NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

The Working Paper Series was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

U. S. Department of Education
Office of Educational Research and Improvement

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

Imputation of Test Scores in the National Education Longitudinal Study of 1988 (NELS:88)

Working Paper No. 2001–16 September 2001

Contact: Ralph Lee
Statistical Standards Program
E-mail: ralph.lee@ed.gov

U. S. Department of Education
Office of Educational Research and Improvement

U.S. Department of Education
Rod Paige
Secretary

Office of Educational Research and Improvement
Grover J. Whitehurst
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Acting Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Office of Educational Research and Improvement
U.S. Department of Education
1990 K Street NW
Washington, DC 20006

September 2001

The NCES World Wide Web Home Page is
http://nces.ed.gov

Suggested Citation

U.S. Department of Education. National Center for Education Statistics. *Imputation of Test Scores in the National Education Longitudinal Study of 1988 (NELS:88)*. Working Paper No. 2001–16, by Maxime C. Bokossa and Gary G. Huang. Project Officer, Ralph Lee. Washington, DC: 2001.

Foreword

In addition to official NCES publications, NCES staff and individuals commissioned by NCES produce preliminary research reports that include analyses of survey results, and presentations of technical, methodological, and statistical evaluation issues.

The ***Working Paper Series*** was initiated to promote the sharing of the valuable work experience and knowledge reflected in these preliminary reports. These reports are viewed as works in progress, and have not undergone a rigorous review for consistency with NCES Statistical Standards prior to inclusion in the Working Paper Series.

Copies of Working Papers can be downloaded as pdf files from the NCES Electronic Catalog (http://nces.ed.gov/pubsearch/), or contact Sheilah Jupiter at (202) 502–7444, e-mail: sheilah_jupiter@ed.gov, or mail: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 1990 K Street NW, Room 9048, Washington, DC 20006.

Marilyn M. Seastrom
Chief Mathematical Statistician
Statistical Standards Program

Ralph Lee
Mathematical Statistician
Statistical Standards Program

This page intentionally left blank.

Imputation of Test Scores

in the

National Education Longitudinal Study of 1988 (NELS:88)

Prepared by:

Maxime C. Bokossa
Gary G. Huang
Synectics for Management Decisions, Inc.

Prepared for:

U.S. Department of Education
Office of Educational Research and Improvement
National Center for Education Statistics

September 2001

This page intentionally left blank.

Table of Contents

List of Tables

List of Figures

CHAPTER I
INTRODUCTION

The National Education Longitudinal Study of 1988 (NELS:88) is the only current National Center for Education Statistics (NCES) dataset that contains scores from cognitive tests given to the same set of students across multiple points in time. The resulting longitudinal test data offer the possibility of researching cognitive gains from middle school through high school—an attractive feature. However, as is inevitable in any survey, cognitive test data are missing for some individuals in each round; the problem is more severe in the second follow up (F2) than in the earlier rounds. Therefore, NCES decided to use imputation to reduce the bias caused by nonresponse.

This study involved a two-step process for implementing this imputation. The first step, as described in chapter 2, was to conduct a simulation study to evaluate two different imputation procedures currently used at NCES: a *model-based random imputation* method called PROC IMPUTE and a *within-class random hot-deck imputation*. In our simulation study, we first examined and selected a range of auxiliary variables that are conceptually and empirically related to the F2 test scores, and then we imputed the Item Response Theory (IRT) theta scores in math and reading. The findings of the simulation study confirmed that PROC IMPUTE performed better (Hu and Salvucci 1999). The second step, as described in chapter 3, involved using PROC IMPUTE to impute missing F2 cognitive test scores in four subject areas: math, science, reading, and history/citizenship/geography. The results provide end users with complete cognitive test data for both cross-sectional and longitudinal research with the F2 data or the base-year through the second follow-up (BY–F2) panel data. As a future step, other measurement scales (proficiency scores, standardized scores, and the number right scores) may be subsequently converted using the theta scores.

BACKGROUND

In NELS:88, the respondents' cognitive ability and the growth (cognitive gains) from 8th through 12th grades at the group and individual levels were measured by a calibrated scale based on Item Response Theory (IRT). This calibration process requires that items are relatively unifactorial across grades in each subject area; that is, with the same dominant factor underlying all test forms in a given subject, say, math (Rock and Pollack 1995). There should be a common set of "anchor" items across adjacent forms, and most content areas should be represented in all grade forms. In NELS:88, the increasingly difficult levels from 8th through 12th grades were created by raising the problem-solving demands in the existing content areas and adding new content in the later forms, especially at 12th grade.

IRT assumes that a test taker's probability of answering an item correctly is a function of his or her ability and one or more characteristics of the test item itself. The three-parameter IRT logistic model uses the pattern of right, wrong, and omitted responses to the items administered in a test form, and the difficulty, discriminating ability, and "guess-ability" of each item, to place each test taker at a particular point, q (theta), on a continuous ability scale. The probability of a correct answer (called the theta score) on item i can be expressed as:

$$P_i(q) = c_i + \frac{(1 - c_i)}{1 + e^{-1.702a_i(q - b_i)}},$$

where q is the ability of the test taker, a_i is discrimination of item i, or how well the item distinguishes between ability levels at a particular point, b_i is the difficulty of item i, and c_i is the "guess-ability" of item i.

A computer program is used to calculate the marginal maximum-likelihood estimates of the IRT parameters that best fit test takers' responses (Muraki and Bock 1991). To assess the models' match with the test data, one compares the IRT-estimated parameters with the actual proportion of correct answers to a test item for test takers grouped by ability. If the IRT-estimated curves and the actual data points match closely, then the theoretical model represents the data accurately. After the parameters for a set of test items are calibrated on the same scale as the test takers' ability estimates, a test taker's probability of a correct answer to each item in the test battery can be estimated, even for items that were not administered to the test taker. Theta scores can be used to derive other test scores: the IRT-estimated number correct score in a subject area is the sum of the probabilities of correct answers for the items in the area.

NELS:88 nonresponse issues: Nonresponse is always a concern in survey data, and some cases in the NELS:88 cognitive test data are missing in each round due to absence, nonparticipation, or results that were unscorable because of too many unattempted test items. This missingness problem is more severe for math theta scores in F2 (22.9 percent missing scores) than in the earlier two rounds of tests (3.7 percent and 6.0 percent missing scores for the base-year (BY) and the first follow up (F1), respectively), as shown in table 1.

Table 1. Number of students and mean math scores by test missing status

Test missing status	Number of students	Mean math theta scores		
		BY	F1	F2
Completed all tests	11,832	46.16	51.53	54.80
Missing BY only	415 (BY: total missing 610)	--	48.86	51.94
Missing F1 only	444 (F1: total missing 995)	42.60	--	49.40
Missing F2 only	3,117 (F2: total missing 3,775)	43.96	48.62	--
Missing BY and F1	23	--	--	44.63
Missing BY and F2	130	--	44.73	--
Missing F1 and F2	486	40.09	--	--
Missing all tests	42	--	--	--

-- = missing

Note: The above information is based on the total BY–F2 panel of 16,489 students.

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88).

The sample weighting adjustment cannot fully solve the problem resulting from survey nonresponse, neither in theory nor in practice (Rubin 1996). Specifically, the bias generated by missing cognitive scores cannot be corrected by the NELS:88 sampling weights because the weights were constructed to remedy unit nonresponse, not item nonresponse (Ingels et al. 1994, p. 70). In fact, the joint impact of item nonresponse to cognitive tests and unit nonresponse on NELS:88 tends to damage the data quality to a potentially dangerous extent. The weighted percentage of students who took all four cognitive tests in all three waves of the survey was 65 percent of the eligible core panel sample (see Rock and Pollack 1995, table 1.1, p. 2).

In addition, Rock and Pollack (1995, pp. 53–56) demonstrated that the missingness pattern of F2 test scores across demographic subgroups was not completely at random. Our tabulation of the BY–F2 panel data confirms this. Table 2 presents a comparison of the rate of missing F2 test scores for some basic demographic subgroups of students in the BY–F2 panel who completed all three tests and those who missed the F2 test. It shows that minority students and students in the lowest socioeconomic (SES) quartile were more likely than others to miss the test. Thus, NELS:88 estimates of academic performance based on the available cases could be biased.

Table 2. Number of students and mean math theta scores by sex, race/ethnicity, and SES quartile

		Number who completed all 3 tests in BY–F2 panel	Number of students with F2 test score available	Percent of BY–F2 panel with missing F2 test scores	Mean of F2 math test scores
TOTAL		16,489	12,714	22.9%	54.5
Sex	Male	8,349	6,430	23.0%	53.9
	Female	8,140	6,284	22.8%	55.1
Race/ Ethnicity[1]	White and Asian	12,657	9,935	21.5%	56.1
	Black, Hispanic, Indian	3,823	2,773	27.5%	48.6
SES[2]	Lowest quartile	4,121	2,989	27.5%	47.8
	2nd quartile	4,095	3,187	22.2%	52.2
	3rd quartile	4,147	3,260	21.4%	55.5
	Highest quartile	4,125	3,278	20.5%	61.8

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88).

The gain measure, which is of critical utility in NELS:88 longitudinal research, is thus built upon test data with high levels of item nonresponse. To assure NELS:88 data quality, strategies other than weighting are needed to address the item nonresponse problem. Imputation of missing test scores is one viable strategy.

