The 2005 High School Transcript Study User's Guide and Technical Report

NCES 2009-480
U.S. Department of eDUca tion

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

U.S. Department of Education
Institute of Education Sciences
NCES 2009–480

The 2005 High School Transcript Study User's Guide and Technical Report

November 2008

Authors

Carolyn Shettle
Michele Cubell
Katylee Hoover
David Kastberg
Stan Legum
Marsha Lyons
Robert Perkins
Lou Rizzo
Stephen Roey
Diane Sickles
Westat

Janis Brown
Project Director
National Center for Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
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. Unless specifically noted, all information contained herein is in the public domain.

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
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006-5651

November 2008

The NCES World Wide Web Home Page address is http://nces.ed.gov.
The NCES World Wide Web Electronic Catalog is http://nces.ed.gov/pubsearch.

Suggested Citation:
Shettle, C., Cubell, M., Hoover, K., Kastberg, D., Legum, S., Lyons, M., Perkins, R., Rizzo, L., Roey, S., and Sickles, D. (2008). *The 2005 High School Transcript Study: The 2005 High School Transcript Study User's Guide and Technical Report.* (NCES 2009–480). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

For ordering information on this report, write to

U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794-1398

or call toll free 1-877-4ED-Pubs or order online at http://www.edpubs.org.

Content Contact:

Janis Brown
(202) 502-7482
Janis.Brown@ed.gov

TABLE OF CONTENTS

Chapter			Page
1	INTRODUCTION		1
	1.1	Overview of the High School Transcript Study	1
	1.2	Relationship of the HSTS 2005 and NAEP 2005	3
	1.3	Contextual Background Data Provided by HSTS 2005	4
		1.3.1 Participation and Confidentiality of Data	5
	1.4	Classification of Secondary School Courses (CSSC)	6
		1.4.1 Adding and Deleting CSSC Codes	7
	1.5	Comparing HSTS 2005 Results to Other Transcript Studies	8
2	SAMPLE DESIGN FOR THE NAEP 2005 12TH-GRADE ASSESSMENT		9
	2.1	Overview of the 12th-grade Sample Design for NAEP 2005	9
	2.2	Stratification	10
	2.3	Selection of Substitute Schools	11
	2.4	Assignment of Sessions and Sample Type to Schools for NAEP and Student Selection	11
	2.5	Students Not Included in the Assessment	12
3	SAMPLING OR SCHOOLS AND GRADUATES FOR THE NAEP HSTS 2005		13
	3.1	Overview of Sample Design for NAEP HSTS 2005 Sample	13
	3.2	Sampling of Schools	13
	3.3	Sampling of Students	15
	3.4	School Response Rates	15
	3.5	Response Rate for Graduates	19
4	DATA COLLECTION PROCEDURES		25
	4.1	Training NAEP 2005 Field Supervisors as Data Collectors	25
	4.2	Contacts with States, Districts, and Schools	26
	4.3	Obtaining Course Catalogs, Sample Transcripts, and Other School-level Information	27
		4.3.1 Catalogs	28
		4.3.2 Sample Transcripts	28
		4.3.3 SIF and HSTS Questionnaire	29
		4.3.4 School Background Questionnaire	29

TABLE OF CONTENTS (Continued)

Chapter				Page
		4.3.5	SD and LEP Questionnaires	30
	4.4		Identifying the Sample of Students and Obtaining Transcripts	30
		4.4.1	Materials from NAEP 2005 Schools	30
		4.4.2	Schools without NAEP 2005 Materials	33
	4.5		Sending Data for Processing	33
5	DATA PROCESSING PROCEDURES			35
	5.1		NAEP 2005 Questionnaires	35
	5.2		Data Collected Other Than NAEP 2005 Questionnaires	35
		5.2.1	Westat Transcript Entry System (WesTes)	35
		5.2.2	Receipt Control	37
		5.2.3	Catalog Coding	38
		5.2.4	Coding Transcripts	39
		5.2.5	Quality Control Procedures	43
6	WEIGHTING AND ESTIMATION OF SAMPLING VARIANCE			47
	6.1		Overview of Weighting in HSTS 2005	47
	6.2		School Weights	47
		6.2.1	School Base Weights	48
		6.2.2	School Nonresponse Adjustments for HSTS 2005	48
		6.2.3	School Base Weights Adjusted for Nonresponse	49
	6.3		Weights for Graduates	50
		6.3.1	Base Weights	51
		6.3.2	Nonresponse Adjustments	54
		6.3.3	Weight Trimming	56
		6.3.4	Calculating Final Weights for Graduates	57
	6.4		Variance Estimation	58
		6.4.1	Jackknife (JK2) Replication Method	58
		6.4.2	Calculating Replicate Weights	59

TABLE OF CONTENTS (Continued)

Chapter			Page
7	GUIDE TO THE DATA FILES AND CODEBOOKS		63
	7.1 Public-use and Restricted-use Data Files		63
		7.1.1 NAEP Transcript Data Explorer	63
		7.1.2 Restricted-use Data	64
	7.2 Content and Organization of the Restricted-use Files		65
		7.2.1 Course Offerings File	65
		7.2.2 School File	66
		7.2.3 Student File	67
		7.2.4 NAEP Data File	67
		7.2.5 SD/LEP Questionnaire File	69
		7.2.6 Test and Honors File	69
		7.2.7 Transcript File	70
		7.2.8 Master CSSC File	70
	7.3 Additional Information for Researchers Wishing to Use Restricted-use Files		71
		7.3.1 Selecting the Proper Weights	71
		7.3.2 Statistical Software for Use with HSTS 2005	72
		7.3.3 Use of NAEP Scores for Individuals	72
	7.4 HSTS Analysis Reports		73
	REFERENCES		77

List of Appendixes

Appendix		
A	Disclosure Notice	A-1
B	2005 HSTS School Information Form (SIF)	B-1
C	Transcript Request Form (TRF)	C-1
D	NAEP 2005 Administration Schedule	D-1
E	Documentation of Missing Transcripts	E-1

TABLE OF CONTENTS (Continued)

List of Appendixes

Appendix		
F	Summary of School Activities	F-1
G	Letter to NAEP School Coordinator	G-1
H	Transcript Format Checklist	H-1
I	Course Catalog Checklist	I-1
J	Shipping Transmittal Form – Phase 2	J-1
K	School Background Questionnaire	K-1
L	2005 SD Questionnaire	L-1
M	2005 LEP Questionnaire	M-1
N	Quality Checks for the HSTS Sample: Comparison of the Full HSTS School Sample to the Responding HSTS School Sample	N-1
O	2005 High School Transcript Study Classification of Secondary School Courses Hierarchical Listing and Detailed Stub List	O-1
P	2005 High School Transcript Study Codebook for Catalog File	P-1
Q	2005 High School Transcript Study Codebook for Master CSSC File	Q-1
R	2005 High School Transcript Study Codebook for NAEP Data File	R-1
S	2005 High School Transcript Study Codebook for School File	S-1
T	2005 High School Transcript Study Codebook for SD/LEP File	T-1
U	2005 High School Transcript Study Codebook for Student File	U-1
V	2005 High School Transcript Study Codebook for Tests and Honors File	V-1
W	2005 High School Transcript Study Codebook for Transcript File	W-1
X	HSTS 2005 User's Guide and Technical Documentation Glossary	X-1

TABLE OF CONTENTS (Continued)

List of Tables

Table		Page
1	NCES high school transcript studies: Selected years, 1982-2005	2
2	School and student sample sizes for NAEP HSTS 2005 (from the NAEP 2005 12th-grade school sample), by school type: 2005	14
3	Unweighted and weighted response rates for all eligible NAEP HSTS schools and school enrollments, by HSTS status: 2005	16
4	Unweighted and weighted response rates for schools and school enrollments eligible for NAEP HSTS and NAEP, by HSTS status: 2005	16
5	Unweighted and weighted response rates for all eligible NAEP HSTS schools and school enrollments, by school type and HSTS status: 2005	17
6	Unweighted and weighted response rates for schools and school enrollments eligible for NAEP HSTS and NAEP, by school type and school status: 2005	17
7	Weighted HSTS response and eligibility rates for all sampled public schools, by census division: 2005	18
8	Weighted HSTS response and eligibility rates for sampled private schools for NAEP HSTS and NAEP, by school type: 2005	19
9	Weighted HSTS response and eligibility rates for all sampled HSTS private schools, by school type: 2005	19
10	Student samples in HSTS participating schools, by graduation status: 2005	20
11	Graduates from HSTS participating schools, by whether the school participated in NAEP: 2005	20
12	Unweighted and weighted student counts, and student within school response rate, for NAEP HSTS participating schools, by HSTS transcript status: 2005	21
13	Unweighted and weighted students count, and student within-school response rate, for NAEP HSTS participating schools, by school type and HSTS transcript status: 2005	21

TABLE OF CONTENTS (Continued)

List of Tables (continued)

Table		Page
14	Student sample in schools participating in NAEP, by NAEP assessment assignment: 2005	22
15	Unweighted NAEP HSTS student sample, by transcript status and school status: 2005	22
16	Unweighted and weighted student sample, and student within-school response rate, for schools participating in both NAEP HSTS and NAEP, by transcript status: 2005	23
17	Unweighted and weighted student sample and student within-school response rate, for schools participating in both NAEP HSTS and NAEP, by school type and transcript status: 2005	23
18	Unlinked study combined response rates, by school type: 2005	24
19	Linked study combined response rates, by school type: 2005	24
20	HSTS and NAEP record identifier naming conventions: 2005	68

List of Figures

Figure		Page
1	Data flow for the High School Transcript Study 2005	36

1. INTRODUCTION

This technical report documents the procedures used to collect and summarize data from the 2005 High School Transcript Study (HSTS 2005). Chapters detail the sampling of schools and graduates (chapters 2 and 3), data collection procedures (chapter 4), data processing procedures (chapter 5), and weighting procedures (chapter 6). Chapter 7 describes the HSTS 2005 data files and codebooks that are encompassed by this report. Appendices A through J contain the HSTS 2005 data collection and documentation forms, and appendices K through M contain the associated National Assessment of Educational Progress (NAEP) 2005 study questionnaires. Appendix N contains information concerning nonresponse bias associated with creating the HSTS weights. Appendix O describes the Classification of Secondary School Courses (CSSC), which was used to code the courses on the HSTS 2005 transcripts, and provides a complete listing of CSSC codes. The codebooks for all of the HSTS 2005 restricted-use data files are in appendices P through W. A glossary of terms is in appendix X.

This chapter provides an introduction to HSTS 2005. Additional information is contained in later chapters. Initial results are contained in the companion report *The Nation's Report Card: America's High School Graduates: Results from the 2005 NAEP High School Transcript Study* (Shettle et al. 2007), where selected topics are discussed in greater detail.

1.1 Overview of the High School Transcript Study

Over the years, various reform efforts have sought to improve the quality of education across the United States. In the early 1980s, the focus was on statewide curricula in core courses, a response to the watershed report, *A Nation at Risk: The Imperative for Educational Reform* (National Commission on Excellence in Education 1983). Since then, national efforts have addressed several issues concerning quality education, analyzing the content of courses in specific subject areas (e.g., mathematics and science), the number of courses completed, and when courses are completed.

NAEP HSTS is a periodic survey that provides educational professionals, such as administrators, policymakers, and researchers, with information regarding the coursetaking patterns of high school graduates and their grade point averages (GPAs). It can also be used to provide information

on the relationship of graduate coursetaking patterns to achievement as measured by NAEP. NAEP is an ongoing, periodic assessment of educational achievement in U.S. schools.

The transcript studies serve as a barometer for changes in high school graduates' coursetaking patterns. Coursetaking patterns provide valuable information about the rigor of high school curricula followed across the nation. The first national transcript study was conducted by the National Center for Education Statistics (NCES) in 1982 and captured baseline information on high school students' patterns prior to the publication of *A Nation at Risk* and the resulting changes in curricula and educational reform.

For HSTS 2005, complete transcripts for 26,200 graduates from public and private high schools in 2005 were collected from a nationally representative sample of schools from May through October 2005. The survey was conducted in conjunction with NAEP 2005 mathematics and science assessments in the 12th grade. A description of this survey can be found on the NAEP home page at http://nces.ed.gov/nationsreportcard/.

Since similar studies were conducted on the coursetaking patterns of graduates over the years, changes in these patterns can be studied and compared. Table 1 lists the nine NCES studies that have been conducted beginning in 1982 involving the collection of high school transcripts.

Table 1. NCES high school transcript studies: Selected years, 1982-2005

Study	Approximate number of transcripts[1]
1982 High School and Beyond	12,700
1987 NAEP High School Transcript Study	34,100
1990 NAEP High School Transcript Study	21,500
National Education Longitudinal Study of 1988 Second Follow-Up (1992)	17,300
1994 NAEP High School Transcript Study	25,500
1998 NAEP High School Transcript Study	25,000
2000 NAEP High School Transcript Study	21,000
Education Longitudinal Study of 2002 First Follow-Up (2004)	16,400
2005 NAEP High School Transcript Study	27,200

[1] Includes transcripts that were not included in the final reports because they were out of scope.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School and Beyond (HS&B), 1982; National Education Longitudinal Study of 1988 (NELS:88) Second Follow-Up, 1992; Education Longitudinal Study of 2002 (ELS:2002) First Follow-Up, (2004); High School Transcript Study (HSTS), Selected years, 1987-2005.

1.2 Relationship of the HSTS 2005 and NAEP 2005

HSTS is conducted in conjunction with NAEP. HSTS 2005 was designed to allow an analysis of the coursetaking patterns of graduates who graduated from American public and private high schools in 2005. It was further designed so that data on graduates' coursetaking patterns can be linked to the NAEP 2005 assessment results. NAEP provides results about subject-matter achievement, instructional experiences, and school environment, and reports these results for populations of students (e.g., 12th-graders) and selected subgroups of those populations (e.g., male students). Changes in the relationship of HSTS coursetaking to NAEP performance can also be examined for similar studies in 1994, 1998, and 2000.[1]

NAEP provides HSTS with data on assessments in different subjects. For HSTS 2005, the scale scores for mathematics and science were provided.

When schools selected for NAEP do not participate in NAEP, the original NAEP schools are replaced with appropriate substitute schools when feasible. To maintain as many links as possible with NAEP 2005 scores, substitute schools that participated in NAEP 2005 were asked to participate in HSTS. When neither the original NAEP refusal schools nor a NAEP substitute participated in NAEP, the school originally selected for NAEP was asked to participate in HSTS. If this school refused participation in HSTS early in the data collection process, the substitute school was asked to participate. Of the 1,017 eligible schools in the original NAEP sample, 726 schools participated in the HSTS 2005 survey, of which 58 were substitute schools. Of the schools participating in HSTS 2005, there were 677 (93 percent) schools that also participated in NAEP 2005 and retained the information necessary to link HSTS and NAEP.[2]

A total of approximately 29,900 12th grade students were selected for HSTS 2005. Because sampling was performed in most schools prior to graduation, not all sampled students were, in fact, graduates. However, only graduates were eligible for inclusion in the transcript study. Of the students in the original sample, it was determined that approximately 27,200 (91 percent) had graduated by October 2005. Transcripts were received from all but 99 (0.4 percent) of these graduates.

[1] See Legum et al. 1997; Roey et al. 2001b; and Perkins et al. 2004 for information on earlier studies.

[2] The links between the graduates and their IDs are maintained at the schools to preserve the confidentiality of the graduates.

1.3 Contextual Background Data Provided by HSTS 2005

Contextual background data for HSTS 2005 are obtained from the NAEP 2005 questionnaires[3], the high school transcripts, and various school-level forms completed by a school coordinator or counselor.

NAEP 2005 Questionnaires also Completed for Non-NAEP Schools

- The *School Background Questionnaire* contains information about the school, its teachers, and its student body (see appendix K).
- The *Students with Disabilities Questionnaire (SD)* contains information about students classified by their schools as having a disability (see appendix L).
- The *Students with Limited English Proficiency Questionnaire (LEP)* contains information about students classified by their schools as having limited skills in English (see appendix M).

NAEP 2005 Questionnaires not Completed for Non-NAEP Graduates

Students taking NAEP completed *Student Questionnaires* embedded in their test booklets that described their background, demographic characteristics, and educational experiences. Since HSTS does not contact individual graduates, comparable information is not available for graduates that did not participate in NAEP.

Transcripts

The transcripts provided information about the graduate that was coded and entered into the data system by trained personnel. These data included the following:

- date graduate enrolled in high school;
- date graduate graduated;

[3] Additional information about these questionnaires can be found at http://nces.ed.gov/nationsreportcard/bgquest.asp.

- rank in class (where available);
- size of class (where available);
- GPA;
- days absent each year (where available);
- standardized test scores and honors (where available);
- list of courses taken in high school, including the grades received, the number of credits earned for each course, and the grade in which the course was taken; and
- total number of credits received and, in many cases, total number of credits attempted.

School Forms, Catalogs or Course Lists

- *Transcript Request Form (TRF)*: A field worker completed a TRF upon returning to a school to obtain requested graduate transcripts. The form contained graduate demographic data, including Title 1 and National School Lunch Program participation status, as well as the student's graduation status.
- *School Information Form (SIF)*: The completed SIF contained information about the school in general, such as sources of data collection information within the school, course description materials, graduation requirements, and grading practices.
- *School-level Catalog or Course Lists*: These lists contained course titles and descriptions needed to code courses, using the Classification of Secondary School Courses (CSSC).[4]

1.3.1 Participation and Confidentiality of Data

Graduates' transcripts were collected by field workers for the sample of graduates selected for the NAEP 2005 assessment. Unlike NAEP, parental consent is not needed in HSTS, and the schools are provided with information about the Family Educational Rights and Privacy Act (FERPA) that authorizes collection of transcript data without parental consent. Sometimes schools object, and field workers reiterate the FERPA. Generally, schools do not require parental or graduate notification or consent for HSTS because there is no burden placed on the graduate.

