youths excluding youths who were their own household respondents

Household respondent interview	Youth interviews		
	Total	Not-status dropouts	Status dropouts
Survey Estimates (in thousands)			
Total	20,970	19,622	1,348
Not-status dropouts	19,722	19,426	296
Status dropouts	1,248	196	1,052
Sample Sizes			
Total	1,060	921	139
Not-status dropouts	913	895	18
Status dropouts	147	26	121

Source: 1989 National Household Education Survey Field Test

Table 7.—Measures of reporting reliability for event and status dropout statistics excluding youths who were their own household respondents

	Percent[1] of 14- to 21-year-olds	Gross difference rate	Net difference rate	Percent identically reported	Index of inconsistency
Event dropouts					
Estimate	2.4%	1.6%	-0.1%	62.0%	39.3%
Standard error	0.3%	0.4%	0.5%	12.2%	11.2%
Status dropouts					
Estimate	8.3%	2.4%	0.5%	78.0%	21.0%
Standard error	0.7%	0.6%	0.5%	5.7%	4.6%

[1]Based on Youth Interview reports.

Source: 1989 National Household Education Survey Field Test

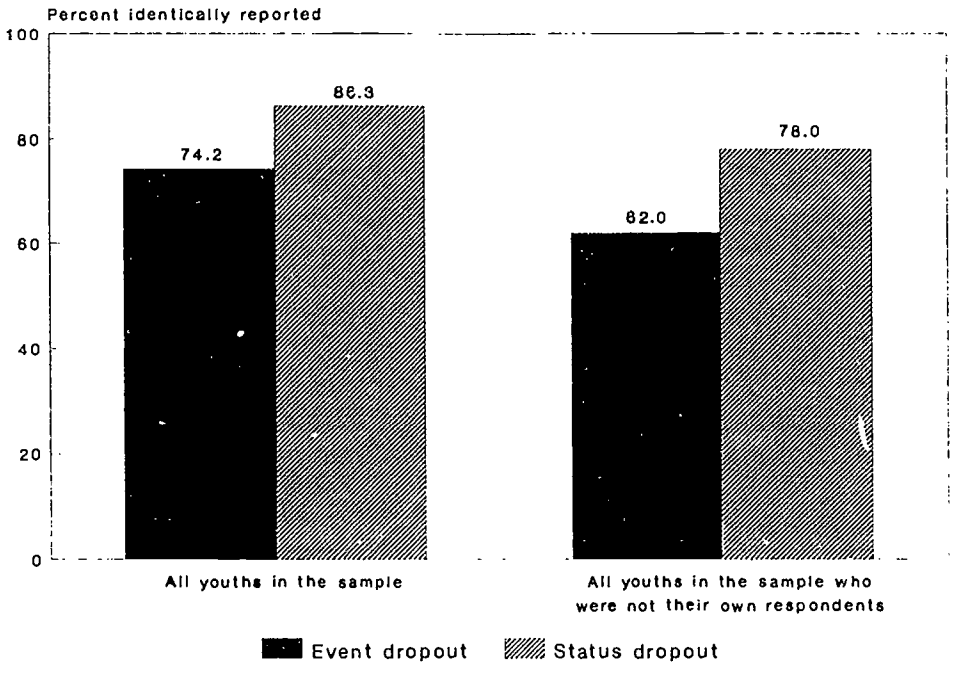
Table 8.—Comparison of proxy and youth reporting of dropout status by the relationship of the proxy respondent to the youth

Household respondent interview	Youth interviews		
	Not status dropout (in thousands)	Status dropout (in thousands)	Percent identical
Survey Estimates			
All (including self)			
Not status dropouts	23,975	296	86.3
Status dropouts	196	1,867	
All proxy respondents[1]			
Not status dropouts	19,287	296	78.0
Status dropouts	196	1,052	
Parent proxy respondents			
Not status dropout	16,039	248	72.4
Status dropout	148	649	
Other proxy respondents[2]			
Not status dropout	3,248	48	89.4
Status dropout	48	403	
Sample Sizes			
All			
Not status dropouts	1,279	18	
Status dropouts	26	214	
All proxy respondents			
Not status dropouts	895	18	
Status dropouts	26	121	
Parent proxy respondents			
Not status dropouts	666	11	
Status dropouts	21	75	
Other proxy respondents[2]			
Not status dropouts	219	7	
Status dropouts	5	46	

[1]Information on relationship of the proxy respondent to the youth was missing for 10 cases.