It is feasible to impute F2 cognitive test scores because a great deal of information is available to reasonably predict the missing scores. This information includes student sociodemographic background, school experience (e.g., coursework, ability and curriculum program placements, and enrichment activity participation), self-reported achievement level, and available scores in other subjects. Furthermore, the general pattern in which such predictive variables relate to achievement is known in the educational research literature. We developed our imputation models based on such knowledge. (Our approach to NELS:88 cognitive test score imputations could be applicable to similar problems likely to arise in the Early Childhood Longitudinal Studies (ECLS), conducted by NCES, which will also include multiple rounds of cognitive tests.)

CHAPTER 2
SIMULATION STUDY COMPARING THE PROC IMPUTE AND HOT-DECK IMPUTATION METHODS

Our simulation study compared PROC IMPUTE and the hot-deck imputation method by imputing the F2 IRT-estimated theta scores in math and reading. To impute missing test scores in a given subject, we used information from available tests in other subjects, student demographic and socioeconomic background, academic coursework, and self-reported grade point averages. We also compared the imputed F2 test scores with BY and F1 test scores in a given subject. We used three criteria to compare the accuracy of the two sets of imputations: the average imputing error, the variance, and the mean bias.

APPROACH

We decided to impute the IRT-estimated theta scores since theta scores are the original estimates of the test takers' probability of correctly answering items in a given set of test items.

As mentioned previously, the F2 missing test scores were not "missing completely at random" (MCAR) as defined by Little and Rubin (1987). That is, the cases that did not have scorable tests in F2 were systematically different from the cases that had completed the three tests in a variety of auxiliary variables, including background and schooling (see table 2 and Rock and Pollack 1995, pp. 53–56). Such non-MCAR missingness patterns call for imputation based on information for a subsample that had completed test scores but shared attributes with the missing cases. Our first step, therefore, was to examine a range of candidate variables in order to select the best auxiliary variables; that is, those which were related to test missingness.

Selection of Auxiliary Variables

The candidate variables were race/ethnicity, sex, SES, coursework in the target subject areas, advanced academic program placement, F1 and F2 dropout status, early graduation status, and BY and F1 cognitive test scores. To determine their utility in the imputation model, we examined bivariate correlations between these variables and the cognitive test scores in two subject areas, math and reading. We then selected variables that correlated highly with the theta scores. Next we identified important predictors of the cognitive test outcome by fitting regression models. The final regression model reflected test scores that were homogeneous within the imputation classes defined by the covariates.

Simulation Study

We studied two imputation techniques, namely, a *model-based imputation* method implemented by computer software called PROC IMPUTE and a *within-class random hot-deck imputation* method. The study included simulating a few levels and patterns of missingness (about 20 percent of the data were made missing) in the NELS:88 BY–F2 panel cases where the BY, F1, and F2 test scores are all nonmissing. We compared statistics derived from the incomplete data with the data after imputing simulated missing cases. Three criteria were used to compare the accuracy of the two types of imputations: the average imputing error, the bias of the variance, and the mean bias.

The relative bias of the variance estimate is defined as

$$Relative\ \ Bias = \frac{(Estimated\ \ Var) - (True\ \ Var)}{True\ \ Var}$$

and the average imputation error is defined as

$$\sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i^* - y_i)^2}$$

where m is the number of missing values, y_i is the true value which is intentionally set to missing, and y_i^* is the imputed value for the i-th missing case. That an imputation method has smaller average imputation errors only implies that the method provides imputations on average closer to the real values. This does not necessarily mean that it gives more accurate estimates for all types of statistics, although that is true in many situations.

Description of Imputation Methods

Within-class random hot-deck imputation: Since we understand reasonably well the factors related to F2 test nonresponse and have data on such factors, we could assume model-based approaches would probably produce more accurate imputation than randomization-based approaches if the model assumptions were satisfied (Hu and Salvucci 1999). Thus, we imputed the IRT-estimated number of the right score in each subject using F2 cross-sectional data on student sociodemographic and socioeconomic background, academic coursework, self-reported grade average point, and available test scores on subjects other than the one to be imputed.

For the implementation of the within-class random hot-deck imputation method, we first sorted the dataset by the auxiliary variables in order to obtain homogeneous cells called imputation classes. To impute a missing value in a

given imputation class, we randomly selected an observed value of the target variable in that class to fill-in for the missing value.

PROC IMPUTE: To overcome the underestimation of variance which is typical in a hot-deck imputation method or a regression-based imputation method, we also added disturbance by using the software package PROC IMPUTE (McLaughlin 1991).

PROC IMPUTE combines the procedures of regression-based and data sampling (often called "hot-deck") methods. Regression involves generating a function, $\hat{y} = f(x_1, x_2, ..., x_p)$, that relates a "target" variable (cognitive test score) to auxiliary variables, then uses the function along with the existing values of the auxiliary variables to compute $\hat{y}$ whenever it is missing. Data sampling involves subsetting the data on the basis of relevant variables and randomly selecting a value for the target variable from an available target variable within the same subset.

PROC IMPUTE considers each variable on the file in turn as a target variable whose missing values are to be filled in, and it uses information on other variables to minimize the error in imputing each target variable. Three steps are taken to impute each variable in PROC IMPUTE.

First, stepwise regression analyses are performed "simultaneously" for each variable. During these analyses, an ordered list of the imputation variables is constructed. The regression analysis for each variable uses as predictors all the complete variables, including the previously imputed variables. The process terminates when there are no more permissible predictors that provide a significant improvement of fit in the prediction of any of the target variables. Second, homogeneous cells (imputation classes) are created for records that have close predicted regression values. Finally, two donors are drawn from the adjacent cells. Each missing record in a given cell is imputed with a weighted average of these two donors with probability proportional to the observed frequencies within the two cells.

PROC IMPUTE runs all the imputation procedures automatically and generates a dataset in which all the records are complete. Imputed data flags are also automatically created by the software and set for each variable; a value of "I" corresponds to imputed values, "R" to reported values, and "A" to skip missing values.

SIMULATION RESULTS

Math Theta Score

We used the F2 panel sample members that had nonmissing math theta scores and nonmissing information for the following auxiliary variables: sex, race/ethnicity, SES, units in foreign languages, units in physics, BY grade composites, and teacher's opinion about student attending college. We selected 1,996 cases, about 20 percent, from the F2 panel members and set their math theta scores as missing. To simulate the actual missingness pattern, the rate of missingness across sex, race/ethnicity, and SES quartiles mimicked that of the actual F2 test missing cases. We used PROC IMPUTE and random hot-deck to impute these simulated missing cases. The mean and variance for the math scores were calculated for the following four groups:

1. A group of 10,248 cases in the F2 panel that reported the math theta scores and auxiliary variables specified above;
2. A group that included the 8,252 cases with actual math theta scores and 1,996 cases with imputed scores using PROC IMPUTE;
3. A group that included the 8,252 cases with actual math theta scores and 1,996 cases with imputed scores using the hot-deck method; and
4. A group of 8,252 cases with actual math theta scores (the 1,996 cases were deleted as "missing"). This group simulates the current scenario in NELS:88 where there are missing test scores, but no imputation has been used.

Group 1 estimates served as the "true scores." Groups 2, 3, and 4 estimates were compared with the true Group 1 estimates to examine if Group 2 (with PROC IMPUTE imputation) did better than Group 3 (with hot-deck imputation) and Group 4 (non-imputed). Table 3 provides the results for average imputation error for the math theta score. Then figure 1 compares the results for the bias of the mean, while table 4 presents the relative bias of the variance for the math theta score. Tables 5 and 6 show, respectively, the mean and standard deviation for the multiple imputation using the PROC IMPUTE and within-class random hot-deck imputation methods. Note that in the race/ethnicity subgroup, whites and Asians were combined because preliminary results had shown that both whites and Asians have on average higher math scores than the other racial/ethnic groups.

About 20 percent of the math scores were imputed using first PROC IMPUTE, and then the random hot-deck imputation method. The average imputation error is consistently lower for PROC IMPUTE than it is for hot-deck in each sociodemographic subgroup, and overall (see table 3).