[4] See section 4.3.1 for additional information on this process.

The data obtained from the transcript study were kept strictly confidential. Student names and any other identifiable information were masked on the copies of the transcripts before these materials left the schools. Furthermore, in schools that participated in the NAEP assessments, each student received a NAEP ID that was also used in HSTS. The list that linked the student's name with that NAEP ID remained in the school. HSTS staff did not have access to that list and could not recreate it if it were lost.

The restricted-use HSTS 2005 data files do not contain the graduates' names or other variables that directly identify the sampled graduates. Data files do contain the graduates' NAEP IDs, which enable researchers to link the transcript data to the NAEP data. HSTS follows NCES' strict procedures regarding the confidentiality of data files.

1.4 Classification of Secondary School Courses (CSSC)

To compare transcripts from different schools, it was necessary to code each of the courses entered from the transcripts using a common course coding system. The coding system employed for this purpose was a modification of the system presented in *A Classification of Secondary School Courses* (Ludwig et al. 1982). The CSSC, which contains over 2,200 course codes, is a modification of the college course classification system presented in *Classification of Instructional Programs* (Morgan, Hunt, and Carpenter 1991). Both course coding systems use a three-level, six-digit system for classifying courses. The CSSC uses the same first two levels as the Classification of Instructional Programs (CIP), which is represented by the first four digits of each code.[5] The third level of the CSSC (the fifth and sixth digits of the course code) is unique to the CSSC and represents specific high school courses.

A taxonomy of course subject areas was developed for HSTS 1987. This taxonomy, documented in the HSTS 1987 tabulations (Thorne 1988), was developed with an emphasis on academic courses. Computer-related courses were considered as constituting a separate nonvocational subject area, and there were fewer subgroups defined for vocational and personal courses. This taxonomy was applied to data from the High School and Beyond (HS&B) 1982 First Follow-Up Study and the HSTS 1987 data. HSTS 1990 used a slightly expanded version of the same taxonomy in its reports.[6]

[5] Specifically, the CSSC uses the first two levels of the CIP as it existed in 1982. The CIP has undergone some modification since then. In addition, three sets of codes at the top level have been added to the CSSC to provide a means of classifying courses specifically designed for students with disabilities.

[6] The 1990 study added 18 new codes to the CSSC and to the taxonomy. The full taxonomy is documented in Legum et al. 1993a and Legum et al. 1993c.

Starting with the 1994 study, HSTS switched over to the Secondary School Taxonomy (SST). Originally developed in the late 1980s by the National Assessment of Vocational Education,[7] SST has a less purely academic emphasis and a more richly defined group of vocational education categories than the taxonomy developed for the earlier HS&B and HSTS studies. Computer-related courses became vocational courses, and general skills and military science courses became new subject areas. To maintain comparability among the transcript studies, the HSTS 1987 and 1990 studies, along with the HS&B 1982 study, were recoded using the SST.

With more than 2,200 codes in the CSSC, it is often neither practical nor desirable to tabulate estimates of each possible CSSC code. It is typically more useful, however, to analyze the courses in larger subject areas such as English, social studies, mathematics, or science. There is also interest in subgroups of these subject areas, such as biology, chemistry, and physics. The taxonomy presented in appendix O provides the structure for aggregating the courses to subject areas.

1.4.1 Adding and Deleting CSSC Codes

Codes are added to the CSSC whenever courses are found in the catalogs that have no match in the CSSC. Highly trained coders coded the school catalogs received from the field workers. These coders reviewed the catalogs, matching the appropriate CSSC codes to the courses offered, according to the content and description of the course. If a course that was offered did not have a matching CSSC code in the existing list, the coders wrote that course description in a special suggestion list. After the catalogs were reviewed, and all but those courses on the suggestion list were coded, a coding specialist reviewed the suggestion list and tried to match these courses to existing CSSC codes. If a course did not have a matching CSSC code, and if this course also appeared in several other schools, a new CSSC code was generated. If the "new" course was limited to just a few schools, the CSSC code that most closely described the course was assigned.

In 1994, 18 new CSSC codes were added to the list. In 1998, the CSSC's computer science curriculum changed dramatically. New courses such as Web Design, Java Programming, and C++ Programming were added. Many courses that were labeled as honor courses in the past were reclassified as Advanced Placement (AP) courses. Many International Baccalaureate (IB) courses were added as well.

[7] A description of the development of the SST is provided in *The Secondary School Taxonomy Final Report* (Gifford, Hoachlander, and Tuma 1994).

In all, a total of 83 new or revised codes were added to the CSSC in 1998. In 2000, two CSSC codes were added, one in science and one in computer-related studies. In 2005, 18 new codes were added. Five new codes reflected the increase in AP and IB courses available to students. Other courses were added when courses were encountered on the transcripts that were clearly different from codes already contained in the master CSSC list. These courses included leadership, military drill team, teacher training, and computer hardware and repair. No new subject areas were identified in HSTS 2005. Three duplicate and unused codes were dropped in 2005.

1.5 Comparing HSTS 2005 Results to Other Transcript Studies

Between 1982 and 2005, NCES has conducted nine high school transcript studies: the HS&B survey in 1982, the Second Follow-Up to the National Educational Longitudinal Study (NELS:88) in 1992, First Follow-Up to the Education Longitudinal Study of 2002 (ELS:2002) in 2004; and NAEP HSTS in 1987, 1990, 1994, 1998, 2000, and 2005. One research objective of NAEP HSTS 2005 was to study changes in the coursetaking patterns among high school graduates over time, comparing its results with the other NCES-conducted high school transcript studies. While results are reported for trends over time, it should be noted that some differences exist among the high school transcript studies and some direct comparisons are cautioned.

For more information about comparisons among the different HS&B and HSTS studies, please refer to chapter 1 of *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990-2000* (Perkins et al. 2004).[8] For discussion about comparisons with the transcript component of the Second Follow-Up to NELS:88, please refer to appendix A of *National Education Longitudinal Study of 1988, Second Follow-Up: Transcript Component Data File User's Manual* (Ingels et al. 1995). The similarities and differences between the high school transcript studies' data (NAEP, NELS, HS&B) are also described extensively in the *NCES Handbook of Survey Methods* (Thurgood et al. 2003). The handbook looks at the comparability of the high school transcript studies' data based upon five criteria: (1) sample sizes, (2) oversampling of subgroups, (3) eligibility criteria for inclusion in the studies, (4) representativeness of cross-sectional and longitudinal populations, and (5) coding differences.

[8] This report can be found at http://nces.ed.gov/nationsreportcard/.

2. SAMPLE DESIGN FOR THE NAEP 2005 12TH-GRADE ASSESSMENTS

The 2005 High School Transcript Study (HSTS 2005) sample consists of a subsample of 12th-grade schools and students selected for participation in the 2005 National Assessment of Educational Progress (NAEP) operational science and mathematics assessments. This chapter describes aspects of the NAEP 2005 sample design that affect the HSTS 2005 sample. The focus of chapter 3 is on aspects of the selection of schools and students that are specific to HSTS 2005.

All public and private high schools in the United States with one or more graduates in 2005 were eligible for HSTS 2005. Graduates were defined as persons receiving a special education, regular education, or honors diploma. Graduates who were considered ineligible for NAEP (e.g., because of a disability) were considered eligible for HSTS. Eligible graduates with incomplete transcripts were considered nonrespondents. For analyses in which the user wishes to link results of the NAEP assessments with HSTS information, graduates were considered eligible if they met both the HSTS and the NAEP eligibility criteria.

HSTS 2005 used all eligible public schools (i.e., schools with 12th-grade NAEP mathematics and/or science assessments) and a subsample of private schools from the 12th-grade NAEP 2005 assessment. The HSTS 2005 graduate sample consisted of the NAEP 2005 student sample in these subsampled schools.

2.1 Overview of the 12th-grade Sample Design for NAEP 2005

The 12th-grade sample for NAEP 2005 was a two-stage probability-based sample of students.[9] This was a national sample in which schools were the first-stage sampling units selected with probability proportional to a measure of size based on the estimated grade-specific enrollment in the schools. The second stage involved selection of students within schools and their assignment to an assessment subject. In previous NAEP studies, the sample design included an initial sample of primary sampling units (PSUs) from across the nation. For NAEP 2005, the PSU sampling stage was eliminated for operational and statistical reasons.

[9] The procedures for the 2005 NAEP selection differed from what was used in 2000.

As in past assessments, modest oversampling of Black and Hispanic students was undertaken in this sample and was carried out at the school level in order to provide adequate information of these groups for analysis. Each school with both more than 15 percent Black and Hispanic students and 10 or more minority students was considered a high minority school for these purposes and was given twice the selection probability of a low minority school of comparable size. This means that while about 40 percent of the student population (including over 95 percent of the Black and Hispanic students) were in high minority schools, about 60 percent of the sampled students were from these schools.

2.2 Stratification

Sampling was done separately for public and private schools. The grade 12 public sample had an implicit stratification, using a hierarchy of stratifiers and a serpentine sort. The top of the hierarchy was census division (9 implicit strata). The next stratifier in the hierarchy was type of location, which had 8 categories. Of the 72 potential type-of-location strata nested within census divisions, several were collapsed with neighboring type-of-location cells, always within census division, giving a total of 55 to 60 census division-location type strata.

These geographic strata were subdivided into 110 to 120 strata by a dichotomous high minority status category. Schools were in the high minority stratum if they had more than 10 minority eligible students and greater than 15 percent minority eligible students (minority defined as Black or Hispanic). Otherwise the school was put in a low minority stratum. If the expected sample size within these strata was less than 8.0, they were left as is. If the expected sample size was greater than 8.0, then the high or low minority stratum was subdivided into a maximum of four substrata (two for expected sample size up to 12.0, three for expected sample size up to 16.0, and four for expected sample size greater than 16.0). For the low minority strata, the subdivision was by state or groups of contiguous states. For the high minority strata, the subdivision was by minority percentage. In total there were between 160 and 180 implicit strata. Within these substrata, the schools were to be sorted by estimated grade enrollment using a serpentine sort within the school type substrata.

The private schools were explicitly stratified by type of private school (Catholic, Lutheran, Conservative Christian, other private). Within each school type, stratification was by census division (9 categories), type of location (8 categories), and by proportion of minority enrollment, used as a continuous sorting variable. The final number of strata was dependent on the proportion of minority

students (Black/Hispanic/Native American) among those schools within each cell defined by private school type, census division, and type of location. In general, where there were few or no schools in a given stratum, categories were collapsed together.

2.3 Selection of Substitute Schools

Though efforts were made to secure the participation of all schools selected, it was anticipated that not all schools would choose to participate. Therefore, as each school was selected in the sample, the two neighboring schools in the sampling frame (immediately preceding and following it) were designated as replacement schools. If an original school refused to participate, the first replacement was then contacted. If that school also refused to participate, the second school was then contacted. There were several constraints on the assignment of substitutes. One sampled school was not allowed to substitute for another, and a given school could not be assigned to substitute for more than one sampled school.

2.4 Assignment of Sessions and Sample Type to Schools for NAEP and Student Selection

The public school sample at grade 12 was assigned three session types: Operational Reading, Mathematics and Science reporting samples (RS); Mathematics, Civics, History, and Economics pilot tests (PT); and Science Bridge (SB). Most of the sample schools received RS and PT session assignments, with many also receiving SB. Some very small schools received only SB. Up to 144 students were selected within schools. For schools with more than 144 students, a systematic equal probability sample of 135 students was selected. If the school had 54 students or more, 22 percent of the students were assigned to an SB session, with 62 percent of students assigned to an RS session and 16 percent of students assigned to a PT session. If the school had 36 to 53 students, a third of the students were assigned to SB and the rest were split between RS and PT in a 4 to 1 ratio. If the school had 24 to 35 students, half of the students were given SB, 40 percent were assigned RS, and 10 percent were assigned PT. Schools with fewer than 24 students had all students assigned to SB.

For private schools, the assignment was similar. Up to 136 students were selected per school. For schools with more than 136 students, a systematic equal probability sample of 120 students was selected. For schools with 97 or more students, approximately 13 percent of students were assigned SB

and the remainder RS and PT. Students were allocated between the latter two sessions in a ratio of 9 to 1. If the school had 36 to 47 students, one-third of the students received SB and the remaining two-thirds received RS and PT. Schools with 24 to 35 students had half of the students assigned to SB and the other half to RS and PT. If the school had less than 24 students, all students took SB. In all cases, students were allocated between the RS and PT sessions in a 9 to 1 ratio.

2.5 Students Not Included in the Assessment

School staff members were asked to determine whether any of the students identified as having a disability or with limited English language proficiency could not participate in the assessment. They needed to determine if a student could not participate meaningfully, or if the accommodations required for the student to participate were not available. These students were not invited to the assessment and coded as "excluded" to distinguish them from absent students. Although school staff are encouraged to follow NAEP standards regarding which students should be excluded from testing, the final decision is made by school personnel.

From the schools selected in the HSTS school sample, approximately 3 percent of the students were excluded from the NAEP assessment. As the transcript study attempted to collect high school transcripts for all students selected for the assessment, whether or not they participated, transcripts for these students are included in the transcript study.

3. SAMPLING OF SCHOOLS AND GRADUATES FOR THE NAEP HSTS 2005

3.1 Overview of Sample Design for NAEP HSTS 2005 Sample

The sample for High School Transcript Study (HSTS) was designed to achieve a nationally representative sample of public and private school high school graduates in the Class of 2005. The target population for the 2005 national assessments included all graduates in public and private schools who were enrolled in 12th-grade in 2004-05, and who graduated in 2005. The samples were selected based on a two-stage sample design: selection of schools and selection of graduates within schools.

3.2 Sampling of Schools

For public schools, the HSTS sample was, in fact, the National Assessment of Educational Progress (NAEP) 2005 12th-grade public school sample for the operational math and science assessments.[10] All participating NAEP 2005 12th-grade public schools were part of the initial HSTS sample regardless of whether they were original or substitute NAEP schools. If neither the original nor the substitute school selected participated in NAEP, the original school was included in the initial 12th-grade public school sample.

In NAEP 2005, private schools were heavily oversampled to meet explicit target sample sizes for reporting group (Catholic, Lutheran, Conservative Christian, Other Religious, Nonsectarian, and Independent) in order to provide reliable NAEP estimates for such students. In HSTS 2005, however, the oversampling of private schools was reversed so that the private school students in HSTS were represented in proportion to their prevalence in the general 12th-grade student population. Table 2 presents the subsampling rates and the calculations that generated those rates.

Probabilities of selection were determined for each school before the school sample was selected. The final probabilities of selection for the 2005 NAEP HSTS school sample were the products of the 2005 NAEP probabilities of school selection and the conditional probabilities of selection in the

[10] Note that this excludes schools that were sampled for only pilot or bridge studies. Public schools with less than 24 students in 12th grade had a two-ninths chance of being assigned a bridge-only session, and private schools with less than 24 students in 12th grade had a one-eighth chance of being assigned a bridge-only session.

Table 2. School and student sample sizes for NAEP HSTS 2005 (from the NAEP 2005 12th-grade school sample), by school type: 2005

School type	NAEP 2005 12th-grade national student sample size	Percent of NAEP 2005 12th-grade national sample size	National estimated grade enrollment in 12th-grade	Percent of estimated grade enrollment in 12th-grade	Proportional student sample size (making sample size proportional to population)	Percent subsampled to obtain sample size proportional to population $(f)_c$	NAEP 2005 12th-grade national school sample size	NAEP HSTS 2005 school sample size
Total	24,500	100.0	3,325,080	100.0	21,454		1,323	1,024
Public	19,600	80.0	3,037,705	91.4	19,600	100.0	829	829
Catholic	2,450	10.0	143,205	4.3	924	37.7	79	30
Lutheran	245	1.0	5,583	0.2	36	14.7	14	2
Conservative Christian	735	3.0	36,085	1.1	233	31.7	132	42
Other private	1,470	6.0	102,502	3.1	661	45.0	244	110
Unknown	—	—	—	—	—	45.0	25	11

— Not available

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

HSTS sample. The subsampling process (using the designated subsampling rates) was a systematic sample within each private school stratum listed in Table 2. The ordering for this systematic sample was the frame ordering from the NAEP 2005 private school sampling process. Schools with unknown affiliation were sampled at the same rate as "other" private school types (45 percent).

3.3 Sampling of Students

For those HSTS sample schools that cooperated in the NAEP assessment, all graduates who were assessed in the operational mathematics and science assessments and also graduated in 2005 were included in the HSTS sample of graduates within the school. For HSTS sample schools that did not cooperate in the NAEP assessment but agreed to cooperate in HSTS, a subsample of 50 graduates was typically drawn from their 12th-graders who graduated in 2005.[11] If the list contained 50 or fewer graduates, all graduates were selected.

3.4 School Response Rates

Nonresponse is a serious concern in any probability sample, as differential response rates within important subgroups may generate biases that are difficult to measure and control through adjustment. NAEP HSTS 2005 had generally very high response rates, but there are two particular areas of concern. The first area of concern is private schools, where response was low. The second area of concern is the decision by one large state not to participate in HSTS (though it participated in the NAEP 2005 12th-grade reading and mathematics operational assessment). Tables 3 and 4 present response rates for the main HSTS study and for the NAEP-HSTS linked study respectively. The first set of response rates is for the HSTS study as a whole, counting as respondents those who participated in HSTS, regardless of their participation in NAEP. The second set of response rates is for the NAEP-HSTS link study, counting as respondents those schools that participated in both HSTS and NAEP, and where linkage of the NAEP assessment and the HSTS study transcript information is possible.[12]

[11] If there were between 50 and 60 graduates, the school had the option of including up to 60 graduates.