[2]Includes grandparent, other relative, and nonrelative.

Figure 1. — Percent identically reported for event and status dropout statistics



Source: 1989 National Household Education Survey Field Test

A wide variety of statistics are reported on the 1970 Census data collection[8]. In the Census report, statistics are classified into three groups depending on the index of inconsistency. If the estimated index is less than 20, the statistics are considered to have low levels of response variance. Those with indices between 20 and 45 are considered to have moderate response variances. Any statistics above 45 are classified as highly inconsistent. When these guidelines are used with the NHES data on dropouts, most statistics are classified as having moderate response variances.

Tables 9 and 10 show the estimates of dropout reporting for two age categories, 14- to 17-year-olds and 18- to 21-years-olds. Table 11 provides the statistical measures of reporting reliability computed separately for 14- to 17-year-olds and 18- to 21-year-olds. The reporting error is seen to be smaller for 18- to 21-year-olds than for 14- to 17-year-olds. Youths 18 to 21 years old were more likely to be their own HRI and, thus, to have a smaller reporting error than 14- to 17-year-olds. Figure 2 shows the percentage identically reported by age category. The percentage reported identically for 18- to 21-year-olds is about 94 percent, whereas it is only 63 percent for 14- to 17-year-olds.

The gross difference rate for 14- to 17-year-olds is over twice as large as for 18- to 21-year-olds. The net difference rate is about the same for both age categories. The index of inconsistency for status dropouts who were 14 to 17 years of age is about 37 percent but only about 6 percent for 18- to 21-year-olds. The statistics provided in table 11 confirm that the reporting reliability is higher for 18- to 21-year-olds than for 14- to 17-year-olds.

Another important characteristic that is often related to measurement error is race and ethnicity of the respondent. Tables 12, 13, and 14 show the same statistics separately for blacks/Hispanics and non-black, non-Hispanics. Blacks and Hispanics were not examined separately because of sample size limitations. Table 12 shows estimates of event dropouts separately for each of the two race/ethnicity categories. The item nonresponse rate was very low for race/ethnicity (less than 1 percent) so that nearly every person enumerated in the study was classified.

Table 9.—Estimates of event dropouts as reported by household respondents and youths, by age

Household respondent interview	Youth interviews					
	Total		Not-event dropouts		Event dropouts	
	14- to 17-year-olds	18- to 21-year-olds	14- to 17-year-olds	18- to 21-year-olds	14- to 17-year-olds	18- to 21-year-olds
Survey Estimates (in thousands)						
Total	13,241	13,094	12,971	12,772	270	322
Not-event dropouts	12,974	12,751	12,847	12,725	127	26
Event dropouts	267	343	124	47	143	296
Sample Sizes						
Total	471	1,066	449	1,031	22	35
Not-event dropouts	440	1,029	435	1,026	5	3
Event dropouts	31	37	14	5	17	32

Source: 1989 National Household Education Survey Field Test

Table 10.—Estimates of status dropouts as reported by household respondents and youths, by age

Household respondent interview	Youth interviews					
	Total		Not-status dropouts		Status dropouts	
	14- to 17-year-olds	18- to 21-year-olds	14- to 17-year-olds	18- to 21-year-olds	14- to 17-year-olds	18- to 21-year-olds
Survey Estimates (in thousands)						
Total	13,241	13,094	12,717	11,455	524	1,639
Not-status dropouts	12,776	11,495	12,583	11,392	194	103
Status dropouts	465	1,599	134	62	331	1,537
Sample Sizes						
Total	471	1,066	429	876	42	190
Not-status dropouts	417	880	411	868	6	12
Status dropouts	54	186	18	8	36	178