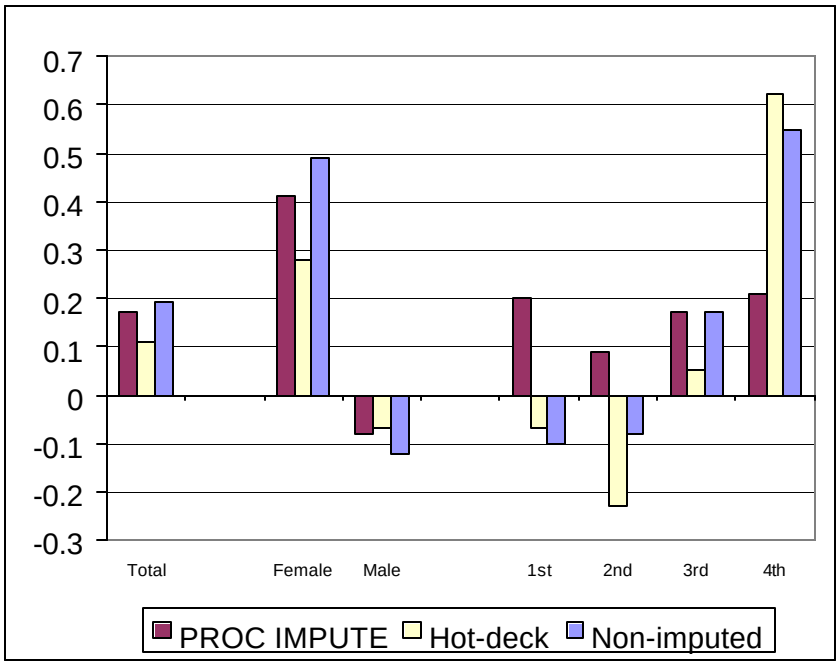
Table 3. Percentage of missing values and average imputation error for math score

		Number of students	Percent of imputed values	Average imputation error PROC IMPUTE	Hot-deck
TOTAL		10,248	19.5%	13.56	14.50
Sex	Female	5,139	20.2%	13.23	14.51
	Male	5,109	18.8%	13.90	14.49
Race/ ethnicity	White and Asian	8,196	19.0%	13.58	14.32
	Black, Hispanic, Indian	2,052	21.3%	13.49	15.10
SES	Lowest quartile	2,176	20.3%	13.82	14.34
	2nd quartile	2,596	19.7%	14.16	14.98
	3rd quartile	2,734	19.3%	12.77	14.18
	Highest quartile	2,742	18.8%	13.51	14.47

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Figure 1 shows the bias of the mean after using PROC IMPUTE and the random hot-deck imputation method, as well as the bias of the mean for the incomplete math score without any imputation. No one of the three methods shows a consistent improvement in the mean bias across the sociodemographic subgroups or overall.

Figure 1. Comparison of bias of the mean for math theta score



Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 4 shows that the relative bias of the variance is consistently smaller for PROC IMPUTE than it is for hot-deck and the non-imputed group, in each of the sociodemographic subgroups of study, and overall, with the exception of the highest quartile of the SES subgroup.

Table 4. Comparison of relative bias of variance for math theta score

		Relative bias of variance		
		Non-imputed	**PROC IMPUTE**	**Hot-deck**
TOTAL		0.055	0.001	0.060
Sex	Female	-0.005	0.053	0.069
	Male	0.010	0.061	0.056
Race/ethnicity	White and Asian	0.018	0.059	0.068
	Black, Hispanic, Indian	0.021	0.046	0.076
SES	Lowest quartile	-0.003	0.036	0.051
	2nd quartile	0.009	0.053	0.049
	3rd quartile	0.005	0.062	0.076
	Highest quartile	-0.021	0.002	-0.009

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 5 presents the resulting mean for a set of five imputations on the math theta score using the PROC IMPUTE and within-class random hot-deck imputation method. As we can see, the multiple imputation means based on the PROC IMPUTE method are consistently closer to the true means than are the means based on the within-class random hot-deck imputation method. This observation is valid for each of the study's sociodemographic subgroups, and overall.

Table 5. Comparison of mean for multiple imputation for math theta score

		Mean bias		
		True	**PROC IMPUTE**	**Hot-deck**
TOTAL		55.16	55.17	55.27
Sex	Female	54.59	54.60	54.86
	Male	55.74	55.74	55.69
Race/ethnicity	White and Asian	56.62	56.62	56.76
	Black, Hispanic, Indian	49.36	49.37	49.35
SES	Lowest quartile	48.78	48.79	48.62
	2nd quartile	52.65	52.64	52.54
	3rd quartile	55.88	55.87	56.09
	Highest quartile	61.90	61.91	62.32

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

With a set of five imputations on the math theta score using the PROC IMPUTE and within-class random hot-deck imputation methods, we calculated the resulting standard deviations (see table 6). From table 6, it is clear that the multiple imputation standard deviations based on the PROC IMPUTE method are consistently closer to the true standard deviations than are the standard deviations based on the within-class random hot-deck imputation method. This held true for all the sociodemographic subgroups of study.

Table 6. Comparison of standard deviation for multiple imputation math theta score

		Standard Deviation		
		True	**PROC IMPUTE**	**Hot-deck**
TOTAL		10.27	10.28	10.52
Sex	Female	9.92	9.94	10.15
	Male	10.58	10.59	10.86
Race/ethnicity	White and Asian	10.03	10.05	10.28
	Black, Hispanic, Indian	9.09	9.11	9.30
SES	Lowest quartile	8.69	8.72	8.82
	2nd quartile	9.40	9.42	9.64
	3rd quartile	9.31	9.32	9.49
	Highest quartile	8.96	8.96	9.02

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Reading Theta Score For the reading cognitive test score simulation study, we used the F2 panel sample members that had nonmissing reading theta scores and nonmissing auxiliary variables. The auxiliary variables considered here were sex, race/ethnicity, SES, units in foreign languages, units in reading, units in chemistry, grade composites from base-year, and teacher's opinion about student attending college. We selected 2,017 cases, about 20 percent, from the F2 panel members and set their reading theta scores as missing. We used PROC IMPUTE and random hot-deck to impute these simulated missing cases. The mean and variance for the reading scores were calculated for the following four groups:

(1) a group of 10,249 cases in the F2 panel that reported the reading theta scores and auxiliary variables specified above;
(2) a group of 8,232 cases with actual reading theta scores and 2,017 cases with imputed scores using PROC IMPUTE;
(3) a group of 8,232 cases with actual reading theta scores and 2,017 cases with imputed scores using the hot-deck method; and
(4) a group of 8,232 cases with actual reading theta scores.

Table 7 provides the calculated average imputation error for the reading theta score, figure 2 displays the calculated bias of the mean, and table 8 presents the calculated relative bias of the variance for the reading theta scores when non-imputed and when imputed using PROC IMPUTE and random hot-deck. Table 9 shows the mean for a set of five imputations using the PROC IMPUTE and within-class random hot-deck imputation methods, and table 10 shows the corresponding standard deviations. Note that, unlike the math test score, the race/ethnicity variable here is categorized by whites on one hand and the other racial/ethnic groups on the other hand.

As in the simulation of math theta scores, around 20 percent of the reading scores were set to missing and imputed using first the PROC IMPUTE and then the random hot-deck imputation methods. The average imputation error is consistently lower for PROC IMPUTE than it is for hot-deck, in each sociodemographic subgroup, and overall (see table 7).

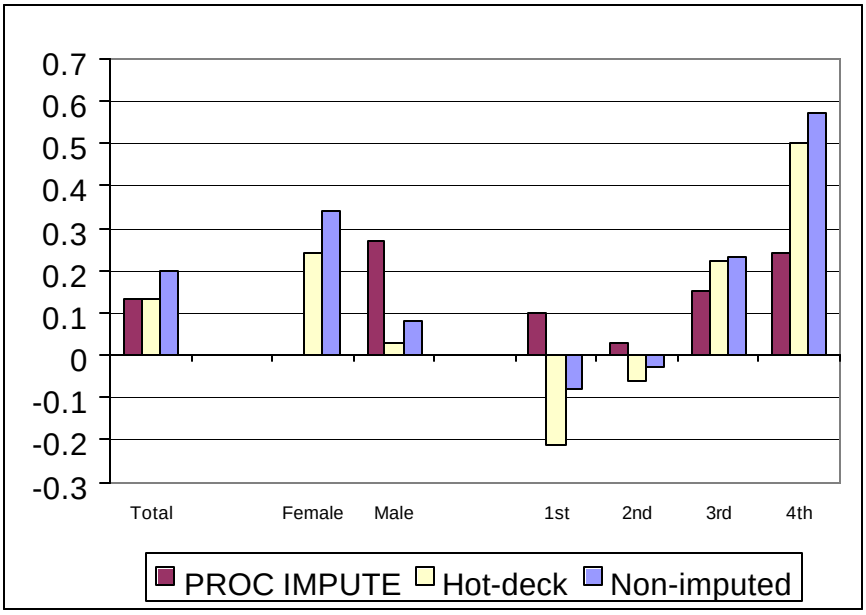
Table 7. Percentage of missing values and average imputation error for reading score

		Number of students	Percent of imputed values	Average imputation error PROC IMPUTE	Hot-deck
TOTAL		10,249	19.7%	13.86	14.70
Sex	Female	5,144	20.0%	13.86	14.50
	Male	5,105	19.4%	13.85	14.90
Race/ ethnicity	White	7,594	19.3%	13.63	14.48
	Asian, Black, Hispanic, Indian	2,655	20.8%	14.44	15.27
SES	Lowest quartile	2,178	20.0%	14.36	14.69
	2nd quartile	2,594	19.5%	14.14	15.66
	3rd quartile	2,738	20.2%	13.51	14.27
	Highest quartile	2,739	19.1%	13.51	14.19

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

In figure 2, note that the bias of the mean for female reading theta score is zero for PROC IMPUTE. Nevertheless, the bias of the mean does not show that any particular method is consistently better across all sociodemographic subgroups.

Figure 2. Comparison of bias of the mean for reading theta score



Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

However, the relative bias of the variance is consistently smaller for PROC IMPUTE than it is for the hot-deck and the non-imputed groups, in each sociodemographic subgroup, and overall, with the exception of the third and fourth quartile of the socioeconomic status subgroup (see table 8).