[12] See Appendix N for the nonresponse bias analysis done for HSTS 2005.

Table 3. Unweighted and weighted response rates for all eligible NAEP HSTS schools and school enrollments, by HSTS status: 2005

HSTS status	Unweighted number of schools selected	Unweighted percent of selected schools	Weighted number of schools selected	Weighted percent of selected schools	Weighted enrollment at selected schools	Weighted percent of enrollment at selected schools
Total	891	100.0	24,731	100.0	3,177,283	100.0
Respondent	726	81.5	19,120	77.3	2,675,008	84.2
Eligible nonrespondent	165	18.5	5,610	22.7	502,274	15.8

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 4. Unweighted and weighted response rates for schools and school enrollments eligible for NAEP HSTS and NAEP, by HSTS status: 2005

HSTS status	Unweighted number of schools selected	Unweighted percent of selected schools	Weighted number of schools selected	Weighted percent of selected schools	Weighted enrollment at selected schools	Weighted percent of enrollment at selected schools
Total	891	100.0	24,731	100.0	3,177,283	100.0
Respondent	677	76.0	17,699	71.6	2,525,904	79.5
Eligible nonrespondent	214	24.0	7,032	28.4	651,379	20.5

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

The first set of rates calculated within each table is based on sample counts; the second set of rates is weighted by the school base weight; and the third set of rates is weighted by school base weight multiplied by grade enrollment. The second set of rates shows estimates at the population level with each school counted as a unit. The third set of rates shows estimates at the population level with each school counted by its number of enrolled students. In calculating these rates, the 133 high schools that were ineligible for HSTS 2005 because they did not have any graduating students are excluded.

Tables 5 and 6 present aggregate school counts by school type by HSTS status for the unlinked and linked studies respectively. The unweighted counts are based on the number of eligible schools in the sample. The weighted school counts are weighted by the school base weights (i.e., the

Table 5. Unweighted and weighted response rates for all eligible NAEP HSTS schools and school enrollments, by school type and HSTS status: 2005

School type and HSTS status	Unweighted number of schools selected	Unweighted percent of selected schools	Weighted number of schools selected	Weighted percent of selected schools	Weighted enrollment at selected schools	Weighted percent of enrollment at selected schools
Total	891	100.0	24,731	100.0	3,177,283	100.0
Public						
Total	744	100.0	17,968	100.0	2,911,954	100.0
Respondent	643	86.4	15,712	87.4	2,510,485	86.2
Eligible nonrespondent	101	13.6	2,256	12.6	401,469	13.8
Private						
Total	147	100.0	6,763	100.0	265,328	100.0
Respondent	83	56.5	3,408	50.4	164,523	62.0
Eligible nonrespondent	64	43.5	3,355	49.6	100,805	38.0

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 6. Unweighted and weighted response rates for schools and school enrollments eligible for NAEP HSTS and NAEP, by school type and school status: 2005

School type and HSTS status	Unweighted number of schools selected	Unweighted percent of selected schools	Weighted number of schools selected	Weighted percent of selected schools	Weighted enrollment at selected schools	Weighted percent of enrollment at selected schools
Total	891	100.0	24,731	100.0	3,177,283	100.0
Public						
Total	744	100.0	17,968	100.0	2,911,954	100.0
Respondent	609	81.9	15,021	83.6	2,382,548	81.8
Eligible nonrespondent	135	18.1	2,948	16.4	529,406	18.2
Private						
Total	147	100.0	6,763	100.0	265,328	100.0
Respondent	68	46.3	2,679	39.6	143,355	54.0
Eligible nonrespondent	79	53.7	4,084	60.4	121,973	46.0

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

inverse of the probability of the school being selected into the sample). The weighted enrollment counts are weighted by the product of the school base weight and the 12th-grade enrollment of the school.

Nonresponse in public schools was concentrated in two census divisions—the Northeast Division and the West Division, with the Northeast suffering much more extensive nonresponse—as can be seen in table 7. It was necessary to determine whether the nonresponse adjustments were sufficient. A nonresponse bias analysis was conducted to review the impact of all nonresponse in the study and to determine whether the nonresponse adjustments were sufficient. The conclusion was that the estimates for the national figures were within acceptable bounds, but that the estimates for the Northeast region were not (see appendix N).

Table 7. Weighted HSTS response and eligibility rates for all sampled public schools, by census division: 2005

Public schools	Weighted enrollment estimate	Weighted eligibility rate	Weighted response rate
Total	3,077,044	94.6	86.2
Census division			
New England	153,577	93.2	100.0
Northeast	387,247	95.6	55.5
East North Central	511,266	94.7	87.6
West North Central	247,194	83.2	93.2
South Atlantic	507,108	97.7	95.6
East South Central	160,609	97.8	95.6
West South Central	369,881	96.7	94.4
Mountain	218,704	97.0	98.4
West	521,460	93.2	77.3

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Response rates for private schools eligible for HSTS were generally much lower than those for public school (62 for eligible private schools vs. around 86 for eligible public schools) as can be seen in tables 7 and 9.

Table 8. Weighted HSTS response and eligibility rates for sampled private schools for NAEP HSTS and NAEP, by school type: 2005

Private school type	Total weighted sample	Weighted percent eligible	Weighted response rate (percent)
Total private	314,297	84.4	54.0
Unknown affiliation	31,235	22.0	#
Roman Catholic	139,530	98.2	64.0
Lutheran	4,876	100.0	100.0
Other private	58,815	78.9	41.9
Other religious private	48,892	86.3	37.0
Conservative Christian	30,949	90.3	56.4

Rounds to zero

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 9. Weighted HSTS response and eligibility rates for all sampled HSTS private schools, by school type: 2005

Private school type	Total weighted sample	Weighted percent eligible	Weighted response rate (percent)
Total private	314,297	84.4	62.0
Unknown affiliation	31,235	22.0	#
Roman Catholic	139,530	98.2	70.4
Lutheran	4,876	100.0	100.0
Other private	58,815	78.9	51.6
Other religious private	48,892	86.3	45.7
Conservative Christian	30,949	90.3	71.4

Rounds to zero

NOTE: Detail may not sum to totals because of rounding.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

3.5 Response Rate for Graduates

For the HSTS main study, a graduate was nonresponding only if the selected graduate was eligible for the transcript study and no usable transcript was available for the graduate. For graduates in schools that participated in NAEP, graduates were considered to be selected for HSTS if they had been selected for NAEP even if they did not participate in NAEP. For schools that did not participate in NAEP, graduates were selected systematically when field personnel visited the schools. A student selected for

HSTS was only eligible for the study if he or she graduated in 2005. Tables 10 and 11 present the breakout of the graduate sample according to graduate eligibility, and a breakout of the eligible graduates by whether or not they were in a school that also participated in NAEP.

Table 10. Student samples in HSTS participating schools, by graduation status: 2005

Graduation status	Number of students in sample
Total in sample	29,868
Graduated	27,150
Did not graduate	2,718

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 11. Graduates from HSTS participating schools, by whether the school participated in NAEP: 2005

NAEP school status	Number of students in sample
Eligible students in sample	27,150
In school participating in NAEP	25,233
In school not participating in NAEP	1,917

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 12 presents a breakout of the main study students by eligibility and response status. A student was generally eligible if he or she had graduated, and ineligible if graduation had not been achieved in 2005. A transcript was only usable if at least 75 percent of the credits necessary for graduation were represented on the transcript. A small number of transcripts for eligible students were not usable and were coded out as nonresponse. The overall weighted response rate was 99.7 percent.

Table 13 presents a breakout of HSTS eligibility separately by public and private schools. The percentage of transcripts that were incomplete was about the same in public and private schools, but the ineligibility rate (the percentage who had not graduated) was much higher in public schools.

Table 12. Unweighted and weighted student counts, and student within school response rate, for NAEP HSTS participating schools, by HSTS transcript status: 2005

HSTS transcript status	Unweighted number of students	Unweighted percent of students	Weighted number of students	Weighted percent of students	Student within school response rate
Total	29,868	100.0	2,973,436	100.0	†
Eligible with complete transcript	27,051	90.6	2,723,399	91.6	99.7
Eligible, transcript not complete	99	0.3	9,415	0.3	†
Ineligible	2,718	9.1	240,622	8.1	†

† Not applicable.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 13. Unweighted and weighted student counts, and student within-school response rate, for NAEP HSTS participating schools, by school type and HSTS transcript status: 2005

School type and HSTS status	Unweighted number of students	Unweighted percent of students	Weighted number of students	Weighted percent of students	Student within school response rate
Public					
Total public	27,919	100.0	2,748,422	100.0	†
Eligible with complete transcript	25,166	90.1	2,504,649	91.1	99.7
Eligible, transcript not complete	91	0.3	8,568	0.3	†
Ineligible	2,662	9.5	235,205	8.6	†
Private					
Total private	1,949	100.0	225,013	100.0	†
Eligible with complete transcript	1,885	96.7	218,750	97.2	99.6
Eligible, transcript not complete	8	0.4	847	0.4	†
Ineligible	56	2.9	5,417	2.4	†

† Not applicable.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

For the NAEP-HSTS link study, a graduate had to also participate in a NAEP assessment for the graduate's transcript to be part of the link study (as information from both the transcript and the NAEP assessment is required). Table 14 presents the numbers of sampled graduates in NAEP link schools who were assigned to a mathematics assessment and who were assigned to a science assessment.

Table 14. Student sample in schools participating in NAEP, by NAEP assessment assignment: 2005

NAEP assignment	Unweighted number of students
Total	27,778
Assigned to mathematics	11,353
Assigned to science	16,425

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 15 presents the relationship between HSTS transcript status and HSTS/NAEP link transcript/assessment status. Among HSTS graduating students with complete transcripts, they linked to an assessment in 69 percent of the cases. In 3 percent of the cases, the student was deemed not eligible to take the NAEP assessment due to a disability or limited English proficiency. In 28 percent of the cases, the student did not take the NAEP assessment (usually a refusal or absence). Some transcripts that were incomplete, making them nonresponsive for the HSTS main study, correspond to students deemed ineligible to take the NAEP assessment. For the link study, these will be counted as ineligible.

Table 15. Unweighted NAEP HSTS student sample, by transcript status and school status: 2005

Transcript status and school status	Unweighted number of students	Unweighted percent of students
Total	27,778	100.0
HSTS eligible with complete transcript	25,135	90.5
NAEP/HSTS respondent	17,416	69.3
NAEP/HSTS nonrespondent	7,061	28.1
NAEP/HSTS ineligible	658	2.6
HSTS eligible, transcript not complete	98	0.4
NAEP/HSTS respondent	0	#
NAEP/HSTS nonrespondent	90	91.8
NAEP/HSTS ineligible	8	8.2
HSTS ineligible	2,545	9.2
NAEP/HSTS respondent	0	#
NAEP/HSTS nonrespondent	0	#
NAEP/HSTS ineligible	2,545	100.0

Rounds to zero

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 16 provides a breakdown of the NAEP-linked school student sample by their NAEP transcript status, both unweighted and weighted (with weights defined using the linked student base weights). Table 17 provides the same breakdown by school type.

Table 16. Unweighted and weighted student sample, and student within-school response rate, for schools participating in both NAEP HSTS and NAEP, by transcript status: 2005

Transcript status	Unweighted number of students	Unweighted percent of students	Weighted number of students	Weighted percent of students	Student within school response rate (percent)
Total	27,778	100.0	2,956,139	100.0	†
Assessed and complete transcript	17,416	62.7	1,889,320	63.9	71.2
Incomplete transcript, or no assessment	7,151	25.7	763,828	25.8	†
Ineligible (for NAEP, or nongraduate)	3,211	11.6	302,991	10.2	†

† Not applicable.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 17. Unweighted and weighted student sample and student within-school response rate, for schools participating in both NAEP HSTS and NAEP, by school type and transcript status: 2005

School type and transcript status	Unweighted number of students	Unweighted percent of students	Weighted number of students	Weighted percent of students	Student within school response rate (percent)
Public					
Total public	25,829	100.0	2,692,192	100.0	†
Assessed and complete transcript	15,843	61.3	1,674,178	62.2	69.9
Incomplete transcript, or no assessment	6,835	26.5	721,910	26.8	†
Ineligible (for NAEP, or nongraduate)	3,151	12.2	296,105	11.0	†
Private					
Total private	1,949	100.0	263,948	100.0	†
Assessed and complete transcript	1,573	80.7	215,143	81.5	83.7
Incomplete transcript, or no assessment	316	16.2	41,918	15.9	†
Ineligible (for NAEP, or nongraduate)	60	3.1	6,887	2.6	†

† Not applicable.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Tables 18 and 19 summarize the results for the unlinked and linked studies respectively, presenting the school, student-within-school, and overall student-level response rates.

Table 18. Unlinked study combined response rates, by school type: 2005

	School type		
Unlinked study response rates	Public (percent)	Private (percent)	Total (percent)
Weighted school level	86.2	62.0	84.2
Weighted student within school	99.7	99.6	99.7
Combined response	85.9	61.8	83.9

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

Table 19. Linked study combined response rates, by school type: 2005

	School type		
Linked study response rates	Public (percent)	Private (percent)	Total (percent)
Weighted school level	81.8	54.0	79.5
Weighted student within school	69.9	83.7	71.2
Combined response	57.2	45.2	56.6

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

4. DATA COLLECTION PROCEDURES

This chapter discusses the procedures used in the data collection for the 2005 High School Transcript Study (HSTS 2005). Included are sections on training field workers, contacting the schools, obtaining course catalogs and other school materials, collecting graduate transcripts, sending the data for processing, and receipting and reviewing the data.

4.1 Training NAEP 2005 Field Supervisors as Data Collectors

The field workers for HSTS 2005 were drawn from the pool of 2005 National Assessment of Educational Progress (NAEP 2005) field supervisors. In December 2004, they were trained in the HSTS 2005 data collection procedures. Conducted by HSTS home office staff, the training consisted of several in-person training sessions, each completed in about half a day. In addition, field workers received an HSTS manual that outlined detailed procedures for collecting the data.

The training sessions established the background knowledge needed to help field workers make informed decisions about collecting information in the schools and to explain why attention to detail and accuracy would be crucial in ensuring the quality of HSTS 2005 data. The training also familiarized field workers with the HSTS 2005 materials and forms. The field workers were given examples of various types of high school records and materials, including school- and district-level catalogs, course lists, transcripts, and all the forms used for HSTS 2005. The field workers learned how the data on each of these materials became the information needed at the school and graduate levels. A PowerPoint presentation with the details of data collection was included in the training sessions.

Also during the trainings, the field workers were provided an opportunity to work with practice forms similar to actual materials used for HSTS 2005. The field workers completed sets of exercises designed to provide them with hands-on experience in examining school materials and filling out the forms that they would use.

4.2 Contacts with States, Districts, and Schools

State coordinators in each state were informed about HSTS and were responsible for telling the public school districts in their states about the study. The home office provided them with a Summary of School Activities (see appendix F). The summary provided information about participating in HSTS 2005, including the amount and nature of school staff and time required for participation and procedures that would be used to ensure confidentiality of the data.

Westat field workers contacted school personnel at the school level. Field workers followed the same procedures for contacts with both public and private school personnel. Workers were provided with an informational letter to principals and a Summary of School Activities. They gave these materials to the school principals and school coordinators during their initial contact with the schools.

Initial HSTS information requested by field workers from schools included school specific information that school personnel were asked to provide on the School Information Form (SIF). They were also asked to provide their school's course catalogs for the four most recent school years, including 2004-05, and sample transcripts. This initial information was collected by field workers either through the mail or in person at the time of their first visit. At a later date, the schools were also asked to provide a complete transcript for each graduate in the HSTS 2005 sample. Information provided on the SIF indicated the appropriate date for the HSTS 2005 field workers to obtain these transcripts.

For eligible participating NAEP schools that agreed to cooperate, students sampled for the mathematics and science operational portion of NAEP 2005 were included in the HSTS 2005 sample, and a brightly colored Disclosure Notice was placed in their folder by a NAEP 2005 field worker or school staff member. This notice served two functions:

- It alerted the school personnel that information contained in the student's folder would be used for HSTS 2005.
- Because of its color, it also served as a visible marker for identifying the folders of students in the HSTS 2005 sample to facilitate finding their transcripts at a later date.

Notification describing the student sampling process and the confidentiality safeguards were sent to schools that participated in NAEP (including schools that were substitutes for the original school selected for NAEP) and to schools that were substitutes for schools that participated in NAEP but refused participation in HSTS. Specifically, the notification stated that the intent was to select fifty 12th-grade

students from the school, and student names would be removed from any papers that left the school. Field workers also emphasized that a school's participation in HSTS 2005 would not involve any student time.

For both NAEP 2005 participating and nonparticipating schools (whether original or substitute schools), the initial contact by the field worker included a discussion of the following:

- procedures for obtaining transcripts for the selected students and the method for reimbursing the school for the expense; and
- the availability of a course catalog or course description list.

An appointment was then set to visit the school to prepare the transcript requests and obtain the course catalogs.

4.3 Obtaining Course Catalogs, Sample Transcripts, and Other School-level Information

Prior to HSTS field data collection, Westat contacted schools and requested that they send copies of their catalogs to Westat. This permitted Westat to start catalog coding prior to field data collection.