Source: 1989 National Household Education Survey Field Test

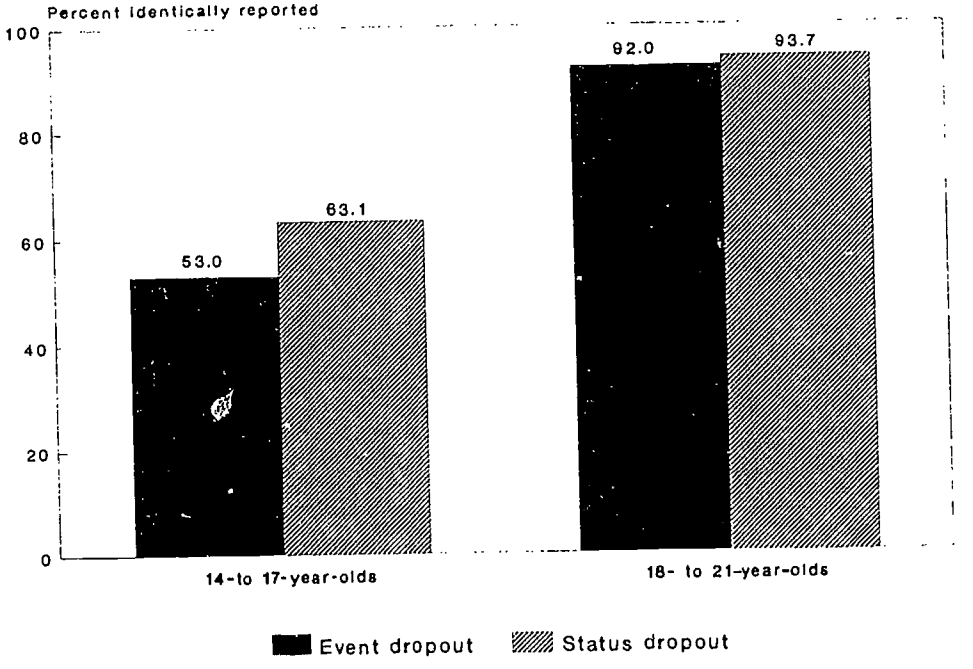
Table 11.—Measures of reporting reliability for event and status dropout statistics, by age

	Percent[1] of 14-21 year olds		Gross difference rate		Net difference rate		Percent identically reported		Index of inconsistency	
	14- to 17-year-olds	18- to 21-year olds	14- to 17-year-olds	18- to 21-year olds	14- to 17-year-olds	18- to 21-year olds	14- to 17-year-olds	18- to 21-year olds	14- to 17-year-olds	18- to 21-year olds
Event dropouts										
Estimate	2.0%	2.8%	1.9%	0.6%	0.0%	-0.2%	53.0%	92.0%	47.9%	10.9%
Standard error	--	--	0.7%	0.2%	0.7%	0.2%	14.6%	4.4%	17.4%	4.1%
Status dropouts										
Estimate	3.9%	12.3%	2.5%	1.3%	0.5%	0.3%	63.1%	93.7%	36.5%	5.9%
Standard error	--	--	0.8%	0.4%	0.8%	0.3%	11.3%	2.2%	11.9%	1.6%

[1]Based on Youth Interview reports.

Source: 1989 National Household Education Survey Field Test

Figure 2. — Percent identically reported for event and status dropout statistics, by age



Source: 1989 National Household Education Survey Field Test

The sample sizes associated with blacks/Hispanics who were reported as event dropouts by youths but reported as not-event dropouts by household respondents are too small to yield reliable estimates of measurement error. There were only nineteen blacks/Hispanics in these cells. Therefore, statistics provided for event dropouts by race/ethnicity should not be looked upon as stable or reliable estimates, and conclusions for race/ethnicity categories are only based on the status dropout sample.

Table 13 provides estimates of status dropouts separately by the two categories of race and ethnicity. The number of status dropouts estimated from the Youth Interview is slightly greater than the number estimated from the HRI for both race/ethnicity groups. This result is consistent with the findings from table 3.