Table 8. Comparison of relative bias of variance for reading theta score

		Relative bias of variance		
		Non-imputed	**PROC IMPUTE**	**Hot-deck**
TOTAL		0.034	-0.009	0.037
Sex	Female	0.005	0.035	0.031
	Male	-0.015	0.028	0.039
Race/ethnicity	White	-0.001	0.035	0.038
	Asian, Black, Hispanic, Indian	0.004	0.038	0.035
SES	Lowest quartile	0.021	0.024	0.030
	2nd quartile	-0.003	0.035	0.021
	3rd quartile	-0.036	0.018	0.029
	Highest quartile	-0.038	-0.011	-0.002

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 9 provides the mean for multiple imputation on the reading theta score using the PROC IMPUTE and within-class random hot-deck imputation method. In most of the sociodemographic subgroups of study, and overall, the multiple imputation means based on the PROC IMPUTE method are closer to the true means than are the means based on the within-class random hot-deck imputation method.

Table 9. Comparison of mean for multiple imputation for reading theta score

			Mean Bias	
		True	**PROC IMPUTE**	**Hot-deck**
TOTAL		53.71	53.78	53.88
Sex	Female	54.82	54.85	55.10
	Male	52.59	52.70	52.66
Race/ethnicity	White	54.86	54.90	55.04
	Asian, Black, Hispanic, Indian	50.41	50.58	50.57
SES	Lowest quartile	47.99	48.03	47.90
	2nd quartile	51.49	51.58	51.55
	3rd quartile	54.33	54.36	54.46
	Highest quartile	59.74	59.84	60.27

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

With a set of five imputations on the math theta score using the PROC IMPUTE and within-class random hot-deck imputation methods, we calculated the resulting standard deviations (see table 10). It is clear that the multiple imputation standard deviations based on the PROC IMPUTE method are consistently closer to the true standard deviations than are the standard deviations based on the within-class random hot-deck imputation method. This held true for all the sociodemographic subgroups of study.

Table 10. Comparison of standard deviation for multiple imputation reading theta score

		Standard deviation		
		True	PROC IMPUTE	Hot-deck
TOTAL		10.61	10.59	10.78
Sex	Female	10.17	10.17	10.31
	Male	10.92	10.89	11.10
Race/ethnicity	White	10.33	10.32	10.51
	Asian, Black, Hispanic, Indian	10.71	10.69	10.86
SES	Lowest quartile	9.43	9.43	9.53
	2nd quartile	10.05	10.04	10.23
	3rd quartile	10.06	10.01	10.15
	Highest quartile	9.29	9.23	9.28

Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

CONCLUSION OF THE SIMULATION STUDY

Using PROC IMPUTE to impute the missing math and reading cognitive test scores produced better results than using the random hot-deck imputation method or no imputation in the simulation study that we conducted using NELS:88 second follow-up (F2) data. We therefore chose PROC IMPUTE as our method of imputing the NELS:88 theta scores in the second part of this project. Those results are discussed in the next chapter.

CHAPTER 3
IMPUTATION OF NELS:88 2ND FOLLOW-UP THETA SCORES USING PROC IMPUTE

The results of the simulation study described in the previous chapter showed that PROC IMPUTE was the appropriate choice of imputation techniques for imputing the missing test score for the second follow-up in the NELS:88. It generated the "best" scores based on the criteria used; that is, PROC IMPUTE was the method with the least average imputing error and mean bias and with the least distortion in variance. Hence, in this chapter, we used PROC IMPUTE to impute the missing test scores in the four tested F2 subject areas: math, science, reading, and history/ citizenship/geography.

Math Theta Score

We used PROC IMPUTE to impute the 3,775 missing cases for the math theta score. We started by using the full BY–F2 panel sample members and the following auxiliary variables:

- from F2—sex, race/ethnicity, SES, units in foreign languages, units in math, units in geometry, units in chemistry, and units in physics;
- from F1—the teacher's opinion about whether the student will go to college or not, number of course the student took in geometry, and math theta score; and
- from BY—grade composite variable and math theta score.

We then computed the overall mean and standard deviation for the math theta, and also the mean and standard deviation for the math theta score across sex, race/ethnicity, and SES quartiles. Those were compared for the following two groups:

1. A group of 12,714 cases in the BY–F2 panel that reported the math theta scores; and
2. A group that included the 12,714 cases with actual math theta scores and 3,775 cases with imputed scores using PROC IMPUTE.

The mean and standard deviation of the math theta score for both groups defined above are shown in tables 11 and 12, respectively.

Table 11. Comparison of mean for math theta score before and after imputation

		Number of students		Mean math theta score	
		Overall	With missing math score	Non-imputed	PROC IMPUTE
TOTAL		16,489	3,775	54.50	53.79
Sex	Female	8,349	1,919	53.90	53.30
	Male	8,140	1,856	55.10	54.30
Race/ ethnicity[1]	White and Asian	12,657	2,722	56.13	55.57
	Black, Hispanic, Indian	3,823	1,050	48.64	47.92
SES[2]	Lowest quartile	4,121	1,132	47.84	47.33
	2nd quartile	4,095	908	52.23	51.57
	3rd quartile	4,147	887	55.52	54.90
	Highest quartile	4,125	847	61.76	61.35

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 12. Comparison of standard deviation (SD) for math theta score before and after imputation

		Number of students		SD math theta score	
		W/nonmissing math score	W/missing math score	Non-imputed	PROC IMPUTE
TOTAL		12,714	3,775	10.50	10.69
Sex	Female	6,430	1,919	10.21	10.37
	Male	6,284	1,856	10.76	10.99
Race/ ethnicity[1]	White and Asian	9,935	2,722	10.28	10.49
	Black, Hispanic, Indian	2,773	1,050	9.10	9.11
SES[2]	Lowest quartile	2,989	1,132	8.72	8.75
	2nd quartile	3,187	908	9.43	9.61
	3rd quartile	3,260	887	9.45	9.68
	Highest quartile	3,278	847	9.17	9.45

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Reading Theta Score We used PROC IMPUTE to impute the 3,771 missing cases for the reading theta score. We started by using the full BY–F2 panel sample members and the following auxiliary variables:

- from F2—sex, race/ethnicity, SES, units in foreign languages, and units in chemistry;
- from F1—the teacher's opinion about whether the student will go to college or not, number of course the student took in foreign languages, and reading theta score;
- from BY—grade composite variable and reading theta score.

We then computed the overall mean and standard deviation for the reading theta, and also the mean and standard deviation for the reading theta score across sex, race/ethnicity, and SES quartiles. Those were compared for the following two groups:

1. A group of 12,718 cases in the BY–F2 panel that reported the reading theta scores; and
2. A group that included the 12,718 cases with actual reading theta scores and 3,771 cases with imputed scores using PROC IMPUTE.

The mean and standard deviation of the reading theta score for both groups defined above are shown in tables 13 and 14, respectively.

Table 13. Comparison of mean for reading theta score before and after imputation

		Number of students		Mean reading theta score	
		Overall	With missing reading score	Non-imputed	PROC IMPUTE
TOTAL		16,489	3,771	53.17	52.58
Sex	Female	8,349	1,913	54.22	53.60
	Male	8,140	1,858	52.09	51.53
Race/ ethnicity[1]	White and Asian	12,657	2,717	54.62	54.13
	Black, Hispanic, Indian	3,823	1,051	47.97	47.45
SES[2]	Lowest quartile	4,121	1,135	47.29	46.80
	2nd quartile	4,095	905	51.11	50.63
	3rd quartile	4,147	882	54.01	53.59
	Highest quartile	4,125	848	59.68	59.26

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 14. Comparison of standard deviation (SD) for reading theta score before and after imputation

		Number of students		SD reading theta score	
		W/nonmissing reading score	W/missing reading score	Non-imputed	PROC IMPUTE
TOTAL		12,718	3,771	10.81	10.96
Sex	Female	6,436	1,913	10.43	10.66
	Male	6,282	1,858	11.09	11.17
Race/ ethnicity[1]	White and Asian	9,940	2,717	10.59	10.74
	Black, Hispanic, Indian	2,772	1,051	9.97	10.08
SES[2]	Lowest quartile	2,986	1,135	9.48	9.70
	2nd quartile	3,190	905	10.18	10.29
	3rd quartile	3,265	882	10.17	10.30
	Highest quartile	3,277	848	9.46	9.58

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Science Theta Score We used PROC IMPUTE to impute the 3,858 missing cases for the science theta score. We started by using the full BY–F2 panel sample members and the following auxiliary variables:

- from F2—sex, race/ethnicity, SES, units in foreign languages, units in math, units in geometry, units in chemistry, and units in physics;
- from F1—the teacher's opinion about whether the student will go to college or not, number of course the student took in geometry, and science theta score;
- from BY—grade composite variable and science theta score.

We then computed the overall mean and standard deviation for the science theta, and also the mean and standard deviation for the science theta score across sex, race/ethnicity, and SES quartiles. Those were compared for the following two groups:

1. A group of 12,631 cases in the BY–F2 panel that reported the science theta scores; and
2. A group that included the 12,631 cases with actual science theta scores and 3,858 cases with imputed scores using PROC IMPUTE.

The mean and standard deviation of the science theta score for both groups defined above are shown in tables 15 and 16, respectively.