Field workers requested sample materials for HSTS 2005 when they first contacted a school for HSTS and collected any materials not previously mailed to Westat when they visited the school in the spring or summer. The sample materials included a course catalog (or a list of courses) offered for each of four consecutive years, from school year 2001-02 through school year 2004-05; a completed SIF; and three sample transcripts, one representing a student taking "regular" courses, one with honors courses, and one with special education courses. Since these materials were unique to each school, acquiring them before the collection of the actual transcripts enabled HSTS 2005 staff to examine them and call a field worker or the school to resolve any questions early in the process. Early collection of the catalogs also permitted inputting catalog information prior to receipt of the transcripts, thereby enabling coding to end as soon as possible after data collection.

The field worker also gathered general information about class periods, course credits, graduation requirements, and other aspects of school policy. Sometimes this information was documented in the course catalog and at other times in a separate school policy document.

4.3.1 Catalogs

Course catalogs were carefully reviewed at the school. Field workers verified that the catalogs contained all of the courses that 12th-graders could have taken in high school, including vocational, remedial, honors, special education, off-campus courses, or courses taught in a language other than English. If these course listings were not in the catalog, every effort was made to obtain additional information from school personnel or, in some cases, through web searches to document the existence of such courses and to describe them.

HSTS requested course catalogs containing the most comprehensive information about the courses offered by the schools. Ordered from most to least complete, the requested types of catalogs are as follows:

- a school-level catalog providing course titles and descriptions;
- a district-level catalog, if it indicated which courses were offered at the HSTS participating school;
- a course list by department that included general descriptions of course offerings by department;
- a school-level course list without descriptions; or
- a district-level catalog without any indication of which courses were offered in specific schools.

All catalogs and course lists that were received by field workers were forwarded to HSTS 2005 data processing staff.

4.3.2 Sample Transcripts

Since transcript format varies greatly among school districts throughout the country, three transcripts of previous graduates were obtained from each school by the NAEP field workers during the initial call or visit to the school. The three transcripts requested from each school included one that contained honors-level courses, one that contained special education courses, and one that contained just the "regular" courses. The HSTS field workers marked each transcript to indicate where on the transcript the needed information was found and how information regarding course level was coded. Attached to

each marked-up transcript was a Transcript Format Checklist (appendix H) indicating the key transcript information and whether or not that information was found or, if so, whether it was marked on the school's transcripts.

4.3.3 SIF and HSTS Questionnaire

The SIF was forwarded for data processing along with the other preliminary materials as described above. The SIF was completed by the field worker. Along with general school information, the completed SIF contained the following information:

- sources of information within the school (if needed to complete HSTS 2005 data collection);
- graduation requirements;
- grading practices at the school;
- format of the school's transcripts; and
- name and position of the school's HSTS 2005 coordinator who helped complete the form.

The field workers were instructed to fill out the SIF completely or to indicate clearly on the SIF where the requested information could be found in the other materials provided by the school.

4.3.4 School Background Questionnaire

The School Background Questionnaire (see appendix K) is a NAEP 2005 questionnaire that collected information about school, teacher, and home factors that might relate to student achievement. It was completed by a school official (usually the principal) as part of NAEP 2005 for the NAEP participating schools. Field workers asked HSTS schools that did not participate in NAEP 2005 to complete a School Questionnaire.

4.3.5 SD and LEP Questionnaires

The questionnaires that NAEP 2005 used to collect information from school staff about students with disabilities and students with limited English proficiency are called the SD Questionnaire and LEP Questionnaire, respectively (see appendices L and M). Schools were asked to have the person most knowledgeable about a disabled or limited English proficient student complete the questionnaire(s). In large schools, this person was typically a counselor, a special education teacher, or a teacher of English as a second language. In smaller schools, this person was typically a classroom teacher. For schools participating in NAEP 2005, the SD and LEP Questionnaires were collected as part of the NAEP procedures.

4.4 Identifying the Sample of Students and Obtaining Transcripts

There were 726 schools that participated in HSTS, and 677 of these schools participated in NAEP 2005 and HSTS. Of these 677 schools, 44 were substitute schools.

HSTS 2005 used the NAEP 2005 sample for selecting schools and students in NAEP participating schools. For schools that participated in NAEP 2005, the student sample was recorded on the NAEP 2005 Administration Schedules. For schools that did not participate in NAEP 2005, the field worker drew a sample of graduates at the school. Details on how this sample was drawn can be found in section 3.2 and 3.3. The procedures for identifying graduates in schools with NAEP 2005 materials and in schools without NAEP 2005 materials are described in detail in separate sections that follow.

4.4.1 Materials from NAEP 2005 Schools

Transcripts were requested for all students who were sampled for the operational mathematics and science part of NAEP 2005. They included all assessed students, sampled students who were absent during the NAEP assessment, and SD and/or LEP students who were excluded by the school from participating in the assessment.

Once graduation information was posted on transcripts, a field worker returned to the school to obtain the requested transcripts. At that time, the field worker used a Transcript Request Form (TRF)

(see appendix C) to obtain basic information about the sampled students that was not available from NAEP 2005 data files. In addition to student name and NAEP ID, it contained columns for entering graduation status, gender, birth month and year, race/ethnicity, SD status, LEP status, Title 1 participation, and National School Lunch Program participation for each listed student. Data available from NAEP 2005 files (NAEP ID and demographic variables) were preprinted on the form.[13] The completed TRFs contained the following information:

- **Student Name –** The field worker recorded the first name, middle initial, and last name of each assessed, absent, or excluded student listed on the NAEP 2005 Administration Schedule. These entries were made to correspond to the preprinted NAEP ID.

- **NAEP ID –** The 10-digit NAEP 2005 assessment booklet numbers and SD and LEP questionnaire numbers for students excluded from the assessment were preprinted in ID order. This column on the TRF identified all students for whom transcripts were needed.

- **Exit Status –** Sometimes the exit status was determined directly from the transcripts, and sometimes it was determined by other records or provided by school personnel Using this information, field workers assigned one of the following codes to describe each student's outcome at the school:

  - graduated with a standard diploma
  - graduated with an honors diploma
  - received a diploma with special education adjustments
  - received a certificate of attendance
  - still enrolled in this school
  - dropped out
  - transferred
  - withdrawn
  - GED
  - other or reason unknown

[13] To ensure consistency between NAEP 2005 and HSTS 2005, the field staff were instructed not to change the preprinted demographic information.

- **Birthdate, Gender, and Race/Ethnicity –** Demographic information was generally preprinted for each sampled student. If not preprinted, it was recorded from the NAEP 2005 Administration Schedule.

- **SD and LEP Status –** For each student, it was recorded whether or not the student was classified by the school as SD and/or LEP.

- **National School Lunch Program and Title I –** Field workers recorded yes or no for participation in each of these programs.

- **Transcript Received –** Field workers checked this column to document that the transcript for a given student had been received.

Once the TRF was completed by carefully transferring student information from the Administration Schedules, the field worker filled out the summary box at the top of the form and requested transcripts according to the procedures set forth by the school. As already noted, the Disclosure Notice placed in students' folders at the time of the first visit helped to facilitate transcript collection in participating NAEP schools.

Once the field worker filled in the names of the students, some schools were able to access an electronic data file and print the transcripts. In other schools, the school coordinators pulled transcripts from their folders and photocopied them at the school.

When the request for transcripts was filled, the field worker reviewed the transcripts to ensure that a transcript had been received for each 12th-grade student selected for the operational mathematics or science portion of the NAEP 2005 assessment, whether or not that student had graduated. Even though nongraduate transcripts were not included in HSTS, each student graduation status needed to be accounted for and verified, so that weighting could be done correctly. Each transcript was checked for eligibility, understandability (e.g., all the codes on it were defined on the transcript or explained in the SIF), and completeness. The field worker then labeled each transcript with preprinted labels containing the School ID and the NAEP ID for the student. The field worker completed a Documentation of Missing Transcripts form to explain the reasons the school gave for any missing transcripts.

4.4.2 Schools without NAEP 2005 Materials

The procedures for schools that did not participate in NAEP were similar to those for schools that participated with the following exceptions:

- As discussed in chapter 3, field staff were responsible for selecting a systematic sample of graduates, since there were no students designated to receive operational mathematics or science assessments.
- Demographic information on the TRF had to be obtained for the sampled graduates.
- The school was asked to complete the NAEP 2005 school questionnaire and an SD or LEP questionnaire for any of the graduates sampled who were classified as SD or LEP.
- Data collection for non-NAEP schools started later than for the NAEP schools. Therefore, catalogs were frequently collected at the same time as the transcripts. In this case, the field worker annotated three actual transcripts from among those that were collected.

4.5 Sending Data for Processing

As with NAEP 2005, safeguards were built into the procedures for the transcript study to ensure that applicable privacy requirements were met. After transcripts were collected and all information on sampled graduates recorded, field workers prepared the transcripts for transmittal to the data processing staff. They first compared the graduate ID and name on the transcripts to the TRF to verify that they had obtained and correctly labeled the transcripts. At the same time, they noted on the TRF which transcripts were received and which were not. They then cut off the left hand column of the TRF, which contained the names of the graduates. The list of names remained in the schools (and was ultimately destroyed) and the remainder of the TRF was placed in the package to send to the HSTS 2005 field officer for data processing.

A Shipping Transmittal Form (appendix J) accompanied all shipments to the data processing staff and summarized the types and number of materials being sent. This form also gave information on whether the transcripts were from the NAEP 2005 list or a new sample and, if the school did not participate in NAEP 2005, whether course catalogs and a SIF were included in the shipment.

This page intentionally blank.

5. DATA PROCESSING PROCEDURES

As discussed in chapter 4, schools provided a wide variety of data for use in the 2005 High School Transcript Study (HSTS 2005). This chapter explains how these data were processed to produce the study's data files. Figure 1 depicts the data flow for the project.

5.1 NAEP 2005 Questionnaires

The main National Assessment of Educational Progress (NAEP) study provided HSTS staff with data files for schools and students included in NAEP 2005. The School Questionnaires and the Students with Disabilities (SD) and Limited English Proficiency (LEP) Questionnaires[14] collected in HSTS 2005 for non-NAEP schools were the same as the ones used in NAEP 2005. These questionnaires were electronically scanned, using the same procedures used in NAEP. Information from these scanned questionnaires was added to the appropriate NAEP data files to constitute HSTS 2005.

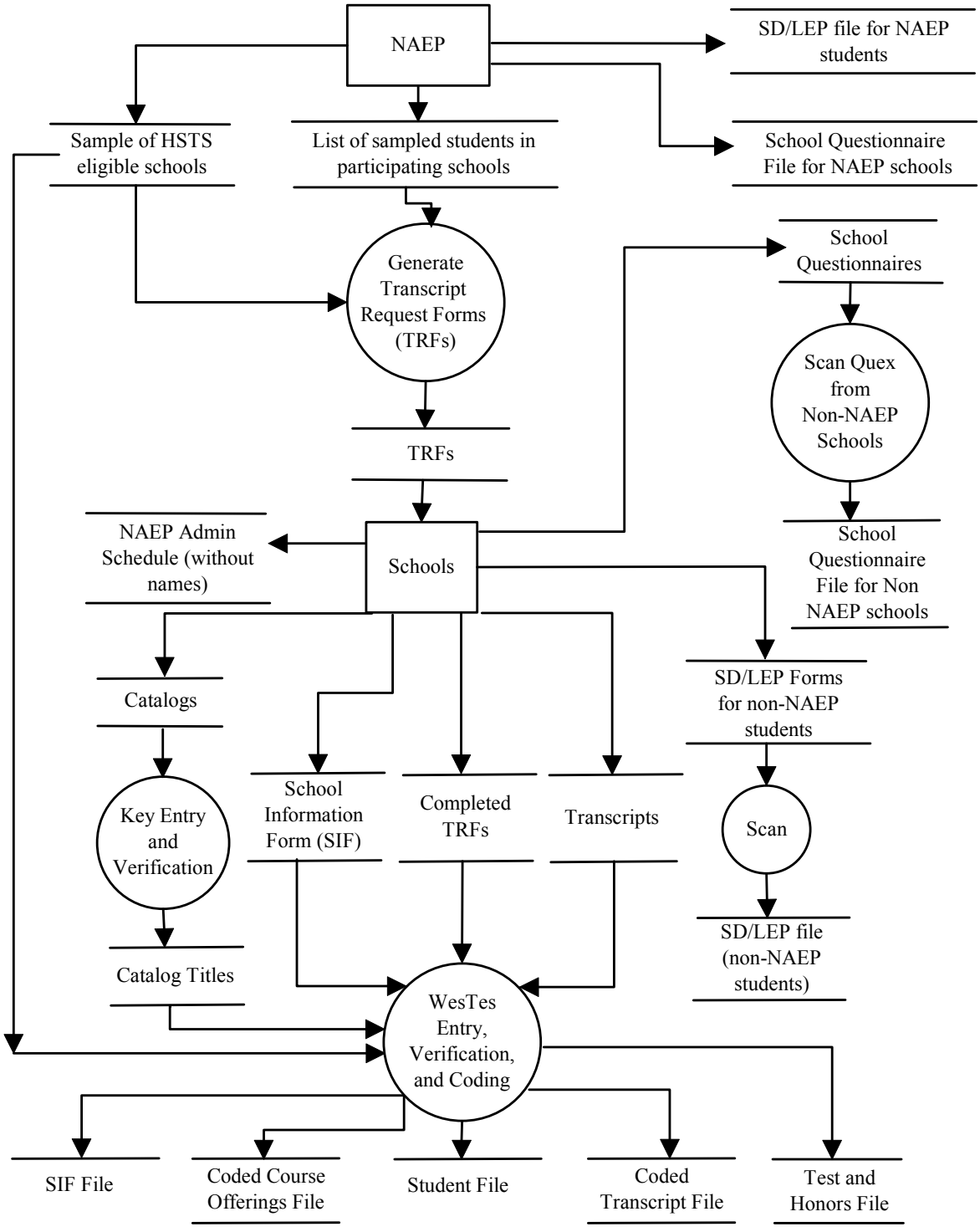
5.2 Data Collected Other Than NAEP 2005 Questionnaires

5.2.1 Westat Transcript Entry System (WesTes)

WesTes is a custom-built Structured Query Language (SQL) server application specifically designed for processing large-scale transcript-based studies in an accurate and efficient manner. It stores most of the school and graduate information collected for the study in a single integrated relational database. It is used to ensure that the data collected by HSTS is properly tracked and to assist the data entry and coding personnel in the prompt and accurate completion of their tasks.

[14] See section 1.3 for a description of these questionnaires and appendices K, L, and M for copies of them.

Figure 1. Data flow for the High School Transcript Study 2005



NOTE: NAEP = National Assessment of Educational Progress; SD = Students with Disabilities; LEP = Limited English Proficiency; Quex = Questionnaires; WesTes= Westat Transcript Entry System
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Transcript Study (HSTS), 2005.

5.2.2 Receipt Control

This section discusses the receipt control tasks associated with HSTS 2005. These include using controls built into WesTes and entering School Information Forms (SIFs), catalogs, other non-transcript data and transcripts.

5.2.2.1 Preparation of WesTes for Receipt Control

The NAEP samples of schools and students selected for the HSTS 2005 sample were loaded into WesTes. When a school refused participation and a substitute school was identified during data collection, the substitute school was added to WesTes and the original school was deleted. The list of schools was used to ensure that data could be entered into the system only for schools sampled for the NAEP HSTS 2005 sample or designated substitute schools. After a valid school ID had been entered by a staff member, the system allowed staff to enter data only for graduates in the sample for that school.

5.2.2.2 SIFs, Course Catalogs, and Other Nontranscript Data

When the packages containing the SIF and catalogs were received at Westat, receipt clerks selected the school ID in WesTes and entered the receipt date for each of these materials. Receipt clerks were also responsible for entering the data from the SIF (see appendix B) into WesTes. These data included substantial amounts of information needed to correctly interpret catalog and transcript entries. Of particular importance is the information on the number of credits given by the school for one Carnegie unit and the number of credits required to graduate. These data were 100 percent verified.

5.2.2.3 Transcripts

When transcript study materials arrived for data processing, a receipt clerk carefully reviewed all items for accuracy and completeness. Transcripts were matched to the Transcript Request Form (TRF). Field workers were contacted immediately if further clarification was needed.

After reviewing the transcript materials, the clerk recorded the transcript materials and followup requests (when required) using the WesTes transcript receipt module. For each school, the clerk compared the number of transcripts requested with the number actually received and reviewed and verified the list of all student IDs assigned to that school. The clerk entered the receipt date for the package and whether or not it contained a TRF. For each student ID in the school, the receipt clerk indicated whether or not it had been received, the receipt date, and the exit status as indicated on the TRF. The receipt staff could update the exit status of students based upon determinations from the coding and transcript management staff if the transcript was found to be not usable or incomplete. Schools were reimbursed for the cost of producing the transcripts within 2 weeks after their materials were received for data processing.

5.2.3 Catalog Coding

This section describes the process of coding catalogs.

Catalog coding staff highlighted the course titles and course numbers, if available, in the catalogs and sent them for key entry. Data entry personnel keyed the catalog titles and, when available, course numbers assigned by the school into ASCII files and key verified the data. The files containing the course titles were then uploaded into WesTes.

After the course titles were loaded into WesTes, a catalog coder reviewed all the materials from a school to obtain an understanding of the school's curriculum and any special circumstances that would clarify the nature of the content of specific courses. Using the WesTes Catalog Coding Screen, the catalog coder displayed each catalog title individually, reviewed the corresponding entry in the school's catalog and then entered the CSSC code that best matched that description.[15]

The catalog coder also set flags to indicate whether the course was a special education course, whether the course was part of a sequence of closely related courses, taught off campus, taught in English, and the level of the course.[16]

[15] See section 1.4 for additional information about the CSSC.