Table 14 shows the statistics on measurement reliability separately for the two race/ethnicity categories. The percentage identically reported is higher for blacks/Hispanics when compared with the non-black, non-Hispanic group, although this difference may be attributed to sampling error. The gross difference rates are about the same for the two groups. The index of inconsistency is somewhat lower for blacks/Hispanics. Once again, these results may be a product of the large sampling error for these measures.

The event and status dropouts were composite variables based on school attendance and whether youths had a high school diploma or equivalent. As stated earlier, a status dropout is defined as a 14- to 21-year-old who was not enrolled in school during the data collection period (October 1989) and did not have a high school diploma or equivalent. Event dropouts are defined as the subset of status dropouts who were enrolled in school 1 year prior to the data collection period (October 1988).

The appendix to this report provides the statistics on measurement errors arising from using household respondents to collect information on the individual variables used to derive the estimated numbers of status and event dropouts. These tables help to establish the level of measurement error associated with each of the variables and shows how the measurement error typically increases when variables are constructed from other variables each of which has its own measurement error. In general, the measurement errors for these variables are relatively small.

Conclusions and Remarks

In general, the relative reporting reliability of the estimates for status dropouts was greater than that for event dropouts. The percentage reported identically and the index of inconsistency for event dropout statistics showed that these estimates were subject to more measurement error than the comparable estimates for status dropout statistics. The gross difference rate and the net difference rate estimates were larger for status dropout estimates, but when these statistics are compared relative to the size of the estimate then the status dropout statistics are relatively more reliable.

For older youths, household respondent responses were very close to the youths' responses, especially for status dropouts. Part of this similarity was because older youths had a higher chance of being their own HRI than younger youths. The bias after the offsetting misclassification was taken into account was small for both status dropouts and event dropouts. The 14- to 17-year-olds showed lower reporting reliability for household respondents. The percentage identically reported was about 30 percent to 40 percent lower than the ones for 18- to 21-year-olds. Furthermore, the index of inconsistency was much higher for 14- to 17-year-olds when compared to 18- to 21-year-olds.

No substantial differences in reporting reliability were observed among the different race/ethnic groups. However, analysis of the differences were limited by the small sample sizes for blacks/Hispanics.

The three factors mentioned in the methodology section should be considered together with the results when drawing conclusions about this study. The completion rate associated with the Youth Interview was about 85 percent. The above comparisons were based on the assumption that the nonrespondents were similar to those who responded, and that the respondent sample was a representative sample of all youths selected for the extended interview. Some of the estimated reliability measures may be biased, to the extent that these assumptions are violated.

The differences in reporting reliability should also be considered against the additional cost of collecting data from youths themselves. Furthermore, the analysis summarized in this report assumed that the youths' responses were the standard. This assumption does not alter the overall conclusions significantly. It was also assumed that dropping out of school did not have any substantial "social undesirability" for youths that affected their reporting.

Table 12.—Estimates of event dropouts as reported by household respondents and youths, by race/ethnicity

Household respondent interview	Youth interviews					
	Total		Not-event dropouts		Event dropouts	
	Black/ Hispanic	non-Black, non-Hispanic	Black/ Hispanic	non-Black, non-Hispanic	Black/ Hispanic	non-Black, non-Hispanic
Survey Estimates (in thousands)						
Total	6,693	19,514	5,485	19,137	209	377
Not-event dropouts	6,469	19,128	6,446	19,004	23	124
Event dropouts	224	386	39	133	186	253
Sample sizes						
Total	384	1,145	365	1,108	19	37
Not-event dropouts	364	1,097	361	1,093	3	4
Event dropouts	20	48	4	15	16	33

Source: 1989 National Household Education Survey Field Test

Table 13.—Estimates of status dropouts as reported by household respondents and youths, by race/ethnicity

Household respondent interview	Youth interviews					
	Total		Not-status dropouts		Status dropouts	
	Black/ Hispanic	non-Black, non-Hispanic	Black/ Hispanic	non-Black, non-Hispanic	Black/ Hispanic	non-Black, non-Hispanic
Survey Estimates (in thousands)						
Total	6,693	19,514	5,955	18,102	739	1,411
Not-status dropouts	5,966	18,191	5,893	17,968	73	223
Status dropouts	727	1,323	62	135	666	1,188
Sample sizes						
Total	384	1,145	317	982	67	163
Not-status dropouts	315	976	310	963	5	13
Status dropouts	69	169	7	19	62	150