Table 15. Comparison of mean for science theta score before and after imputation

		Number of students		Mean science theta score	
		Overall	With missing science score	Non-imputed	PROC IMPUTE
TOTAL		16,489	3,858	53.70	52.91
Sex	Female	8,349	1,958	52.09	51.47
	Male	8,140	1,900	55.35	54.39
Race/ ethnicity[1]	White and Asian	12,657	2,778	55.50	54.78
	Black, Hispanic, Indian	3,823	1,077	47.21	46.72
SES[2]	Lowest quartile	4,121	1,159	47.51	46.97
	2nd quartile	4,095	929	51.73	50.95
	3rd quartile	4,147	904	54.87	54.15
	Highest quartile	4,125	865	60.06	59.54

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 16. Comparison of standard deviation (SD) for science theta score before and after imputation

		Number of students		SD science theta score	
		W/nonmissing science score	W/missing science score	Non-imputed	PROC IMPUTE
TOTAL		12,631	3,858	10.64	10.82
Sex	Female	6,391	1,958	10.06	10.21
	Male	6,240	1,900	10.97	11.22
Race/ ethnicity[1]	White and Asian	9,879	2,778	10.23	10.49
	Black, Hispanic, Indian	2,746	1,077	9.50	9.50
SES[2]	Lowest quartile	2,962	1,159	9.27	9.29
	2nd quartile	3,166	929	9.76	9.87
	3rd quartile	3,243	904	9.88	10.13
	Highest quartile	3,260	865	9.52	9.87

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

History/citizenship/ geography Theta Score

We used PROC IMPUTE to impute the 3,917 missing cases for the history/citizenship/geography theta score. We started by using the full BY–F2 panel sample members and the following auxiliary variables:

- from F2—sex, race/ethnicity, SES, units in foreign languages, units in math, units in geometry, units in chemistry, and units in physics;
- from F1—the teacher's opinion about whether the student will go to college or not, number of course the student took in foreign languages, number of course the student took in geometry, and history/citizenship/geography theta score;
- from BY—grade composite variable and history/citizenship/ geography theta score.

We then computed the overall mean and standard deviation for the history/citizenship/geography theta, and also the mean and standard deviation for the history/citizenship/geography theta score across sex, race/ethnicity, and SES quartiles. Those were compared for the following two groups:

1. A group of 12,572 cases in the BY–F2 panel that reported the history/citizenship/geography theta scores; and
2. A group that included the 12,572 cases with actual history/citizenship/geography theta scores and 3,917 cases with imputed scores using PROC IMPUTE.

The mean and standard deviation of the history/citizenship/geography theta score for both groups defined above are shown in tables 17 and 18, respectively.

Table 17. Comparison of mean for history/citizenship/geography theta score before and after imputation

		Number of students		Mean history theta score	
		Overall	With missing history score	Non-imputed	PROC IMPUTE
TOTAL		16,489	3,917	55.41	54.76
Sex	Female	8,349	1,983	54.58	54.08
	Male	8,140	1,934	56.27	55.45
Race/ ethnicity[1]	White and Asian	12,657	2,820	56.71	56.18
	Black, Hispanic, Indian	3,823	1,094	50.74	50.05
SES[2]	Lowest quartile	4,121	1,180	49.72	49.31
	2nd quartile	4,095	943	53.53	53.00
	3rd quartile	4,147	921	56.25	55.68
	Highest quartile	4,125	872	61.56	61.03

[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

Table 18. Comparison of standard deviation (SD) for history/citizenship/geography theta score before and after imputation

		Number of students		SD history theta score	
		W/nonmissing history score	W/missing history score	Non-imputed	PROC IMPUTE
TOTAL		12,572	3,917	9.92	10.09
Sex	Female	6,366	1,983	9.46	9.63
	Male	6,206	1,934	10.30	10.49
Race/ ethnicity[1]	White and Asian	9,837	2,820	9.73	9.90
	Black, Hispanic, Indian	2,729	1,094	9.16	9.23
SES[2]	Lowest quartile	2,941	1,180	8.73	8.89
	2nd quartile	3,152	943	9.09	9.32
	3rd quartile	3,226	921	9.17	9.37
	Highest quartile	3,253	872	8.78	8.97

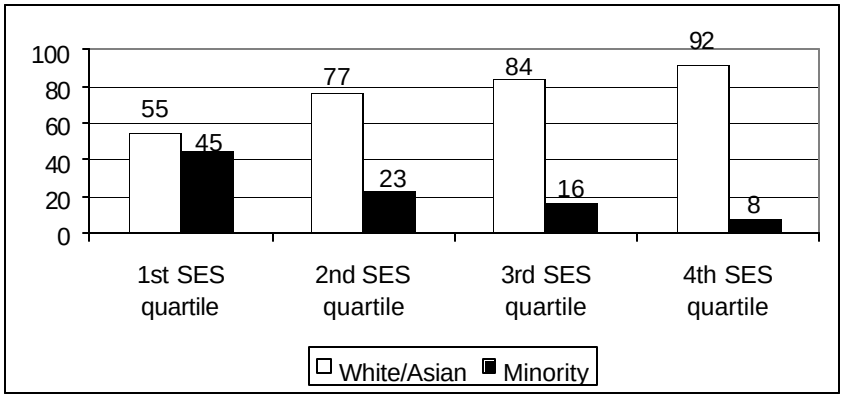
[1] There are 9 cases with missing data on race/ethnicity.
[2] There is 1 case with missing data on SES.
Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), original and imputed data.

CONCLUSION

The SES variable is associated with the race/ethnicity variable (with Pearson chi-squared p-value=0.0001). As seen in figure 3, as the SES quartile increases, the proportion of minorities in that SES quartile decreases. Also the proportion of minorities that have missing values for each subject theta score is higher than the corresponding proportion of minorities that have nonmissing values for that given subject, as shown in figure 4.

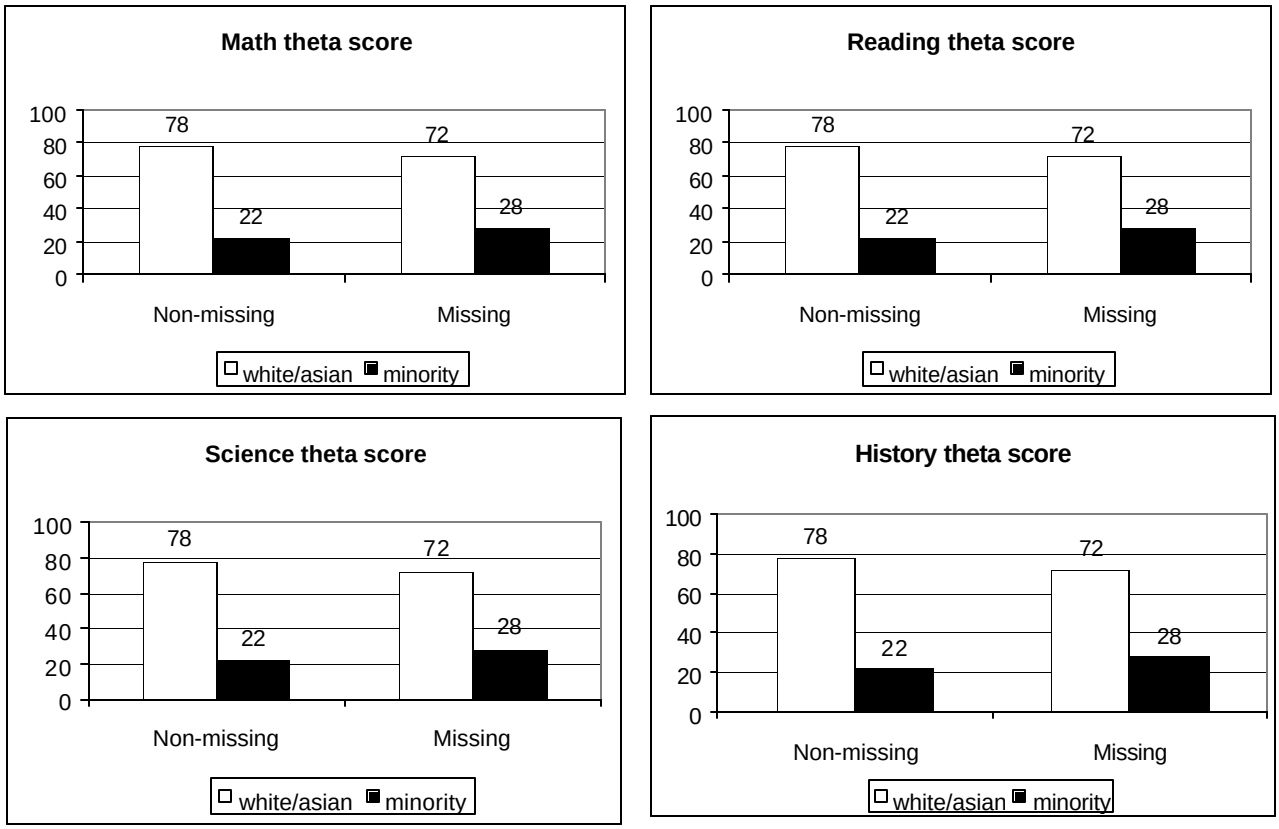
Since the mean theta score increases for each subject as the socioeconomic status quartile increases, we would expect (as is the case in tables 11, 13, 15, and 17) the mean theta score to be slightly lower after imputation than before imputation. That is, the higher proportion of minority students with missing test scores have a slightly lower overall average test score after imputation.