[16] See appendix O for information on these code values.

When coding a course, catalog coders matched the course description in the high school's catalog to the course description in the CSSC.[17] Using course descriptions rather than titles for purposes of assigning CSSC codes to catalog courses is important because the course title often does not provide sufficient information to differentiate between codes. For example, a course with a name such as Algebra 1 could be a remedial course, a reduced-pace algebra course, a first semester algebra course, or the first year of a series of algebra courses. To facilitate the matching process, the full CSSC was available online to the catalog coders. If a transcript course cannot be found in the course catalog from that school, the course is added to the course offerings file. This typically occurs when course catalogs are not up to date or complete.

5.2.4 Coding Transcripts

Transcripts may contain a variety of information in addition to lists of courses taken and the grades and credits earned for each course. Many include information such as graduation date and class rank, for which only one entry is made per graduate, as well as the names of tests taken by the graduate, test scores, and honors awarded. This section describes the coding process for these additional items as well as for the course information.

5.2.4.1 Single-entry Items

The following single-entry items were recorded for each graduate when they were available on the transcript:

- graduation date;
- class rank;
- size of class;
- grade point average (GPA);
- adjusted GPA (as reported by the school);
- days absent in 9th grade;

[17] See appendix O for information on these code values.

- days absent in 10th grade;
- days absent in 11th grade;
- days absent in 12th grade;
- total credits received;
- total credits attempted;
- whether the graduate received a General Equivalency Diploma (GED); and
- date of GED completion.

5.2.4.2 Honors

If a transcript listed honors, the date (month and year) and a description of the honor were entered. In order to speed data entry, the following common descriptions were included on a drop-down list:

- national honors;
- athletic honors;
- academic honors;
- honor roll; and
- other.

When "other" was selected, the data entry clerk typed in the name of the award. Many of these referred to specific subject matter such as English, algebra, or chemistry and had names like English 9 Award. Others were not subject-matter specific, for example, the Golden State Seal Merit Diploma.

5.2.4.3 Tests

Tests were recorded in much the same way as honors. A pick list was provided containing the names of the most common tests that appear on transcripts.[18] Test scores were recorded when available.

5.2.4.4 Course Entry

Transcript courses required the most extensive portion of the data entry effort. This effort was because the graduates' transcripts, on average, included 46 distinct course entries. For each course, the transcript entry staff recorded the grade level (9th, 10th, 11th, or 12th), the year in which the course was taken, the term (e.g., fall semester, summer school), the course name, the grade, and the number of credits earned. The transcript entry staff also set a flag to indicate whether or not a course was transferred from another school. In addition, since information on whether a course was a special education course, was taught off campus, was taught as English as a second language course or in a foreign language, and the level of the course is recorded on some transcripts, the transcript entry staff set flags representing each type of information. These flags had the same possible values as the corresponding flags used to code courses listed in the schools' catalogs.

5.2.4.5 Assigning CSSC Codes through Title Matching

One of the most challenging aspects of the transcript coding process is linking the course titles on the transcripts to the appropriate catalog course title in order to assign each transcript course an appropriate CSSC code. This was done through a process known as title matching, which was performed by coders who were trained for title matching after the preceding transcript information had been entered for all graduates from a school. To the greatest extent possible, title matchers worked with the same high school catalogs that they coded in the spring. This permitted them to capitalize on their knowledge of state education systems, graduation requirements, and acronyms.

[18] The test names on the pick list included ACT Composite, English, Mathematics, Reading, and Science; PLAN Composite; PSAT Writing; SAT Math and Verbal; and Stanford Language, Mathematics, Science, Social Science, and Total Reading.

Because course titles on transcripts are frequently different from the course titles a school uses in its catalog, it is usually not practical to fully automate the title matching process. In addition, even when it is possible to automate title matches, a review by knowledgeable staff of the resulting matches is an important quality control check. For these reasons, title matching was performed by experienced catalog coders using computer-assisted matching tools built into WesTes.

WesTes presented the title matcher with a list of all the unique title and flag combinations appearing on a school's transcript and a list of the coded catalog titles for the school. The title matcher's task was to match each of the titles from the transcripts to a catalog course. Title matchers matched nontransfer course titles on transcripts to the high school's catalog. They matched transfer courses directly to the most likely CSSC course description found in the generic catalog. The generic catalog was the most current version of the CSSC file.

They matched transfer courses directly to the most likely CSSC course description. In this case, the CSSC was used like a course catalog.

Title matchers used all of the title and flag information that was comparable in both the catalogs and the transcripts for a school. For example, if the school distinguished between a regular English 9 course and an honors English 9 course in the catalog and on its transcripts, title matchers would ensure that a transcript course named English 9 with the level flag set to the honors code would be matched with the catalog course English 9 Honors (CSSC code 230111) and not with the average or remedial English 9 courses in the catalog. However, if the catalog did not distinguish between the different levels of English 9 but the transcripts did, the catalog coders would match all the English 9 courses on the transcripts to the catalog English 9 course, even though the catalog had the flag set to the default regular level. Often courses in the catalog represent classes in which the student ability is mixed. The flag for the course in the catalog is set at the lowest level, thus a class that offers honors or even AP within a regular class is given a flag for general level. If the transcript indicates that a particular student received credit at a level that differs from the general catalog level, the transcript flag for that student is set at the appropriate level. When these data are analyzed, the level assigned to the student for the course is based upon the transcript if it is different than the level flag found in the course catalog.

5.2.5 Quality Control Procedures

Procedures designed to ensure a high quality data processing operation include the careful hiring and training of HSTS staff, rekeying data for verification purposes, automated tests to identify records for review, and the use of logs to identify and rectify problems. Each of these quality assurance measures is discussed in a separate section below.

5.2.5.1 Hiring and Training

Central to quality control is having thoroughly trained, well-qualified staff. Westat, therefore, spent considerable effort on selecting and training data processing staff. Two distinct groups of staff members were recruited and trained for HSTS 2005: catalog coders and data entry staff. Catalog coders matched the course descriptions in each school's catalog to the corresponding code in the CSSC, and a subgroup of them also matched the course titles on the transcripts to the corresponding titles in the school's catalog. Data entry staff entered the transcript information into the project database.[19]

Transcript Data Entry Staff

Transcript data entry staff members were selected for their ability to enter data accurately and consistently. They were then provided with extensive training that explained the study and taught them how to enter data from high school transcripts.

The transcript entry training spanned 5 days, with new concepts introduced in the morning and practical application exercises performed in the afternoon. The primary function of the training was instructing the staff in the use of WesTes for entering data found on transcripts with an emphasis on hands-on practical experience. The secondary function of the training was covering the basic concepts and challenges the staff would encounter while entering high school transcripts.

[19] In addition to staff described here, other Westat data entry clerks did some straightforward data inputting tasks, such as typing the names of course titles.

Catalog Coders

Requirements for employment as a catalog coder included a minimum of a bachelor's degree in a social science, with a master's degree preferred, and 2 or more years of teaching experience. The interview process paid special attention to experience in teaching a variety of core content areas as well as special education, knowledge of school curricula and procedures, attention to detail, application of analytical skills, and dedication to accuracy.

Training lasted for 5 days. It was guided by a training manual covering the following topics: high school catalog components; state requirements for graduation; special education programs; course levels and flags; transfer courses; honors and advanced placement programs; and vocational, work, and career-related programs.

Training activities involved informative presentations, visual demonstrations, and practical applications. Examples illustrating salient points were drawn from actual materials. Coders learned to use the WesTes coding system, employing its category and subcategory search. Coders also became familiar with the CSSC and the importance of studying a course's full description before assigning a CSSC code. A final exercise was given on the last day of training as a graded evaluation of coding analysis and application to verify that coders were able to meet the accuracy standards for the project.

Title-Matching Staff

Six catalog coders were selected to perform the title-matching task. The training for title matching was similar in both form and procedure to catalog coder training. During a 4-day period, title matchers learned to match the transcript information entered during transcript entry and verification with the CSSC catalog course codes that they had assigned during catalog coding.

5.2.5.2 Rekeying for Verification

All data entry from transcripts was verified by a staff member other than the one who initially entered the data. This required blindly rekeying most of the information on the transcript. However, the names of honors, tests, and courses were displayed during verification, because the verifier

needed to see them to make meaningful entries in the related fields. These fields were only rekeyed when the verifier believed that an error had been made in entering them. Since course titles were keyed in full and were later matched to catalog titles, verifiers paid particular attention to any data entry errors in the titles that might potentially lead to incorrect matches. The system alerted the verifier to any differences between his or her entries and the original entries. The verifier then had the opportunity to confirm or change the most recent entry.

Verification showed that initial data entry was quite accurate. For example, verifiers changed 2 percent of grades and less than 1.5 percent of course credits.

5.2.5.3 Automated Tests to Identify Records for Review

Several automated reports were developed to identify records to be manually reviewed. Some reports were developed to assist coders in identifying records to be re-examined before submitting cases for review. Other reports were designed for supervisors to use to identify courses for their review.

The automated reports highlighted items that appeared to be inconsistent or to have unusually high or low values. For instance, reports were generated of transcripts within a school that had not yet been coded. A more complex type of report was a list of course titles containing the word "honors" without the honors flag having been set in the catalog.

5.2.5.4 Use of Logs to Identify and Rectify Problems

Specific problems encountered during data entry and coding operations were entered on one of two logs: (1) Supervisor's Problem Log, and (2) System's Problem Sheet. Problems identified in the Supervisor's Problem Log were discussed in a meeting of HSTS staff, and the resolutions were indicated on each problem sheet. A System's Problem Sheet was submitted when the problem encountered was with WesTes. A description of the problem and a screen shot (when applicable) were included and the systems specialists corrected these problems.

This page intentionally blank.

6. WEIGHTING AND ESTIMATION OF SAMPLING VARIANCE

6.1 Overview of Weighting in HSTS 2005

This chapter explains how the weights associated with the 2005 High School Transcript Study (HSTS 2005) were calculated. Restricted-file users do not need to understand all the material in this chapter; however, they do need to be aware of which types of weights are appropriate to use with the analyses they wish to do. The appropriate weights to use are discussed in section 7.3.1.

The next sections of this chapter discuss school weights and graduate weights. The final section presents additional information about the replicate weights. This chapter uses the following terminology in discussing the weights:

- **Linked weights:** Weights that should be used for point estimates for a variable that is only available for schools or graduates in schools that can be linked to the National Assessment of Educational Progress (NAEP) (such as NAEP scores or parental education).
- **Unlinked weights:** Weights that should be used for point estimates for a variable available for all responding schools or all graduates in the HSTS sample (such as highest science course completed or grade point average [GPA]).
- **Replicate weights:** Weights used when estimating variances for point estimates.

6.2 School Weights

The final school weights consist of the product of base weights that reflect differences in the probability of schools being selected for HSTS and weights that adjust for differential nonresponse rates for different types of HSTS schools.

This section first discusses the school base weights and then discusses adjustments for nonresponse. Finally, it presents the formulas used for calculating the school weights from the base weights and the nonresponse adjustment factors.

6.2.1 School Base Weights

The school base weights w_s and the 62 corresponding replicate weights $w_s(r)$ were computed, using the following factors:

- a factor equal to the inverse of the school's probability of being sampled for NAEP [W_s and $W_s(r)$];
- a factor equal to the inverse of the school's probability of being subsampled for HSTS (W_c); and
- a factor equal to the inverse of the school's probability of having students selected for participation in either the NAEP mathematics or science assessment ($SCHSES_s$).[20]

The formulas for calculating the HSTS base weights and replicate weights for the schools are as follows:

$$w_s = W_s * W_c * SCHSES_s$$
$$w_s(r) = W_s(r) * W_c * SCHSES_s$$

6.2.2 School Nonresponse Adjustments for HSTS 2005

This section describes the weighting cell adjustment used for adjusting for school nonresponse. The starting point for a cell structure was the strata from the original NAEP school sampling process (for 12th-grade public and 12th-grade private schools). This was also the cell structure used in the development of nonresponse cells for NAEP. When cells based on the strata were too small to allow for stable nonresponse adjustment, cells were collapsed. The final school nonresponse weighting cells for use with the unlinked HSTS responding school sample were designated as $SNRADJ_c^{(U)}$, c=1,…,$C^{(U)}$, where $C^{(U)}$ is the total number of weighting cells, S_c is the set of all eligible original HSTS schools in cell c, $R_c^{(U)}$ is the set of all unlinked responding schools (with responding substitutes replacing original nonrespondents) within S_c ; and w_s is the HSTS school full sample base weight.

[20] Schools with fewer than 24 students in the 12th grade did not necessarily participate in either the operational science or mathematics assessments. Two-ninths of the public schools with less than 24 students were assigned to a bridge session only, and one-eighth of the private schools with less than 24 students were assigned to a bridge session only. Thus, schools with fewer than 24 students were weighted by the inverse of the probability of having an operational mathematics or science assessment.

$$SNRADJ_c^{(U)} = \frac{\sum_{s \in S_c} w_s}{\sum_{s \in R_c^{(U)}} w_s}$$

In words, this calculation indicates that the adjusted weight equals the total base weight for all schools divided by the total base weight for all responding schools in the cell.

Because many analyses are restricted to schools for which student NAEP scores can be linked to HSTS data, a similar nonresponse adjustment was done for schools in the linked sample. The corresponding cells for the linked NAEP-HSTS responding school sample were $SNRADJ_c^{(L)}$, c=1,…,$C^{(L)}$, where $C^{(L)}$ is the total number of weighting cells and ($R_c^{(L)}$) is the set of all linked responding school (with responding substitutes replacing original nonrespondents). The school nonresponse adjustments were computed as follows for each nonresponse weighting cell:

$$SNRADJ_c^{(L)} = \frac{\sum_{s \in S_c} w_s}{\sum_{s \in R_c^{(L)}} w_s}$$

For the unlinked weights, schools were eligible if they had at least one graduate in their class, and if they satisfied the NAEP criteria for eligibility. A school was cooperating if it cooperated with the HSTS survey (whether or not it also participated in NAEP 2005).

For the linked weights, the eligibility criteria were the same as for the unlinked weights. A school was considered cooperating only if it cooperated both with NAEP 2005 and HSTS 2005 and retained the information needed to link the NAEP and HSTS records.

6.2.3 School Base Weights Adjusted for Nonresponse

The unlinked (linked) school nonresponse-adjusted weight $SCHWGT_s^{(U)}$ ($SCHWGT_s^{(L)}$) is equal to

$$SCHWGT_s^{(U)} = w_s * SNRADJ_c^{(U)}$$
$$SCHWGT_s^{(L)} = w_s * SNRADJ_c^{(L)}$$

The replicate school nonresponse adjustments were computed as follows for each nonresponse weighting cell:

$$SNRADJ_c^{(U)}(r) = \frac{\sum_{s \in S_c} w_s(r)}{\sum_{s \in R_c^{\{U\}}} w_s(r)}, \qquad r = 1,...,R$$

$$SNRADJ_c^{(L)}(r) = \frac{\sum_{s \in S_c} w_s(r)}{\sum_{s \in R_c^{\{L\}}} w_s(r)}, \qquad r = 1,...,R$$

The unlinked (linked) school nonresponse-adjusted replicate weights $SCHWGT_s^{(U)}(r)$ ($SCHWGT_s^{(L)}(r)$) are equal to

$$SCHWGT_s^{(U)}(r) = w_s(r) * SNRADJ_c^{(U)}(r)$$
$$SCHWGT_s^{(L)}(r) = w_s(r) * SNRADJ_c^{(L)}(r)$$

Appendix N presents estimates of enrollment, by grade and standard errors, using these weights and replicate weights. As can be seen, the school nonresponse adjustments had the effect of allowing the responding schools to represent the full set of schools. The replicate adjustments were designed to produce variance estimates reflecting the component of variability added by the nonresponse adjustment process.

6.3 Weights for Graduates

The final weights for graduates consist of the product of the base weights, which are equal to the inverse of the probability of the graduate being selected for HSTS; a nonresponse adjustment factor; and a trimming factor used to ensure that individual graduate weights are not excessively large. For graduates, the following types of final weights are calculated:

- weights for all graduates in the HSTS 2005 sample;

- weights for graduates in the HSTS 2005 sample for whom NAEP mathematics scores are available;
- weights for graduates in the HSTS 2005 sample for whom NAEP science scores are available; and
- weights for graduates in the HSTS 2005 sample for whom either NAEP mathematics or science scores are available.

All of these weights are designed to estimate variables for all graduates. Which type of weight should be used depends upon what type of data the user is analyzing. For example, in estimating the GPA of graduates, the first weight would be used. If, however, the user wishes to explore the relationship between NAEP mathematics scores and GPA, the user will use the second type of weight, because these estimates must be based on the subsample of all HSTS graduates who also took the NAEP mathematics assessment. Similarly, if the user wishes to explore the relationship between NAEP science scores and GPA, the user will use the third type of weight, because these estimates must be based on the subsample of all HSTS graduates who also took the NAEP science assessment. Finally, the fourth weight is used when the user wishes to estimate a variable (such as responses to the student questionnaire) available for all graduates who took either the mathematics or the science assessment but not for other HSTS graduates.

6.3.1 Base Weights

The number of weights calculated for a graduate in HSTS 2005 depended upon which of the four possible samples the graduate belonged to (i.e., all graduates, all graduates who took the NAEP mathematics assessment, all graduates who took the science assessment, and all graduates who took either the mathematics or the science assessment). All sample members have a weight (referred to as the unlinked weight) used to estimate statistics for all graduates. Since students could not participate in both a mathematics and a science assessment, those graduates with linked NAEP data will have two weights in addition to the unlinked weight – either the mathematics or science linked weight and the combined linked weight. The summation of the graduate base weights over a particular subgroup is an unbiased[21] estimator of the total number of graduates in that subgroup in the population.