Source: 1989 National Household Education Survey Field Test

Table 14.—Measures of reporting reliability for event and status dropout statistics, by race/ethnicity of the youth

	Percent[1] of 14-21 year olds		Gross difference rate		Net difference rate		Percent identically reported		Index of inconsistency	
	Black/ Hispanic	non-Black non-Hispanic	Black/ Hispanic	non-Black non-Hispanic	Black/ Hispanic	non-Black non-Hispanic	Black/ Hispanic	non-Black non-Hispanic	Black/ Hispanic	non-Black non-Hispanic
Event dropouts										
Estimate	3.2%	1.8%	0.9%	1.3%	-0.1%	0.3%	89.0%	67.2%	14.2%	33.9%
Standard error	--	--	0.5%	0.4%	0.5%	0.5%	6.3%	12.1%	6.1%	10.3%
Status dropouts										
Estimate	10.9%	7.5%	2.0%	1.8%	0.2%	0.5%	90.1%	84.2%	10.1%	14.5%
Standard error	--	--	0.9%	0.5%	0.7%	0.5%	4.3%	5.1%	4.0%	3.8%

Note: Sample sizes for event dropouts were too small to draw any conclusions for this group.

[1]Based on Youth Interview reports.

Source: 1989 National Household Education Survey Field Test

The responses for the HRI were reliable for estimating the number of status dropouts and, to a lesser extent, event dropouts. The net biases, as measured by the net difference rates, were typically small for these statistics. However, the index of inconsistency suggests that the event dropout statistics are subject to considerable measurement error. Analyses which tabulate the data by several different variables are more seriously affected by these types of errors. Also, as the sample size for the study increases, the measurement error for these statistics will become a more and more dominant factor since the measurement error is not reduced by increasing the sample size.

If a future study is designed to estimate the number of event dropouts and their characteristics, it will be possible to estimate the size of the measurement error associated with using a household respondent relative to the youth themselves from these results. The findings of the Field Test suggest that it may be prudent to collect data from both a household respondent and youths to improve the reliability of dropout estimates. This is particularly true for event dropout estimates.

References

1. Waksberg, J., "Sampling Methods for Random Digit Dialing," *Journal of the American Statistical Association*, Vol. 73, No. 361, March 1978.

2. U.S. Census Bureau, "The Current Population Survey Reinterview Program, January 1961 through December 1966, Technical Paper 19," 1968.

3. Hansen, Hurwitz, and Pritzker, "The Estimation and Interpretation of Gross Differences and the Simple Response Variance," Pergamon Press, 1964.

4. U.S. Bureau of the Census, "Evaluating Censuses of Population and Housing, Statistical Training Document," Washington D.C., 1985.

5. Mathiowetz, N. and R.M. Groves, "The Effects of Respondent Rules on Health Survey Reports," Chapter V in C.F. Cannell et. al. *An Experimental Comparison of Telephone and Personal Health Surveys*. NCHS, Rockville, Md., 1983.

6. U.S. Department of Commerce, "Consistency of Reporting of Ethnic Origin in the Current Population Survey, Technical Paper 31," 1974.

7. U.S. Department of Commerce, "An Error Profile: Employment as Measured by the Current Population Survey, Statistical Policy Working Paper 3," 1978.

8. U.S. Department of Commerce, "Accuracy of Data for Selected Housing Characteristics as Measured by Reinterviews," PHC(E)-10, 1975.