Figure 3. Percentage of racial/ethnic subgroups by socioeconomic status for all F2 panel respondents



Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), imputed data.

Figure 4. Percentage of racial/ethnic subgroups by missing status for all F2 panel respondents



Source: U.S. Department of Education, National Center for Education Statistics, National Education Longitudinal Study of 1988 (NELS:88), imputed data.

REFERENCES

Hu, M., and Salvucci, S. (1999). Evaluation of Some Popular Imputation Algorithms. *1998 Proceedings of the Section on Survey Research Methods* (pp. 308–313). Alexandria, VA: American Statistical Association.

Ingels, S.J., Dowd, K.L., Baldridge, J.D., Stipe, J.L., Bartot, V.H., and Frankel, M.R. (1994). *National Education Longitudinal Study of 1988: Second Follow-Up: Student Component Data File User's Manual* (NCES 94–374). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Little, R.J.A., and Rubin, D.B. (1987). *Statistical Analysis with Missing Data.* New York: John Wiley & Sons, Inc.

McLaughlin, D.H. (1991). *Imputation for Non-Response Adjustment*. Internal Report, American Institute for Research: Palo Alto, California.

Muraki, E.J., and Bock, R.D. (1991). PARSCLE: Parameter scaling of rating data (computer software). Chicago, IL: Scientific Software, Inc.

Rock, D.A., and Pollack, J.M. (1995). *Psychometric Report for the NELS:88 Base Year Through Second Follow-Up* (NCES 95–382), U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Rubin, D.B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association 91*: 473–489.

Listing of NCES Working Papers to Date

Working papers can be downloaded as pdf files from the NCES Electronic Catalog (http://nces.ed.gov/pubsearch/). You can also contact Sheilah Jupiter at (202) 502–7444 (sheilah_jupiter@ed.gov) if you are interested in any of the following papers.

Listing of NCES Working Papers by Program Area

No.	Title	NCES contact
Baccalaureate and Beyond (B&B)		
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Beginning Postsecondary Students (BPS) Longitudinal Study		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper
Common Core of Data (CCD)		
95–12	Rural Education Data User's Guide	Samuel Peng
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97–15	Customer Service Survey: Common Core of Data Coordinators	Lee Hoffman
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
2000–12	Coverage Evaluation of the 1994–95 Common Core of Data: Public Elementary/Secondary School Universe Survey	Beth Young
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2001–09	An Assessment of the Accuracy of CCD Data: A Comparison of 1988, 1989, and 1990 CCD Data with 1990–91 SASS Data	John Sietsema
2001–14	Evaluation of the Common Core of Data (CCD) Finance Data Imputations	Frank Johnson
Data Development		
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Decennial Census School District Project		
95–12	Rural Education Data User's Guide	Samuel Peng
96–04	Census Mapping Project/School District Data Book	Tai Phan
98–07	Decennial Census School District Project Planning Report	Tai Phan
2001–12	Customer Feedback on the 1990 Census Mapping Project	Dan Kasprzyk
Early Childhood Longitudinal Study (ECLS)		
96–08	How Accurate are Teacher Judgments of Students' Academic Performance?	Jerry West
96–18	Assessment of Social Competence, Adaptive Behaviors, and Approaches to Learning with Young Children	Jerry West
97–24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97–36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West

No.	Title	NCES contact
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
2001–03	Measures of Socio-Emotional Development in Middle Childhood	Elvira Hausken
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
Education Finance Statistics Center (EDFIN)		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
High School and Beyond (HS&B)		
95–12	Rural Education Data User's Guide	Samuel Peng
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
HS Transcript Studies		
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
International Adult Literacy Survey (IALS)		
97–33	Adult Literacy: An International Perspective	Marilyn Binkley
Integrated Postsecondary Education Data System (IPEDS)		
97–27	Pilot Test of IPEDS Finance Survey	Peter Stowe
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–14	IPEDS Finance Data Comparisons Under the 1997 Financial Accounting Standards for Private, Not-for-Profit Institutes: A Concept Paper	Peter Stowe
National Assessment of Adult Literacy (NAAL)		
98–17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999–09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999–09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999–09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999–09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999–09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999–09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999–09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek
2000–05	Secondary Statistical Modeling With the National Assessment of Adult Literacy: Implications for the Design of the Background Questionnaire	Sheida White
2000–06	Using Telephone and Mail Surveys as a Supplement or Alternative to Door-to-Door Surveys in the Assessment of Adult Literacy	Sheida White
2000–07	"How Much Literacy is Enough?" Issues in Defining and Reporting Performance Standards for the National Assessment of Adult Literacy	Sheida White
2000–08	Evaluation of the 1992 NALS Background Survey Questionnaire: An Analysis of Uses with Recommendations for Revisions	Sheida White
2000–09	Demographic Changes and Literacy Development in a Decade	Sheida White
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White

No.	Title	NCES contact
National Assessment of Educational Progress (NAEP)		
95–12	Rural Education Data User's Guide	Samuel Peng
97–29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Steven Gorman
97–30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Steven Gorman
97–31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Steven Gorman
97–32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questionnaires)	Steven Gorman
97–37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Steven Gorman
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
National Education Longitudinal Study of 1988 (NELS:88)		
95–04	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95–05	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings
95–06	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95–07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95–12	Rural Education Data User's Guide	Samuel Peng
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98–06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
2001–16	Imputation of Test Scores in the National Education Longitudinal Study of 1988	Ralph Lee
National Household Education Survey (NHES)		
95–12	Rural Education Data User's Guide	Samuel Peng
96–13	Estimation of Response Bias in the NHES:95 Adult Education Survey	Steven Kaufman
96–14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96–21	1993 National Household Education Survey (NHES:93) Questionnaires: Screener, School Readiness, and School Safety and Discipline	Kathryn Chandler
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler

No.	Title	NCES contact
96–29	Undercoverage Bias in Estimates of Characteristics of Adults and 0- to 2-Year-Olds in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
96–30	Comparison of Estimates from the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97–02	Telephone Coverage Bias and Recorded Interviews in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–03	1991 and 1995 National Household Education Survey Questionnaires: NHES:91 Screener, NHES:91 Adult Education, NHES:95 Basic Screener, and NHES:95 Adult Education	Kathryn Chandler
97–04	Design, Data Collection, Monitoring, Interview Administration Time, and Data Editing in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–05	Unit and Item Response, Weighting, and Imputation Procedures in the 1993 National Household Education Survey (NHES:93)	Kathryn Chandler
97–06	Unit and Item Response, Weighting, and Imputation Procedures in the 1995 National Household Education Survey (NHES:95)	Kathryn Chandler
97–08	Design, Data Collection, Interview Timing, and Data Editing in the 1995 National Household Education Survey	Kathryn Chandler
97–19	National Household Education Survey of 1995: Adult Education Course Coding Manual	Peter Stowe
97–20	National Household Education Survey of 1995: Adult Education Course Code Merge Files User's Guide	Peter Stowe
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
97–28	Comparison of Estimates in the 1996 National Household Education Survey	Kathryn Chandler
97–34	Comparison of Estimates from the 1993 National Household Education Survey	Kathryn Chandler
97–35	Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey	Kathryn Chandler
97–38	Reinterview Results for the Parent and Youth Components of the 1996 National Household Education Survey	Kathryn Chandler
97–39	Undercoverage Bias in Estimates of Characteristics of Households and Adults in the 1996 National Household Education Survey	Kathryn Chandler
97–40	Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1996 National Household Education Survey	Kathryn Chandler
98–03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
National Longitudinal Study of the High School Class of 1972 (NLS-72)		
95–12	Rural Education Data User's Guide	Samuel Peng
National Postsecondary Student Aid Study (NPSAS)		
96–17	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
2000–17	National Postsecondary Student Aid Study:2000 Field Test Methodology Report	Andrew G. Malizio
National Study of Postsecondary Faculty (NSOPF)		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
Postsecondary Education Descriptive Analysis Reports (PEDAR)		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Private School Universe Survey (PSS)		
95–16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95–17	Estimates of Expenditures for Private K–12 Schools	Stephen Broughman
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
96–26	Improving the Coverage of Private Elementary-Secondary Schools	Steven Kaufman
96–27	Intersurvey Consistency in NCES Private School Surveys for 1993–94	Steven Kaufman
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman

No.	Title	NCES contact
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Recent College Graduates (RCG)		
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
Schools and Staffing Survey (SASS)		
94–01	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94–02	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94–03	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94–04	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94–06	Six Papers on Teachers from the 1990–91 Schools and Staffing Survey and Other Related Surveys	Dan Kasprzyk
95–01	Schools and Staffing Survey: 1994 Papers Presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95–02	QED Estimates of the 1990–91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95–03	Schools and Staffing Survey: 1990–91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk
95–08	CCD Adjustment to the 1990–91 SASS: A Comparison of Estimates	Dan Kasprzyk
95–09	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95–10	The Results of the 1991–92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95–12	Rural Education Data User's Guide	Samuel Peng
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
95–15	Classroom Instructional Processes: A Review of Existing Measurement Approaches and Their Applicability for the Teacher Follow-up Survey	Sharon Bobbitt
95–16	Intersurvey Consistency in NCES Private School Surveys	Steven Kaufman
95–18	An Agenda for Research on Teachers and Schools: Revisiting NCES' Schools and Staffing Survey	Dan Kasprzyk
96–01	Methodological Issues in the Study of Teachers' Careers: Critical Features of a Truly Longitudinal Study	Dan Kasprzyk
96–02	Schools and Staffing Survey (SASS): 1995 Selected papers presented at the 1995 Meeting of the American Statistical Association	Dan Kasprzyk
96–05	Cognitive Research on the Teacher Listing Form for the Schools and Staffing Survey	Dan Kasprzyk
96–06	The Schools and Staffing Survey (SASS) for 1998–99: Design Recommendations to Inform Broad Education Policy	Dan Kasprzyk
96–07	Should SASS Measure Instructional Processes and Teacher Effectiveness?	Dan Kasprzyk
96–09	Making Data Relevant for Policy Discussions: Redesigning the School Administrator Questionnaire for the 1998–99 SASS	Dan Kasprzyk
96–10	1998–99 Schools and Staffing Survey: Issues Related to Survey Depth	Dan Kasprzyk
96–11	Towards an Organizational Database on America's Schools: A Proposal for the Future of SASS, with comments on School Reform, Governance, and Finance	Dan Kasprzyk
96–12	Predictors of Retention, Transfer, and Attrition of Special and General Education Teachers: Data from the 1989 Teacher Followup Survey	Dan Kasprzyk
96–15	Nested Structures: District-Level Data in the Schools and Staffing Survey	Dan Kasprzyk
96–23	Linking Student Data to SASS: Why, When, How	Dan Kasprzyk
96–24	National Assessments of Teacher Quality	Dan Kasprzyk
96–25	Measures of Inservice Professional Development: Suggested Items for the 1998–1999 Schools and Staffing Survey	Dan Kasprzyk
96–28	Student Learning, Teaching Quality, and Professional Development: Theoretical Linkages, Current Measurement, and Recommendations for Future Data Collection	Mary Rollefson
97–01	Selected Papers on Education Surveys: Papers Presented at the 1996 Meeting of the American Statistical Association	Dan Kasprzyk

No.	Title	NCES contact
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
97–10	Report of Cognitive Research on the Public and Private School Teacher Questionnaires for the Schools and Staffing Survey 1993–94 School Year	Dan Kasprzyk
97–11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97–12	Measuring School Reform: Recommendations for Future SASS Data Collection	Mary Rollefson
97–14	Optimal Choice of Periodicities for the Schools and Staffing Survey: Modeling and Analysis	Steven Kaufman
97–18	Improving the Mail Return Rates of SASS Surveys: A Review of the Literature	Steven Kaufman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
97–23	Further Cognitive Research on the Schools and Staffing Survey (SASS) Teacher Listing Form	Dan Kasprzyk
97–41	Selected Papers on the Schools and Staffing Survey: Papers Presented at the 1997 Meeting of the American Statistical Association	Steve Kaufman
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98–02	Response Variance in the 1993–94 Schools and Staffing Survey: A Reinterview Report	Steven Kaufman
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
98–05	SASS Documentation: 1993–94 SASS Student Sampling Problems; Solutions for Determining the Numerators for the SASS Private School (3B) Second-Stage Factors	Steven Kaufman
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
98–12	A Bootstrap Variance Estimator for Systematic PPS Sampling	Steven Kaufman
98–13	Response Variance in the 1994–95 Teacher Follow-up Survey	Steven Kaufman
98–14	Variance Estimation of Imputed Survey Data	Steven Kaufman
98–15	Development of a Prototype System for Accessing Linked NCES Data	Steven Kaufman
98–16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999–02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Fieldtest Results to Improve Item Construction	Dan Kasprzyk
1999–10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk
1999–12	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume III: Public-Use Codebook	Kerry Gruber
1999–13	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber
1999–14	1994–95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber
1999–17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
Third International Mathematics and Science Study (TIMSS)		
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein

Listing of NCES Working Papers by Subject

No.	Title	NCES contact
Achievement (student) - mathematics		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Adult education		
96–14	The 1995 National Household Education Survey: Reinterview Results for the Adult Education Component	Steven Kaufman
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
98–03	Adult Education in the 1990s: A Report on the 1991 National Household Education Survey	Peter Stowe
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Adult literacy—see Literacy of adults		
American Indian – education		
1999–13	1993–94 Schools and Staffing Survey: Data File User's Manual, Volume IV: Bureau of Indian Affairs (BIA) Restricted-Use Codebook	Kerry Gruber
Assessment/achievement		
95–12	Rural Education Data User's Guide	Samuel Peng
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
97–29	Can State Assessment Data be Used to Reduce State NAEP Sample Sizes?	Larry Ogle
97–30	ACT's NAEP Redesign Project: Assessment Design is the Key to Useful and Stable Assessment Results	Larry Ogle
97–31	NAEP Reconfigured: An Integrated Redesign of the National Assessment of Educational Progress	Larry Ogle
97–32	Innovative Solutions to Intractable Large Scale Assessment (Problem 2: Background Questions)	Larry Ogle
97–37	Optimal Rating Procedures and Methodology for NAEP Open-ended Items	Larry Ogle
97–44	Development of a SASS 1993–94 School-Level Student Achievement Subfile: Using State Assessments and State NAEP, Feasibility Study	Michael Ross
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Beginning students in postsecondary education		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper

No.	Title	NCES contact
Civic participation		
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
Climate of schools		
95–14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng
Cost of education indices		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
Course-taking		
95–12	Rural Education Data User's Guide	Samuel Peng
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson
Crime		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Curriculum		
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
Customer service		
1999–10	What Users Say About Schools and Staffing Survey Publications	Dan Kasprzyk
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2001–12	Customer Feedback on the 1990 Census Mapping Project	Dan Kasprzyk
Data quality		
97–13	Improving Data Quality in NCES: Database-to-Report Process	Susan Ahmed
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Data warehouse		
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
Design effects		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
Dropout rates, high school		
95–07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
Early childhood education		
96–20	1991 National Household Education Survey (NHES:91) Questionnaires: Screener, Early Childhood Education, and Adult Education	Kathryn Chandler

No.	Title	NCES contact
96–22	1995 National Household Education Survey (NHES:95) Questionnaires: Screener, Early Childhood Program Participation, and Adult Education	Kathryn Chandler
97–24	Formulating a Design for the ECLS: A Review of Longitudinal Studies	Jerry West
97–36	Measuring the Quality of Program Environments in Head Start and Other Early Childhood Programs: A Review and Recommendations for Future Research	Jerry West
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
2001–03	Measures of Socio-Emotional Development in Middle School	Elvira Hausken
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
Educational attainment		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Educational research		
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
Eighth-graders		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Employment		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
Employment – after college		
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Engineering		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Enrollment – after college		
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Faculty – higher education		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
Fathers – role in education		
2001–02	Measuring Father Involvement in Young Children's Lives: Recommendations for a Fatherhood Module for the ECLS-B	Jerry West
Finance – elementary and secondary schools		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
96–19	Assessment and Analysis of School-Level Expenditures	William J. Fowler, Jr.
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman

No.	Title	NCES contact
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
2001–14	Evaluation of the Common Core of Data (CCD) Finance Data Imputations	Frank Johnson
Finance – postsecondary		
97–27	Pilot Test of IPEDS Finance Survey	Peter Stowe
2000–14	IPEDS Finance Data Comparisons Under the 1997 Financial Accounting Standards for Private, Not-for-Profit Institutes: A Concept Paper	Peter Stowe
Finance – private schools		
95–17	Estimates of Expenditures for Private K–12 Schools	Stephen Broughman
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Geography		
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
Graduate students		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
Graduates of postsecondary education		
2001–15	Baccalaureate and Beyond Longitudinal Study: 2000/01 Follow-Up Field Test Methodology Report	Andrew G. Malizio
Imputation		
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meeting	Dan Kasprzyk
2001–10	Comparison of Proc Impute and Schafer's Multiple Imputation Software	Sam Peng
2001–14	Evaluation of the Common Core of Data (CCD) Finance Data Imputations	Frank Johnson
2001–16	Imputation of Test Scores in the National Education Longitudinal Study of 1988	Ralph Lee
Inflation		
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
Institution data		
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
Instructional resources and practices		
95–11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk
International comparisons		
97–11	International Comparisons of Inservice Professional Development	Dan Kasprzyk
97–16	International Education Expenditure Comparability Study: Final Report, Volume I	Shelley Burns
97–17	International Education Expenditure Comparability Study: Final Report, Volume II, Quantitative Analysis of Expenditure Comparability	Shelley Burns
2001–01	Cross-National Variation in Educational Preparation for Adulthood: From Early Adolescence to Young Adulthood	Elvira Hausken
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein

No.	Title	NCES contact
International comparisons – math and science achievement		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Libraries		
94–07	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
Limited English Proficiency		
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Literacy of adults		
98–17	Developing the National Assessment of Adult Literacy: Recommendations from Stakeholders	Sheida White
1999–09a	1992 National Adult Literacy Survey: An Overview	Alex Sedlacek
1999–09b	1992 National Adult Literacy Survey: Sample Design	Alex Sedlacek
1999–09c	1992 National Adult Literacy Survey: Weighting and Population Estimates	Alex Sedlacek
1999–09d	1992 National Adult Literacy Survey: Development of the Survey Instruments	Alex Sedlacek
1999–09e	1992 National Adult Literacy Survey: Scaling and Proficiency Estimates	Alex Sedlacek
1999–09f	1992 National Adult Literacy Survey: Interpreting the Adult Literacy Scales and Literacy Levels	Alex Sedlacek
1999–09g	1992 National Adult Literacy Survey: Literacy Levels and the Response Probability Convention	Alex Sedlacek
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–05	Secondary Statistical Modeling With the National Assessment of Adult Literacy: Implications for the Design of the Background Questionnaire	Sheida White
2000–06	Using Telephone and Mail Surveys as a Supplement or Alternative to Door-to-Door Surveys in the Assessment of Adult Literacy	Sheida White
2000–07	"How Much Literacy is Enough?" Issues in Defining and Reporting Performance Standards for the National Assessment of Adult Literacy	Sheida White
2000–08	Evaluation of the 1992 NALS Background Survey Questionnaire: An Analysis of Uses with Recommendations for Revisions	Sheida White
2000–09	Demographic Changes and Literacy Development in a Decade	Sheida White
2001–08	Assessing the Lexile Framework: Results of a Panel Meeting	Sheida White
Literacy of adults – international		
97–33	Adult Literacy: An International Perspective	Marilyn Binkley
Mathematics		
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
1999–08	Measuring Classroom Instructional Processes: Using Survey and Case Study Field Test Results to Improve Item Construction	Dan Kasprzyk
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
Parental involvement in education		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings

No.	Title	NCES contact
97–25	1996 National Household Education Survey (NHES:96) Questionnaires: Screener/Household and Library, Parent and Family Involvement in Education and Civic Involvement, Youth Civic Involvement, and Adult Civic Involvement	Kathryn Chandler
1999–01	A Birth Cohort Study: Conceptual and Design Considerations and Rationale	Jerry West
2001–06	Papers from the Early Childhood Longitudinal Studies Program: Presented at the 2001 AERA and SRCD Meetings	Jerry West
Participation rates		
98–10	Adult Education Participation Decisions and Barriers: Review of Conceptual Frameworks and Empirical Studies	Peter Stowe
Postsecondary education		
1999–11	Data Sources on Lifelong Learning Available from the National Center for Education Statistics	Lisa Hudson
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Postsecondary education – persistence and attainment		
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
Postsecondary education – staff		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
Principals		
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
Private schools		
96–16	Strategies for Collecting Finance Data from Private Schools	Stephen Broughman
97–07	The Determinants of Per-Pupil Expenditures in Private Elementary and Secondary Schools: An Exploratory Analysis	Stephen Broughman
97–22	Collection of Private School Finance Data: Development of a Questionnaire	Stephen Broughman
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber
2000–15	Feasibility Report: School-Level Finance Pretest, Private School Questionnaire	Stephen Broughman
Projections of education statistics		
1999–15	Projected Postsecondary Outcomes of 1992 High School Graduates	Aurora D'Amico
Public school finance		
1999–16	Measuring Resources in Education: From Accounting to the Resource Cost Model Approach	William J. Fowler, Jr.
2000–18	Feasibility Report: School-Level Finance Pretest, Public School District Questionnaire	Stephen Broughman
Public schools		
97–43	Measuring Inflation in Public School Costs	William J. Fowler, Jr.
98–01	Collection of Public School Expenditure Data: Development of a Questionnaire	Stephen Broughman
98–04	Geographic Variations in Public Schools' Costs	William J. Fowler, Jr.
1999–02	Tracking Secondary Use of the Schools and Staffing Survey Data: Preliminary Results	Dan Kasprzyk
2000–12	Coverage Evaluation of the 1994–95 Public Elementary/Secondary School Universe Survey	Beth Young
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber

No.	Title	NCES contact
Public schools – secondary		
98–09	High School Curriculum Structure: Effects on Coursetaking and Achievement in Mathematics for High School Graduates—An Examination of Data from the National Education Longitudinal Study of 1988	Jeffrey Owings
Reform, educational		
96–03	National Education Longitudinal Study of 1988 (NELS:88) Research Framework and Issues	Jeffrey Owings
Response rates		
98–02	Response Variance in the 1993–94 Schools and Staffing Survey: A Reinterview Report	Steven Kaufman
School districts		
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
School districts, public		
98–07	Decennial Census School District Project Planning Report	Tai Phan
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
School districts, public – demographics of		
96–04	Census Mapping Project/School District Data Book	Tai Phan
Schools		
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
Schools – safety and discipline		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Science		
2000–11	Financial Aid Profile of Graduate Students in Science and Engineering	Aurora D'Amico
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
Software evaluation		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
Staff		
97–42	Improving the Measurement of Staffing Resources at the School Level: The Development of Recommendations for NCES for the Schools and Staffing Survey (SASS)	Mary Rollefson
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
Staff – higher education institutions		
97–26	Strategies for Improving Accuracy of Postsecondary Faculty Lists	Linda Zimbler
Staff – nonprofessional		
2000–13	Non-professional Staff in the Schools and Staffing Survey (SASS) and Common Core of Data (CCD)	Kerry Gruber

No.	Title	NCES contact
State		
1999–03	Evaluation of the 1996–97 Nonfiscal Common Core of Data Surveys Data Collection, Processing, and Editing Cycle	Beth Young
Statistical methodology		
97–21	Statistics for Policymakers or Everything You Wanted to Know About Statistics But Thought You Could Never Understand	Susan Ahmed
Statistical standards and methodology		
2001–05	Using TIMSS to Analyze Correlates of Performance Variation in Mathematics	Patrick Gonzales
Students with disabilities		
95–13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Survey methodology		
96–17	National Postsecondary Student Aid Study: 1996 Field Test Methodology Report	Andrew G. Malizio
97–15	Customer Service Survey: Common Core of Data Coordinators	Lee Hoffman
97–35	Design, Data Collection, Interview Administration Time, and Data Editing in the 1996 National Household Education Survey	Kathryn Chandler
98–06	National Education Longitudinal Study of 1988 (NELS:88) Base Year through Second Follow-Up: Final Methodology Report	Ralph Lee
98–11	Beginning Postsecondary Students Longitudinal Study First Follow-up (BPS:96–98) Field Test Report	Aurora D'Amico
98–16	A Feasibility Study of Longitudinal Design for Schools and Staffing Survey	Stephen Broughman
1999–07	Collection of Resource and Expenditure Data on the Schools and Staffing Survey	Stephen Broughman
1999–17	Secondary Use of the Schools and Staffing Survey Data	Susan Wiley
2000–01	1999 National Study of Postsecondary Faculty (NSOPF:99) Field Test Report	Linda Zimbler
2000–02	Coordinating NCES Surveys: Options, Issues, Challenges, and Next Steps	Valena Plisko
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
2000–12	Coverage Evaluation of the 1994–95 Public Elementary/Secondary School Universe Survey	Beth Young
2000–17	National Postsecondary Student Aid Study:2000 Field Test Methodology Report	Andrew G. Malizio
2001–04	Beginning Postsecondary Students Longitudinal Study: 1996–2001 (BPS:1996/2001) Field Test Methodology Report	Paula Knepper
2001–07	A Comparison of the National Assessment of Educational Progress (NAEP), the Third International Mathematics and Science Study Repeat (TIMSS-R), and the Programme for International Student Assessment (PISA)	Arnold Goldstein
2001–09	An Assessment of the Accuracy of CCD Data: A Comparison of 1988, 1989, and 1990 CCD Data with 1990–91 SASS Data	John Sietsema
2001–11	Impact of Selected Background Variables on Students' NAEP Math Performance	Arnold Goldstein
2001–13	The Effects of Accommodations on the Assessment of LEP Students in NAEP	Arnold Goldstein
Teachers		
98–13	Response Variance in the 1994–95 Teacher Follow-up Survey	Steven Kaufman
1999–14	1994–95 Teacher Followup Survey: Data File User's Manual, Restricted-Use Codebook	Kerry Gruber
2000–10	A Research Agenda for the 1999–2000 Schools and Staffing Survey	Dan Kasprzyk
Teachers – instructional practices of		
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
Teachers – opinions regarding safety		
98–08	The Redesign of the Schools and Staffing Survey for 1999–2000: A Position Paper	Dan Kasprzyk
Teachers – performance evaluations		
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk

No.	Title	NCES contact
Teachers – qualifications of		
1999–04	Measuring Teacher Qualifications	Dan Kasprzyk
Teachers – salaries of		
94–05	Cost-of-Education Differentials Across the States	William J. Fowler, Jr.
Training		
2000–16a	Lifelong Learning NCES Task Force: Final Report Volume I	Lisa Hudson
2000–16b	Lifelong Learning NCES Task Force: Final Report Volume II	Lisa Hudson
Variance estimation		
2000–03	Strengths and Limitations of Using SUDAAN, Stata, and WesVarPC for Computing Variances from NCES Data Sets	Ralph Lee
2000–04	Selected Papers on Education Surveys: Papers Presented at the 1998 and 1999 ASA and 1999 AAPOR Meetings	Dan Kasprzyk
Violence		
97–09	Status of Data on Crime and Violence in Schools: Final Report	Lee Hoffman
Vocational education		
95–12	Rural Education Data User's Guide	Samuel Peng
1999–05	Procedures Guide for Transcript Studies	Dawn Nelson
1999–06	1998 Revision of the Secondary School Taxonomy	Dawn Nelson