[21] This assumes that the school nonresponse adjustments completely adjusted for bias due to school nonresponse.

6.3.1.1 Base Weights for Graduates in Schools without Linked NAEP Data

The unlinked base weight for graduates ($STUWGT_{sk}^{(U)}$) within the unlinked schools was different from the student weight in the NAEP sample because, as explained in chapter 3, the two samples are not identical. The unlinked weight in this case is a product of the following factors (where *s* indicates school; *k* indicates graduate):

- the school nonresponse adjusted weight $SCHWGT_s^{(U)}$;
- the substitute school weighting factor ($SUBADJ_s$) that adjusts for differences in enrollment between the original school and the substitute school; and
- the within-school sampling interval[22] for graduate selection $WINSCHWT_s$.

To summarize:

$$STUWGT_{sk}^{(U)} = SCHWGT_s^{(U)} * SUBADJ_s * WINSCHWT_s$$

The corresponding replicate weights ($r=1,\ldots,R$) for unlinked schools are as follows:

$$STUWGT_{sk}^{(U)}(r) = SCHWGT_s^{(U)}(r) * SUBADJ_s * WINSCHWT_s$$

6.3.1.2 Base Weights for Graduates within Schools with Linked NAEP Data

The unlinked graduate base weight $STUWGT_{sk}^{(U)}$ within the linked schools is a product of the following factors (*s* indicates school; *k* indicates graduate):

- the school nonresponse adjusted weight $SCHWGT_s^{(U)}$ discussed earlier;
- the substitute school weighting factor ($SUBADJ_s$) that adjusts for differences in enrollment between the original school and the substitute school;
- the within-school sampling interval for student selection ($WINSCHWT_s$);

[22] The sampling interval is the reciprocal of the probability of selection.

- the assessment session assignment weighting factor ($STUSESWT_s$) that adjusts for the probability of a selected student being selected for a session that includes operational mathematics and/or science assessment;

- a factor for year-round schools ($YRRND\_FC_s$) that adjusts for students not being available for NAEP testing because they are not enrolled in the school during the semester that NAEP is administered; and

- a factor ($MASC\_AF_s$) for selection of the student to a mathematics or science operational booklet (rather than a reading operational booklet, a pilot booklet, or a bridge booklet), given the student was in a session including operational mathematics and/or science assessment. This factor was equal to 75/49.

Similarly, the linked graduate base weight $STUWGT_{sk}^{(L)}$ within the linked schools is a product of the following factors:

- the school nonresponse adjusted weight $SCHWGT_s^{(L)}$ discussed earlier;

- the substitute school weighting factor ($SUBADJ_s$) that adjusts for differences in enrollment between the original school and the substitute school;

- the within-school sampling interval for student selection ($WINSCHWT_s$);

- the assessment session assignment weighting factor ($STUSESWT_s$) that adjusts for the probability of a selected student being selected for a session that includes operational mathematics and/or science assessment;

- a factor for year-round schools ($YRRND\_FC_s$) that adjusts for students not being available for NAEP testing because they are not enrolled in the school during the semester that NAEP is administered; and

- a factor ($MASC\_AF_s$) for selection of the student to a mathematics or science operational booklet (rather than a reading operational booklet, a pilot booklet, or a bridge booklet), given the student was in a session including operational mathematics and/or science assessment. This factor was equal to 75/49.

To summarize:

$$STUWGT_{sk}^{(U)} = SCHWGT_s^{(U)} * WINSCHWT_s * STUSESWT_s * SUBADJ_s * YRRND\_FC_s * MASC\_AF_s$$

$$STUWGT_{sk}^{(L)} = SCHWGT_s^{(L)} * WINSCHWT_s * STUSESWT_s * SUBADJ_s * YRRND\_FC_s * MASC\_AF_s$$

The corresponding replicate weights are as follows ($r=1,\ldots,R$) for linked schools:

$$STUWGT_{sk}^{(U)}(r) = SCHWGT_{s}^{(U)}(\mathrm{r})*WINSCHWT_{s}*STUSESWT_{s}*SUBADJ_{s} *YRRND\_FC_{s}*MASC\_AF_{s}$$

$$STUWGT_{sk}^{(L)}(r) = SCHWGT_{s}^{(L)}(\mathrm{r})*WINSCHWT_{s}*STUSESWT_{s}*SUBADJ_{s} *YRRND\_FC_{s}*MASC\_AF_{s}$$

6.3.2 Nonresponse Adjustments

The methods used for nonresponse adjustment for the HSTS 2005 graduates were very similar to those used for the NAEP 2005 12th-grade operational studies, with a few minor differences. Nonresponse adjustments were done separately for the unlinked and the linked samples of graduates. For the linked samples, it was necessary for the graduate to be both a NAEP respondent and a graduate with a transcript to be a linked sample respondent.

6.3.2.1 Preliminary Formation of Weighting Cells for Public School Graduates

For unlinked and linked weights for graduates from public schools, the following nesting cell structure was used to define nonresponse weighting cells, following what was done for NAEP 2005:

- SD/LEP status of graduate crossed with subject (SD and/or LEP math, SD and/or LEP science, no SD/LEP);
- school nonresponse cell;
- age of graduate (classed into "older" graduate and "normal age or younger" graduate);
- sex; and
- race (as given on the school administration form).

6.3.2.2 Preliminary Formation of Weighting Cells for Private School Graduates

For unlinked and linked graduate weights within private schools, the following nesting cell structure was used to define nonresponse weighting cells,[23] again following what was done for NAEP 2005:

- school nonresponse cell;
- age of graduate (classified into "older" graduate and "normal age or younger" graduate);
- sex; and
- race (as given on the school administration form).

6.3.2.3 Collapsing of Weighting Cells

When cells based on the nesting structure, (sections 6.3 and 2.2) were too small[24] to allow for stable nonresponse adjustment, cells were collapsed. The final graduate nonresponse weighting adjustments for unlinked and linked weights respectively are designated as $STNRADJ_d^{(U)}, \quad d = 1,...,D^{(U)}$ and $STNRADJ_d^{(L)}, \quad d = 1,...,D^{(L)}$ where $D^{(U)}$ is the total number of weighting cells designated for the unlinked weights and $(D^{(L)})$ is the total number of weighting cells designated for the linked weights.

6.3.2.4 Calculation of Nonresponse Adjustments

The nonresponse adjustments for graduates were computed as follows for unlinked weights:

$$STNRADJ_d^{(U)} = \frac{\sum_{sk \in S_d^{(U)}} STUWGT_{sk}}{\sum_{sk \in R_d^{(U)}} STUWGT_{sk}} \quad STNRADJ_d^{(L)} = \frac{\sum_{sk \in S_d^{(L)}} STUWGT_{sk}}{\sum_{sk \in R_d^{(L)}} STUWGT_{sk}}$$

[23] Logistic regression analysis was performed to identify the significant predictors, and only the significant predictors were used to form the nonresponse adjustment cells.

[24] Less than 20 sampled students or a replicate with less than 15 students, or an adjustment greater than 2.0, or a replicate adjustment greater than 1.5 times the full sample adjustment.

where $S_d^{(U)}$ ($S_d^{(L)}$) is the set of all eligible sampled graduates in unlinked (linked) weight cell *d*, $R_d^{(U)}$ is the set of all usable transcripts in unlinked weight cell *d*, and $R_d^{(L)}$ is the set of all linked transcript-assessment pairs in linked weight cell *d*. The replicate weighting adjustment $STNRADJ_d^{(U)}(r), \quad d = 1,...,D^{(U)}, \quad r = 1,...,R$, $STNRADJ_d^{(L)}(r), \quad d = 1,...,D^{(L)}, \quad r = 1,...,R$ are computed in a similar fashion with $STUWGT_{sk}(r)$ replacing $STUWGT_{sk}$ in the formulas above.

6.3.3 Weight Trimming

A similar trimming procedure was used for HSTS 2005 as was used for NAEP 2005. The weights for graduates were trimmed using the multiple median rule trimming procedure. The trimming procedure detects and truncates excessively large weights. Any weight within a given trimming group greater than a specified multiple of the median weight value of the given trimming group had its weight scaled back to that threshold. The same trimming factor calculated for the full sample weight was applied to each replicate weight within the same trimming group.

A multiple 3.5 was attempted as the cutoff factor. If too many weights are trimmed using this cut (e.g., more than 5%), 4.5 is used as the cutoff factor. The trimming group was defined by school type (public, the various types of private schools).

6.3.3.1 Unlinked Weights

For the unlinked weights, the median $MED_c^{(U)}$ of the nonresponse adjusted weights for the responding graduates within each trimming group c was calculated, then the trimming factor[25] was calculated as follows:

$$STRMADJ_{sk}^{(U)} = \begin{cases} \dfrac{4.5 \times MED_c^{(U)}}{STUNRWGT_{sk}^{(U)}}, & \text{if } (STUNRWGT_{sk}^{(U)} > 4.5^{*}\, MED_c^{(U)}), \\ 1, & \text{otherwise} \end{cases}$$

[25] For the unlinked weight trimming, the 4.5 median rule was used, and 47 cases were trimmed.

6.3.3.2 Linked Weights

For the linked weights, a factor was attached for whether the graduate had a mathematics NAEP assessment or a science NAEP assessment. This factor is designated as $STUMSO_{sk}$. The probability of being assigned a mathematics assessment given assignment to mathematics or science is 0.4081. Thus $STUMSO_{sk}$ for a mathematics assessment graduate is 2.45. The probability of being assigned a science assessment given assignment to mathematics or science is 0.5919, with a corresponding $STUMSO_{sk}$ of 1.69. Then the subject-specific linked weights are computed as follows:

$$SUBJWGT_{sk}^{(L)} = STUNRWGT_{sk}^{(L)} * STUMSO_{sk}$$
$$SUBJWGT_{sk}^{(L)}(r) = STUNRWGT_{sk}^{(L)}(r) * STUMSO_{sk}, \qquad r = 1,...,R$$

The median $MED_c^{(L)}$ of the subject-adjusted weights for the responding graduates within each trimming group was calculated, and then the trimming factor[26] was calculated as follows:

$$STRMADJ_{sk}^{(L)} = \begin{cases} \dfrac{3.5 \times MED_c^{(L)}}{SUBJWGT_{sk}^{(L)}}, & \text{if } (SUBJWGT_{sk}^{(L)} > 3.5 * MED_c^{(L)}), \\ 1, & \text{otherwise} \end{cases}$$

6.3.4 Calculating Final Weights for Graduates

The final graduate weights are calculated by multiplying the appropriate base weights, nonresponse adjusted weights, and trimming factors together:

$$FINWGT_{sk}^{(U)} = STUWGT_{sk}^{(U)} * STNRADJ_{sk}^{(U)} * STRMADJ_{sk}^{(U)},$$

$$FINWGT_{sk}^{(U)}(r) = STUWGT_{sk}^{(U)}(r) * STNRADJ_{sk}^{(U)}(r) \times STRMADJ_{sk}^{(U)}, \qquad r = 1,...,R$$

[26] For the linked weight trimming, the 3.5*median rule was used, and 9 cases were trimmed.

$$FINWGT_{sk}^{(L)} = STUWGT_{sk}^{(L)} * STNRADJ_{sk}^{(L)} * STRMADJ_{sk}^{(L)}, \text{ and}$$

$$FINWGT_{sk}^{(L)}(r) = STUWGT_{sk}^{(L)}(r) * STNRADJ_{sk}^{(L)}(r) \times STRMADJ_{sk}^{(L)}, \qquad r = 1,...,R$$

6.4 Variance Estimation

Replicate weights have been provided for each set of sample weights to allow users to compute variances for HSTS 2005 estimates. The particular method used for HSTS 2005 was the stratified jackknife assuming two primary sampling units (PSUs) per stratum (Krewski and Rao 1981), the same method used for the main NAEP 2005.

Graduate estimates based on HSTS 2005 are subject to sampling error because they are derived from a sample, rather than from the whole population. The variance is a measure of sampling error and, for the most part, determines the reliability of an estimate. Sampling variance indicates how much a population estimate for a given statistic would be likely to change if it were based on another equivalent sample of individuals drawn in exactly the same manner as the actual sample. Since HSTS 2005 used a complex sample design with several stages of sampling, unequal selection probabilities, and complex weighting procedures, use of standard textbook formulas or standard routines in software packages such as SAS and SPSS generally underestimate the true variance of survey estimates and should not be used.

6.4.1 Jackknife (JK2) Replication Method

The basic idea behind replication is to select subsamples repeatedly from the whole sample, calculate the statistic of interest for each subsample, and then use the variability among the subsample or replicate statistics to estimate the variance of the full sample statistic. Different ways of creating subsamples from the full sample result in different replication methods. The subsamples are called replicates, and the statistics calculated from these replicates are called replicate estimates.

The stratified jackknife replication method used for HSTS 2005, known also as the JK2 replication method, assumes that the population of PSUs, the first stage units, is grouped in L variance strata with two PSUs (or variance units) selected from each stratum. In the case of HSTS 2005, the first

stage units are the schools. In general, a replicate estimate is formed by randomly selecting one variance unit in a variance stratum. The weight of the selected variance unit is doubled, the weight of the nonselected variance unit is multiplied by zero, and the weights for the variance units in the remaining variance strata are not modified. This process is repeated for each variance stratum. If there are L variance strata, then L replicates are created.

The JK2 replication method, as well as any of the other replication methods, is implemented by using replicate weights. Each replicate weight corresponds to a given replicate. The estimated sampling variance of some statistic t is calculated by taking the sum of M squared differences (where M is the number of replicate weights developed):

$$\hat{V}ar(t) = \sum_{i=1}^{M} (t_i - t)^2$$

where t_i denotes the statistic of interest obtained using the i^{th} set of replicate weights and t denotes the statistic obtained using the set of full sample weights.

6.4.2 Calculating Replicate Weights

Replicate weights for a given HSTS 2005 sample were created by generating random samples of the original sample. In all, 62 replicate weights were created on each graduate record in an HSTS 2005 data set. Thirty-six replicates were designed to reflect the variance contribution arising from sampling PSUs (generally known as between-PSU variance). The remaining 26 replicates were designed to reflect the variance contribution arising from sampled schools within the 22 certainty PSUs (generally known as within-PSU variance). This variance replication scheme was the same one traditionally used for the national main NAEP 2005 assessment samples.

The creation of the 36 variance strata for the noncertainty PSUs involved pairing noncertainty PSUs in a manner that models a two PSU per stratum design in which PSUs are drawn with replacement. The HSTS 2005 samples used the main NAEP 2005 pairings, where PSUs were paired based on similar stratum characteristics. The 36 pairs of PSUs were formed by putting together PSUs from adjacent strata within NAEP region and metro status. Adjacent strata had similar socioeconomic characteristics such as proportion minority population, population change since 1980, per capita income,

civilian unemployment rate, educational attainment, and unemployment rate. Each PSU in a pair was randomly assigned to one of two different variance units (1 or 2). Each PSU pairing was referred to as a variance stratum, and each PSU in a variance stratum was referred to as a variance unit.

The procedure for creating the 26 variance strata for the certainty PSUs was analogous but somewhat more complex. The first stage units in certainty PSUs were schools, and so schools were paired to form variance strata under the JK2 model. For the 22 certainty PSUs in each HSTS 2005 sample, schools were listed in order of selection, and successive schools were paired within certainty PSUs. If there were an odd number of schools within a certainty PSU, the last three schools were grouped into a triple. Each school grouping was referred to as an initial variance stratum. Each school in a pair (or triple) was randomly assigned to 1 of 2 (3) different variance units [1, 2, (or 3)]. Since the number of initial variance strata greatly exceeded the desired number of variance strata (26), the initial strata were systematically assigned to 26 "combined" variance strata.[27] To distinguish between the two types of variance components, the 26 variance strata for the certainty PSUs were labeled 1 through 26, and the 36 variance strata for the noncertainty PSUs were labeled 27 through 62.

Replicate base weights (i = 1–62) for a graduate assigned to a variance stratum with two first-stage sampling units were calculated as below. *STU_BWT* was the graduate base weight for a given HSTS 2005 sample, as described in section 6.1, which reflected the various stages of selection.

$$STU\_BWT_{rep\_i} = \begin{cases} 0 & \text{if student is in variance unit 1 of variance stratum } i \\ 2 \times STU\_BWT & \text{if student is in variance unit 2 of variance stratum } i \\ STU\_BWT & \text{if the student is not in variance stratum } i \end{cases}$$

When a stratum contained three first-stage sampling units, graduates in the stratum had their weights adjusted for two sets of replicates. Replicate base weights (i = 1–62) for a graduate assigned to variance stratum with three first-stage units were calculated as follows:

$$STU\_BWT_{rep\_i} = \begin{cases} 0 & \text{if student is in variance unit 1 of variance stratum } i \\ 1.5 \times STU\_BWT & \text{if student is in variance unit 2 or 3 of variance stratum } i \\ STU\_BWT & \text{if the student is not in variance stratum } i \end{cases}$$

[27] Initial variance strata comprising three schools were assigned two variance strata so that two replicates are created for each of these strata. This is one common approach to handle three PSUs per stratum.

The final replicate weights for a given HSTS 2005 data set were calculated by applying the same weighting adjustment procedures described in section 6.1 to each set of replicate base weights. By applying the weighting procedures on each set of replicate base weights, variance estimates reflected the intended effects of the weighting adjustments.