Appendix

Statistics on Measurement Errors for Variables Used to Derive Event and Status Dropout Rates

Table A-1.—Estimates of youths attending junior high school one year prior to the data collection as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not in junior high school a year ago	In junior high scho a year ago
Survey Estimates (in thousands)			
Total	26,335	22,079	4,256
Not in a junior high school a year ago	22,204	21,831	374
In junior high school a year ago	4,131	248	3,882
Sample Sizes			
Total	1,537	1,405	132
Not in junior high school a year ago	1,405	1,393	12
In junior high school a year ago	132	12	120

Source: 1989 National Household Education Survey Field Test

Table A-2.—Estimates of youths attending high school one year prior to the data collection as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not in high school a year ago	In high school a year ago
Survey Estimates (in thousands)			
Total	26,335	14,111	12,223
Not in a high school a year ago	13,913	13,325	588
In high school a year ago	12,421	786	11,635
Sample Sizes			
Total	1,537	978	559
Not in high school a year ago	983	935	48
In high school a year ago	554	43	511

Source: 1989 National Household Education Survey Field Test

Table A-3.—Estimates of youths with high school or equivalent diplomas as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	No high school or equivalent diploma	With high school or equivalent diploma
Survey Estimates (in thousands)			
Total	26,335	15,517	10,818
No high school or equivalent diploma	15,454	15,338	116
With high school or equivalent diploma	10,881	179	10,702
Sample Sizes			
Total	1,537	665	872
No high school or equivalent diploma	663	652	11
With high school or equivalent diploma	874	13	861

Source: 1989 National Household Education Survey Field Test

Table A-4.—Estimates of youths attending junior high school at the time of data collection as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not attending junior high school	Attending junior high school
Survey Estimates (in thousands)			
Total	26,335	25,277	1,058
Not attending junior high school	25,087	25,087	0
Attending junior high school	1,246	188	1,058
Sample Sizes			
Total	1,537	1,506	31
Not attending junior high school	1,499	1,499	0
Attending junior high school	38	7	31

Source: 1989 National Household Education Survey Field Test

Table A-5.—Estimates of youths attending high school at the time of data collection as reported by household respondents and youths

Household respondent interview	Youth interviews		
	Total	Not attending high school	Attending high school
Survey Estimates (in thousands)			
Total	26,335	14,087	12,247
Not attending high school	14,271	13,922	349
Attending high school	12,064	165	11,899
Sample Sizes			
Total	1,537	1,137	400
Not attending high school	1,155	1,130	25
Attending high school	382	7	375

Source: 1989 National Household Education Survey Field Test

Table A-6.—Measures of reporting reliability for junior high school or high school attendance statistics for one year prior to the reporting period

	Percent of all 14-21 year olds	Gross difference rate	Net difference rate	Percent identically reported	Index of inconsistency
Junior high school					
Estimate	16.2%	2.4%	+0.5%	91.2%	8.9%
Standard error	1.1%	0.6%	0.5%	2.6%	2.3%
High school					
Estimate	46.4%	5.2%	-0.8%	95.2%	10.5%
Standard error	1.5%	0.8%	0.8%	0.9%	1.7%

Source: 1989 National Household Education Survey (NHES) Field Test

Table A-7.—Measures of reporting reliability for junior high school or high school attendance statistics for the reporting period

	Percent of all 14-21 year olds	Gross difference rate	Net difference rate	Percent identically reported	Index of inconsistency
Junior high school					
Estimate	4.0%	0.7%	-0.7%	100.0%	7.9%
Standard error	0.6%	0.3%	0.3%	0.0%	3.0%
High school					
Estimate	46.5%	2.0%	+0.7%	97.2%	3.9%
Standard error	1.5	0.4%	0.4%	0.6%	0.8%

Source: 1989 National Household Education Survey (NHES) Field Test

Table A-8.—Measures of reporting reliability for statistics on number of youths with high school or equivalent diplomas

	Percent of all 14-21 year olds	Gross difference rate	Net difference rate	Percent identically reported	Index of inconsistency
Youths with diploma					
Estimate	41.1%	1.1%	-0.2%	98.9%	2.3%
Standard error	1.4%	0.4%	0.3%	0.4%	0.8%

Source: 1989 National Household Education Survey (NHES) Field Test

United States
Department of Education
Washington, D.C. 20208–5651

Official Business
Penalty for Private Use, $300

Postage and Fees Paid
U.S. Department of Education
Permit No. G–17

FOURTH CLASS BOOK RATE