This page intentionally blank.

7. GUIDE TO THE DATA FILES AND CODEBOOKS

This chapter describes the content and organization of the 2005 High School Transcript Study (HSTS 2005) data files and codebooks. It also details the process for accessing and obtaining the data files.

7.1 Public-use and Restricted-use Data Files

7.1.1 NAEP Transcript Data Explorer

For the first time, the National Assessment of Educational Progress (NAEP) HSTS data are accessible as public-use data through a web-based analysis tool. Researchers will be able to conduct interactive analyses on the NAEP HSTS 2005 data with the High School Transcript (HSTS) version of the NAEP Data Explorer (NDE). An adaptation of NDE, the NDE for the HSTS is a Data Analysis System (DAS) that enables users to access and analyze the graduate transcript data collected for HSTS 2005. Users can construct tables with as many variables as can be supported by the data, although data disclosure rules will place some limitations on the number of crossed variables, as well as the number of categories within the variables.

The NDE for HSTS is being made available to the public in two phases. In the first phase which has been released, researchers can generate tables of average NAEP assessment score tables for a number of independent variables, which include coursetaking and other transcript information, school and graduate demographic information, and the NAEP questionnaire responses. The data is limited to graduates who participated in both the NAEP assessment and HSTS and were eligible for inclusion in the transcript analysis. All analyses in the first phase use the NAEP scores as the default dependent variable. The NAEP-based data available in the NDE for HSTS are categorical or binary variables and can be used for user table requests.

The second phase of the NDE for HSTS will provide researchers with additional flexibility in variable selection for dependent variables, such as earned course credits and grade point average (GPA). Transcript data will be available from all graduates eligible for inclusion in the transcript analysis, regardless of their participation in NAEP. Phase II will also add 1990 and 2000 data to the NDE for

HSTS, so that users can examine trends in HSTS data. The NDE can be accessed at http://nces.ed.gov/nationsreportcard/nde/help/qs/About_NAEP_Data_Explorer.asp. When the second phase of the NDE for HSTS is available, it will be announced on this website.

7.1.2 Restricted-use Data

By Federal law, the schools and graduates that participated in HSTS 2005 are to remain confidential. However, all NAEP microdata files, including the NAEP HSTS 2005 data files, are available to users as restricted-use data files. Restricted-use data files contain variables for schools and graduates that cannot be released to the public, because of confidentiality concerns, but are made available to educational researchers. Though these data contain direct identifiers of schools, educational researchers using the HSTS 2005 data files must agree not to release any information that directly identifies a school or graduate, such as school name or address.

Because of confidentiality legislation, secondary users who wish to obtain a copy of the restricted-use data files must apply for an National Center for Education Statistics (NCES) restricted data license. If an organization does not already have a restricted data license, it is necessary to obtain a copy of the *Restricted-Use Data Procedures Manual*. There is a four-page checklist in this document that details the steps involved in obtaining a license. The manual may be viewed and downloaded from the NCES website at http://nces.ed.gov/statprog/rudman, or a copy may be requested from the following contact individual:

Cynthia Barton (202) 502-7307
cynthia.barton@ed.gov

If an organization already has a restricted data license, the organization may need only to have the license amended to add new datasets and/or authorized data users. Note that, in college or university settings, only faculty can serve as the primary project officer.

To obtain a restricted data license (or to amend an existing license), a secondary user should apply using the electronic registration process available at http://nces.ed.gov/statprog/instruct.asp.

7.2 Content and Organization of the Restricted-use Files

Data from HSTS 2005 were organized into eight data files:

- Course Offerings File
- School File
- Student File
- NAEP Data File
- SD/LEP Questionnaire File
- Tests and Honors File
- Transcript File
- Master CSSC File

Except for the Master CSSC File (which is not related to individual schools or graduates), all files can be linked by unique school identifiers. The Student, NAEP Data, SD/LEP Questionnaire, Transcript, and Tests and Honors files can be linked by unique student identifiers. The Master CSSC File can be linked to either the Course Offerings or the Transcript File by CSSC number.

Each file contains the appropriate weighting variables and replicate weights.[28] To obtain accurate results, users must select the appropriate weights for the type of analyses they are undertaking.

This section will provide an overview of the information available in each of the data files. More detailed information is available in the codebooks in appendices P to W.

7.2.1 Course Offerings File

The Course Offerings File is a complete listing of courses offered in all participating schools. Organized by school, each of the file's 169,864 records contains the following information:

- school ID;

[28] See Chapter 6 for a discussion of the appropriate weights to use.

- course title;
- course CSSC code;
- special education flag;
- the source of the catalog (e.g., generated from transcripts or from a school-provided catalog);
- the catalog type (whether the catalog is a district-level catalog, a school catalog, or a list of courses generated by the school);
- the location of the course (including various off-campus locations);
- the language of instruction;
- the level of the course (e.g., remedial, regular, enriched, honors); and
- whether it was part of an instructional sequence.

It should be noted that schools may not offer all courses that are on a transcript. For example, in a high school that covers grades 10 through 12, the grade 9 courses that graduates took in junior high school were not treated as transfer courses but appeared as if they were offered by the high school. This treatment provides a more balanced picture of the courses available to graduates in four years of high school than would be provided by treating such courses as transfer courses. For the 55 schools from which no catalogs were received, the list of unique course titles appearing on the sampled transcripts was the only available source of course offering entries.

7.2.2 School File

The school file contains one record for each of the participating schools. The file includes school variables gathered on the School Information Form (SIF) during the transcript study, as well as the school's responses to the NAEP School Questionnaire. Copies of the questionnaires are in appendices K, L, and M. Information collected on the SIF that appear on the HSTS 2005 school file include diplomas offered, school programs offered, the typical number of classes per school day, and the typical length of time for school classes.

7.2.3 Student File

The student file contains a record for each of the high school graduates who were identified for HSTS 2005. Each record in the file contains demographic information, sampling information, graduate weights, and replicate weights for variance estimation. It also contains a flag indicating whether or not the graduate was disabled (SDSTATUS) and a variable indicating the specific nature of the disability when applicable (HCTYPE).[29] The file also contains a series of derived variables, including summaries of the graduate's coursetaking record by major educational topic, as taken from the graduate's transcript data. Because a number of transcripts for graduates were not received or were incomplete, only 26,151 graduates have full transcript information on their graduate records.

7.2.4 NAEP Data File

This file contains data from the NAEP 2005 mathematics and science assessments for the HSTS sampled graduates. Because NAEP scores are designed to provide accurate group estimates rather than student-level information, plausible values for graduates are developed. These plausible values variables are "conditioned" on other variables (e.g., parents' education level and NAEP region) in the NAEP datasets. These plausible values provide more unbiased estimates of graduate scores when NAEP data are analyzed in conjunction with the conditioning variables.[30] The NAEP data file includes the plausible values for NAEP proficiency scores for each 2005 high school graduate who participated in a NAEP assessment in a school that was fully linked to HSTS 2005.

In addition to the variables used to estimate plausible values for the main NAEP study, the following transcript study variables included in the student file were used in the conditioning process:

- ACADTRK Student Program
- CLRANK/CLSIZE Class Rank divided by Class Size

[29] The values of the disabling condition codes in 2005 are 00 = Multidisabled, 01 = Learning Disabled, 02 = Hearing Impaired, 03 = Visual Impaired, 04 = Speech Impaired, 05 = Mental Retardation, 06 = Emotional Disturbance, 07 = Orthopedic Impaired, 08 = Traumatic Brain Injury, 09 = Autism, 10 = Developmental Delay, 11 = Other Health Impaired, 12 = Other, 88 = Not Reported. These codes have been modified since the codes used in HSTS 1998 file.

[30] The plausible value estimation process for NAEP is explained at http://nces.ed.gov/nationsreportcard/pubs/guide97/ques11.asp and in the NAEP technical report for 1996 (http://nces.ed.gov/nationsreportcard/pdf/main1996/1999452b.pdf. Also, see the forthcoming online NAEP 2005 technical report for a detailed discussion of conditioning.

- EXITSTS — Student Exit Status
- TGPA — Calculated GPA
- GRREQFLG — Graduation Requirements Level Flag
- SDSTATUS — Student Disability Status
- LEP — Student Limited English Proficiency Status
- CENSREG — Census Region
- STUB0100 - STUB1600 — These "stub" variables represent the number of credits graduates received in various subject areas. These are defined in detail in appendix O.
- STUB2001 - STUB2005 — New Basics Curriculum categories. These variables represent variants of academically oriented course taking patterns recommended in the *Nation at Risk* report. They are defined in detail in appendix O.

Because of the inclusion of the transcript study variables, the NAEP scores reported in the HSTS files are slightly different from the scores contained in the records for the same graduates distributed solely as NAEP data. The overall national scores from the two studies are marginally different.

If the need arises to match transcript study records with records obtained from NAEP files obtained from other sources, the user will need to take into account the differences in naming conventions for the school and student IDs noted in table 20.

Table 20. HSTS and NAEP record identifier naming conventions: 2005

HSTS transcript study record identifier		NAEP record identifier (other than those distributed with the transcript files)	
Variable name	Field length	Variable name	Field length
SCHOOLID	7	SCHID	7
STUDENTI	10	BOOK	3
		BKSER	6
		CHKDIG	1

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, NAEP High School Transcript Study (HSTS), 2005.

The student identifier in the transcript study, STUDENTI, is created by concatenating the NAEP book number (BOOK, which identifies the form of the assessment which was administered), the book serial number (BKSER), and the check digit (CHKDIG).

7.2.5 SD/LEP Questionnaire File

The SD/LEP questionnaire file contains a record for each of the approximately 3,000 HSTS sampled graduates with student disability and/or limited English proficiency questionnaire data. The file contains all data from the completed questionnaires, including the demographic variables asked for on the cover of the questionnaires.

7.2.6 Test and Honors File

The test and honors file contains information on standardized test scores and honors that appears on high school transcripts. Of the transcripts collected, about 9,400 transcripts (approximately one-third) contained either standardized test scores or notations regarding honors and awards that graduates received. Transcripts without this information may belong to graduates who did not take standardized test scores and/or received no honors; however, they may also be for graduates attending schools that did not report some or all of this information on transcripts. Because of the relatively small percentage of transcripts represented and the uncertainty about the source of missing data, the data in this file should be used with caution.

Graduates in the Test and Honors File are identified by the combination of school and graduate ID variables. Each test or honor entry on a transcript is identified with a unique sequence number. The combination of graduate ID and test/honor sequence number allows for a unique ID number for each test or honor within the file. Each entry also contains an indicator of the record type ("T" = test, "H" = honor), the month and year of the test or honor (if available), and a 50-character description of the honor or the test.

For most tests, scores were provided; however, it was not always possible to give meaningful entries for some test scores. The subtests that are reported also varied tremendously. Complete scores are provided for the Preliminary Scholastic Aptitude Test (PSAT) math and verbal subtests, the Scholastic Aptitude Test (SAT) math and verbal subtests, and the American College Test (ACT) composite subtests that appeared on the transcripts. The remaining test information is less complete. The file contains 46,680 records.

7.2.7 Transcript File

The Transcript File contains a record for each course appearing on the sampled graduates' transcripts. It is an extremely large file, containing over 1.3 million records. Courses are uniquely identified by a course ID number. Each course record includes the following variables:

- student ID number;
- grade level when course was taken;
- school year when course was taken;
- school term when course was taken;
- course title;
- grade received (original and standardized);
- credits received (original and standardized Carnegie units);
- course Classification of Secondary School Courses (CSSC) code;
- whether the course was a special education course;
- whether the course was taught off campus;
- whether the course was taught in a language other than English;
- instruction level of the course; and
- whether the course was a transfer course.

The analyst may wish to use this file to obtain new summary variables for graduates to add to the student file.

7.2.8 Master CSSC File

The Master CSSC File contains all codes in the modified version of the Classification of Secondary School Courses (CSSC) used in this study. The CSSC is described in chapter 1, and additional information on the codes is included in appendix O.

The Master CSSC File is organized by the CSSC code and contains four variables:

- CSSC course code (described in chapter 1);
- special education flag (described in chapter 5);
- standard course title; and
- sequence flag (described in chapter 5).

7.3 Additional Information for Researchers Wishing to Use Restricted-use Files

The HSTS data files contain a wealth of education-based information for researchers to use to understand issues related to coursetaking, access to courses, and achievement. This section addresses some topics that were not addressed in preceding sections such as the use of NAEP scores for individuals.

7.3.1 Selecting the Proper Weights

As discussed in chapter 6, there are multiple weights associated with HSTS 2005. Selecting the appropriate weight to use in analyses involving HSTS 2005 is critical for ensuring accurate results. [31]

Users' wishing to estimate variances as well as point estimates should be aware of the importance of using replicate weights[32] with HSTS 2005. Since HSTS 2005 used a complex sample design with several stages of sampling, unequal selection probabilities,[33] and complex weighting procedures, use of standard textbook formulas or standard routines in software packages such as SAS and SPSS generally underestimate the true variance of survey estimates and should not be used.

[31] Section 6.1 describes the weights to be used for different type of analyses.

[32] See section 6 for additional information on the replicate weights.

[33] See chapters 2 and 3 for more information on the sampling used in HSTS.

7.3.2 Statistical Software for Use with HSTS 2005

Specialized software is required to produce the appropriate statistics from the HSTS 2005 data due to the complex sample design reflected in the jackknife replicate weights and the plausible values of the NAEP scale scores. Standard SAS and SPSS code can produce accurate point estimates but cannot easily produce correct standard errors.

The International Association for the Evaluation of Educational Achievement (IEA) has developed SAS and SPSS macros to work with similar kinds of jackknife replicate weight datasets found in international educational assessments like Third International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS). These files need to be modified for use with the HSTS 2005 data files. The programs can be downloaded from the international website at the IEA International Study Center at Boston College (http://isc.bc.edu/timss2003i/userguide.html).

Commercial software such as WesVar can also be used for analyzing the HSTS data (http://www.westat.com/wesvar/). Other commercially-available software includes SUDAAN v9 (http://www.rti.org/sudaan/) and STATA v9 (http://www.stata.com/).

7.3.3 Use of NAEP Scores for Individuals

The design of the NAEP studies does not allow reporting on the performance of individual students. Rather it assesses student performance in selected academic areas for specific populations of students or subgroups of these students. The NAEP sample includes students from both public and private schools. To maximize student participation, NAEP policy states that a student should be asked to participate in the assessment, unless their inability to do so can be clearly established. Beginning with the 2000 assessment, NAEP HSTS linked analyses have included graduates who took the assessments with accommodations because they had disabilities or were LEP students.

NAEP Scale Scores

Because of the design of the NAEP assessments, each student typically responds to only a few questions within any content area, and not all students are asked the same questions. Unlike many

traditional assessments, there is no linear transformation between correct/incorrect items and a single score. Using a single student-level score would result in misleading estimates of population characteristics. Instead, NAEP constructs sets of plausible values (in sets of five) designed to represent the distribution of performance in the population for each subject assessed. A plausible value is a representative value from the potential scale scores for all students in the population with similar characteristics and identical patterns of item response. Because HSTS collects additional information about the student characteristics and item responses that can be used in this estimation process, plausible values for NAEP scale scores are recalculated for the HSTS sample for use in analyses relating NAEP scores and HSTS transcript data. As a result, NAEP scale scores associated with the HSTS 2005 data differ slightly from NAEP scale scores associated with NAEP 2005 student data.

Since the statistics describing the performance on the NAEP mathematics and science scales are based on the plausible values, the statistical software used to conduct these analyses must properly compute the statistics for the plausible values.

More information about NAEP 2005, including scale scores, plausible values, and jackknife variance replication can be found in the forthcoming online NAEP 2005 technical report.

7.4 HSTS Analysis Reports

Data collected by HSTS offers researchers a unique glance into graduate coursetaking patterns from one study year to the next. Many of the analyses done to date can be found in the publications located on the HSTS website (http://nces.ed.gov/nationsreportcard/hsts/).

For HSTS 2005, the initial release report, *The Nation's Report Card, America's High School Graduates, Results from the 2005 NAEP High School Transcript Study,* provides analyses on course credits, grades, and NAEP achievement. The analyses look at graduates based on gender, race/ethnicity, parent education, and performance over time. The analyses discussed in the HSTS 2005 initial release report represented high school graduates with complete transcripts. Students whose transcripts did not include course-by-course data for at least 3 full years of high school were excluded. To be consistent with

other published analyses, the following rules were adopted for including and excluding students in the analyses that produced the tables:[34]

1. Both public and private school graduates were included.

2. Students with special education diplomas, certificates of attendance, and certificates of completion were excluded. Certificates of completion indicate that a student completed the necessary school requirements for graduation, but failed to successfully complete a required state graduation exam.

3. Graduates with disabilities who received regular or honors diplomas (i.e., those who were not screened out by rule 2) were included.

4. Graduates with fewer than 16 Carnegie Units were excluded. A Carnegie Unit was a factor used to standardize all credits indicated on transcripts across the study. The Carnegie Unit is defined as the number of credits received for 120 hours of classroom instruction over the course of a year.

5. Graduates with zero English credits were excluded.

Prior to finalizing the data file, transcript records were subject to quality control procedures that listed transcripts that needed to be examined because the transcript records were inconsistent with the student's exit status. In a few cases, it was determined that a student initially recorded as a graduate had not actually graduated, and the student's exit status was revised accordingly. Among students with transfer courses, it was sometimes determined that, although a student had fewer credits than were required to graduate, the transcript had all the other attributes of a graduated senior. These attributes included student exit status, graduation date, GPA, and class standing. Credits from transfer schools may not have been recorded on the transcript, or the transferred credits may have had a different credit assignment than the school of graduation. In these cases, if a careful review of the transcript and the data files showed no data entry or coding errors, and the lack of credits resulted from missing or improperly converted Carnegie credits for the transfer courses, the record was updated. An additional transcript record with undifferentiated credit was added, or the existing transfer credit records were modified to assign the actual number of credits the graduate had taken.

In summary, for a transcript to be included in the analyses in the initial release report, it had to meet three requirements: (1) the graduate graduated with either a standard or honors diploma, (2) the graduate's transcript contained 16 or more Carnegie credits, and (3) the graduate's transcript contained

[34] An exception to this is that the 2005 initial release report contains an analysis of graduates with disabilities that included those graduates receiving special education diplomas and certificates of attendance in addition to those receiving honors or standard diplomas.

more than 0 Carnegie credits in English courses. These additional restrictions reduced the number of 2005 graduates in sample used in the report from 27,051 graduates to 26,525 graduates.

For HSTS 2000, there are two publications containing many comparisons and analyses. The first publication, *The High School Transcript Study: A Decade of Change in Curricula and Achievement, 1990–2000*, is a printed report available from the National Center for Education Statistics via its website (http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2004455) and EDPUBS. This report analyzes the changes in course credits earned and GPAs achieved by high school graduates from HSTS 1990 to HSTS 2000. It also looks at correlation values between the NAEP 2000 mathematics and science assessment scores with various student coursetaking variables. The second publication, *The 2000 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 2000, 1998, 1994, 1990, 1987, and 1982 High School Graduates*, is available on the NCES website, NCES Publication #2007463. It details the number of credits earned by high school graduates in various school subject fields and by various school and graduate characteristics, including gender, race/ethnicity, academic track, type of locale, school type (public/private), and region of the country. It also contains tables covering graduation requirements, grade point averages, and NAEP 2000 mathematics and science assessment scores.

The HSTS 2005 datasets offer new possibilities for data analyses that previous HSTS datasets could not offer. Researchers can analyze relationships between the mean NAEP mathematics and science assessment scores by whether or not graduates took selected mathematics or science courses. Incorporating the HSTS 2005 datasets with the previous HSTS datasets, researchers can track courses by grade level across the transcript studies to determine whether course curricula have changed in the past 2 decades. Linking the HSTS files with the corresponding NAEP student questionnaires provides new educational-related variables for data analysis, including parents' education levels, computer usage at home and school, and time spent on homework.

This page intentionally blank.

REFERENCES

Cochran, W.G. (1977). *Sampling Techniques*, 3rd Ed. New York: Wiley.

Frankel, M.R. (1971). *Inference from Survey Samples: An Empirical Investigation*. Ann Arbor, MI: Institute for Social Research.

Gifford, A., Hoachlander, E., and Tuma, J. (1994). *The Secondary School Taxonomy Final Report.* U.S. Department of Education, National Assessment of Vocational Education. Final Report to Congress, Volume II, Chapter 1.

Ingels, S., Dowd, K., L., Taylor, J., Bartot, V., Frankel, M., and Pulliam, P. (1995). *National Education Longitudinal Study of 1988, Second Follow-Up: Transcript Component Data File User's Manual* (NCES 95–377). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Jones, C., Knight, S., Butz, M. Crawford, I., and Stephenson, B. (1983). *High School and Beyond Transcripts Survey (1982): Data File User's Manual* (NCES 84-205). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Kish, L., and Frankel, M. (1974). Inference from Complex Samples. *Journal of the Royal Statistical Society, Series B,* 36, 1-22.

Krenzke, T., Rust, K., Wallace, L., and Yansanah, I. (2000). *1998 NAEP Sampling and Weighting Report.* Rockville, MD: Westat.

Krewski, D., and Rao, J. (1981). Inference from Stratified Samples: Properties of Linearization, Jackknife and Balanced Repeated Replication. *Annals of Statistics,* 9, 1010-1019.

Lee, E.S., Forthofer, R. N., and Lorimer, R.J. (1989). *Analyzing Complex Survey Data*. Newbury Park, CA: Sage Publications.

Legum, S., Caldwell, N., Goksel, H., Haynes, J., Hynson, C., Rust, K., and Blecher, N. (1993a). *The 1990 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1990, 1987, and 1982 High School Graduates* (NCES 93–423). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Legum, S., Caldwell, N., Goksel, H., Haynes, J., Hynson, C., Rust, K., and Blecher, N. (1993b). *The 1990 High School Transcript Study Technical Report* (NCES 93–117). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Legum, S., Caldwell, N., Goksel, H., Haynes, J., Hynson, C., Rust, K., and Blecher, N. (1993c). *User's Manual: 1990 High School Transcript Study* (NCES 93–416). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Legum, S., Caldwell, N., Davis, B., Haynes, J., Hill, T., Litavecz, S., Rizzo, L., Rust, K., and Vo, N. (1997). *The 1994 High School Transcript Study Technical Report* (NCES 97–262). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Legum, S., Caldwell, N., Davis, B., Haynes, J., Hill, T., Litavecz, S., Rizzo, L., Rust, K., and Vo, N. (1998). *The 1994 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1994, 1990, 1987, and 1982 High School Graduates* (NCES 98–532). 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: Wiley.

Ludwig, M.J., Olivetti, L., Sandberg, N., and Waite, B. (1982). *A Classification of Secondary School Courses.* Alexandria, VA: Evaluation Technologies, Inc.

Morgan, R., Hunt, E., and Carpenter, J. (1991). *Classification of Instructional Programs* (NCES 91–396). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

National Assessment of Vocational Education. (1994). *Final Report to Congress, Volume II,* Chapter 1. Washington, DC: U.S. Government Printing Office.

National Commission on Excellence in Education. (1983). *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: U.S. Government Printing Office.

Oh, H.L., and Scheuren, F.S. (1983). Weighting Adjustments for Unit Nonresponse. In W.G. Madow, I. Olkin, and D.B. Rubin (Eds.), *Incomplete Data in Sample Surveys, Vol. II: Theory and Annotated Bibliography*. New York: Academic Press.

Perkins, R., Kleiner, B., Roey, S., and Brown, J. (2004). *The High School Transcript Study: A Decade of Change in Curricula and Achievement*, *1990–2000* (NCES 2004–455). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Perkins, R., Roey, S., and Brown, J. (2005). *The 2000 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 2000, 1998, 1994, 1990, 1987, and 1982 High School Graduates* (upcoming NCES 2005–482). U.S. Department of Education, National Center for Education Statistics.

Roey, S., Caldwell, N., Rust, K., Blumstein, E., Krenzke, T., Legum, S., Kuhn, J., Waksberg, M., and Haynes, J. (2001a). *The 1998 High School Transcript Study User's Guide and Technical Report* (NCES 2001–477). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Roey, S., Caldwell, N., Rust, K., Blumstein, E., Krenzke, T., Legum, S., Kuhn, J., Waksberg, M., and Haynes, J. (2001b). *The 1998 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1998, 1994, 1990, 1987, and 1982 High School Graduates*

(NCES 2001–498). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Roey, S., Caldwell, N., Rust, K., Hicks, L., Lee, J., Perkins, R., Blumstein, E., and Brown, J. (2005). *The 2000 High School Transcript Study User's Guide and Technical Report* (NCES 2005–483). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Shettle, C., Roey, S., Mordica, J., Perkins, R., Nord, C., Teodorovic, J., Brown, J., Lyons, M., Averett, C., Kastberg, D. (2007). *The Nation's Report Card: America's High School Graduates Results from the 2005 NAEP High School Transcript Study*, (NCES 2007-467). U.S. Department of Education, Washington, DC National Center for Education Statistics.

Thorne, J., Burke, J., Rust, K., Marshall, R., Caldwell, N., Sickles, D., and Ha, P. (1988). *High School Transcript Analysis: The 1987 Graduates*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Thorne, J., Burke, J., Rust, K., Marshall, R., Caldwell, N., Sickles, D., and Ha, P. (1989). *1987 High School Transcript Study Technical Report.* Rockville, MD: Westat.

Thurgood, L., Walter, E., Carter, G., Henn, S., Huang, G., Nooter, D., Smith, W., Cash, R.W., and Salvucci, S. (2003). *NCES Handbook of Survey Methods* (NCES 2003–603). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Wolter, K. (1985). *Introduction to Variance Estimation.* New York: Springer-Verlag.

Appendix A

Disclosure Notice

APPENDIX A. DISCLOSURE NOTICE

2005 HIGH SCHOOL TRANSCRIPT STUDY

"A copy of this student's transcript____ will be ____ has been provided to WESTAT, agent for the U.S. Department of Education, National Center for Education Statistics (NCES). The granting of Education Department authority for collection of the transcript data has been made pursuant to the provisions of the Family Education Rights and Privacy Act (FERPA) (20 U.S.C. 1232g) as implemented by 34 CFR 99.31(a)(3)(ii) and 99.35, summarized on the back of this notice. This disclosure statement fulfills the requirements of provision 34 CFR 99.32 of FERPA.

The High School Transcript Study (HSTS), sponsored by NCES, is being conducted to collect information on current course offerings and course taking in the nation's secondary schools. This student has been selected to participate in HSTS, and data from these records will be combined with other into statistical summaries and tables. No individually identifiable information will be released in any form."

Appendix B

2005 HSTS School Information Form (SIF)

2005 HSTS
School Information Form (SIF)

Instructions to Supervisors on filling out the School Information Form (SIF)

1. The SIF is in two parts. Part 1 is to be filled out over the phone with the HSTS School Coordinator. You will call the School Coordinator and tell him/her about HSTS and go through Part 1 of the form and set up a day to visit the school.

2. Part 1 of the SIF is scripted, but you may deviate from the script to address specific aspects of the study, answer questions in detail, etc.

3. One section of the SIF script is to ask the School Coordinator the name of a person at the school who is very knowledgeable about the content of the course catalog, the transcript layout, and details about classes offered at the school. Often, this person is the registrar. After you get this person's name, you should call him/her to let him/her know that you will be at the school and would like to spend some time with him or her obtaining detailed information on the school's courses. Tell him/her that you will send a questionnaire so that he/she will know what types of questions to expect when you visit. This questionnaire will contain some, but not all of the questions contained in Part 2 of the SIF.

4. You will also be asking the HSTS coordinator to send you a copy of the 2004-2005 catalog before your visit to the school, if a catalog has not already been received at Westat. Assuming that the coordinator sends you the catalog, please review the catalog and fill out as many of the Part 2 SIF questions from the catalog as you can in order to minimize the burden on the school.

5. You will visit the school on the agreed-upon date and you will meet with the School Coordinator to complete information on the SIF (such as where the administration schedule will be kept, whether the School Coordinator works in the summer, etc.) and you will obtain sample transcripts.

6. After meeting with the School Coordinator, you will meet with the registrar or other person knowledgeable about school's courses. You will spend some time filling out the remainder of Part 2 of the SIF.

Part 1 of the 2005 HSTS School Information Form (SIF)

Territory: _____ Region:_____ Area:_____
NAEP School ID:_____ NAEP Supervisor:_____ NAEP AC:_____

Materials collected prior to supervisor's call: _____
_____

School Name:_____ School phone: _____
City, State:_____ Fax: _____
Principal:_____ Email: _____
School Coordinator (SC):_____ Web Site: _____
SC Phone Number:_____ Grade Range of School: _____
Is School Participating in NAEP?_____ Assessment Date: _____

Hello, this is [FIRST NAME, LAST NAME] from the National Assessment of Educational Progress or NAEP. I'm calling to tell you about another aspect of NAEP. It is the High School Transcript Study or HSTS. The HSTS is being conducted to provide educational policy makers with information regarding current course offerings and course-taking patterns in secondary schools. This study will also permit researchers to examine the relationship between course-taking patterns and educational achievements. We would like to designate a School Coordinator for this study, someone who is knowledgeable about the courses offered at your school.

1. Would you be the contact person or the School Coordinator for HSTS?

   Yes [GO TO #2] No

   If "No" Who will be the HSTS contact at the school? And at what phone number can I reach him/her?

   Name:_____Telephone number:_____

   Thank you so much for your help. [END CALL]

2. I'd like to explain a little bit about the HSTS to you. Do you have about 10 minutes to talk right now?

   Yes [GO TO #3] No

   If "No" When is the best day and time for me to reach you to discuss this study? Again, I just need about 10 minutes.

   Date:_____ Time:_____

   Thank you so much. I will talk to you soon. [END CALL]

3. There are two phases to the High School Transcript Study. During phase 1 in the month of March, I would like to visit your school to collect information about your school, including course catalogs and three sample transcripts. For phase 2, in the summer, I or another HSTS representative will return to the school to obtain transcripts of those students who were selected for the math or science portion of NAEP. Absolutely no student time is involved in the study and confidentiality of the materials we collect will be strictly maintained. You will be reimbursed for all transcripts. **[IF 2004-2005 CATALOG WAS NOT SENT TO WESTAT:** At this time I'd like to collect the current course catalog for your school. Ideally, the catalog should contain all courses offered at the school including honors, vocational, remedial, special education, and off-campus courses. Our preference would be to obtain a school-level catalog with course names and content descriptions, if such a catalog is available.]

[IF CATALOG HAS NOT BEEN SENT TO WESTAT] Are copies of the current year's school-level catalog available?

Yes [GO TO #4] No

If "No" When will they be available? Date: _____

[If no school-level catalog available, ask about obtaining another type of catalog. Please check which type of catalog is available. The order of preference with the most desirable type of catalog listed first is:

___school-level catalogs that provide course names and content descriptions
___district-level catalogs that provide course names and content descriptions for this particular school clearly marked
___course list by department that includes general descriptions of course offerings by department
___course lists without content descriptions
___district-level catalogs without school-level identification

4. **[IF CATALOG HAS NOT BEEN SENT TO WESTAT]** In what format or formats is the most current catalog available – In hardcopy, on a website, or in an electronic file?

In hardcopy [GO TO #4a] On a web site [GO TO #4b] In an electronic file [GO TO #4C]

4a. If I send you a pre-addressed envelope, would you please send me a copy of the current course catalog?

Yes [ASK FOR ADDRESS TO MAIL ENVELOPE TO] Address: _____
_____

4b. What is the web site address that lists the catalog? _____

4c. Would it be possible for you to email me the electronic file of the catalog? My email address is [EMAIL ADDRESS]

5. I would also like to obtain copies of the three prior years' catalogs. Again, the catalogs should contain all courses offered at the school including honors, vocational, remedial, special education, off-campus courses, distance learning courses, and ESL courses. Our preferences, if available would be school-level catalogs with course names and content descriptions. Do you think I will be able to obtain copies of the 2003-2004, 2002-2003, and 2001-2002 catalogs when I visit?

Yes No or I don't know Other (only certain years, etc.):_____

5a. In what format or formats are these catalogs available – in hardcopy, on a web site, or in an electronic file?

[CIRCLE YEARS CATALOG IS IN THIS FORMAT]

In hardcopy:	2003-2004	2002-2003	2001-2002
On a web site:	2003-2004	2002-2003	2001-2002
In an electronic file:	2003-2004	2002-2003	2001-2002

6. As I mentioned earlier, when I visit your school, I would also like to obtain three sample transcripts for students who have already graduated. Before I leave the building I will be removing identifying information from the transcripts. The sample transcripts should reflect one with regular courses, one with honors courses, and one with special education courses. If there are other special programs offered at the school (IB, performing arts, etc.), I'd appreciate seeing transcripts that include these programs [these could be the same three transcripts or different ones]. The transcripts should also include grades for course taken. Will I be able to obtain such transcripts when I visit?

Yes No

(NOTE to supervisor: if you find a school that cannot provide transcripts that meet these criteria (e.g. the school uses a standards-based transcript), please probe to see if the information can be obtained in some other fashion, If not, please notify your field manager as soon after completion of the form as possible, so that a decision about the feasibility of including this school in HSTS can be made.)

Comments: _____
_____
_____
_____
_____

7. In addition to collecting the course catalogs and transcripts, I will need to spend some time with someone at your school who is very knowledgeable about the details contained in the catalogs, the transcripts and related school information. If you are not that person, can you tell me the best person to talk to about this information? I would like to give him or her a call in advance of my visit and send him or her some information so he/she knows what to expect when I visit.

Name:_____ Telephone Number: _____

8. After graduation, we will return to obtain transcripts of the 12th grade students who were selected for NAEP mathematics or science. No student time is involved, confidentiality is strictly maintained, and you will be reimbursed for all of the transcripts copied. During my initial visit, I would also like to insert disclosure notices as markers in these 12th grade student files.

Comments: _____
_____
_____
_____
_____

9. Our normal procedures for the transcript study, as specified by FERPA, is to provide FERPA notices for each sampled student's file, but not to notify parents of their child's inclusion, because no student time is involved and all transcript information is collected anonymously. Is there any reason that we should use different procedures in your school?

Yes No

[PARENT INFORMATION AND CONSENT LETTERS ARE AVAILABLE IF THE SCHOOL INSISTS ON HAVING PARENTS LETTERS]

Comments: _____
_____
_____
_____
_____

10 Would [date in March] date work for you for me to visit your school? Remember that I will need to spend some time with you or [name of Registrar].

Yes [GO TO #10a] No

If NO, what date would work for you in March? _____

10a. What time should I arrive?_____

11. As part of the HSTS this year, we are collecting information about what textbooks are being used in high school mathematics and science courses. I will mail you these forms in advance of my visit and will ask you to fill them out or have someone else, such as the chairpersons for the mathematics and science departments, fill them out and then fax them to the phone number on the form.

12. Thank you so much for speaking with me today. I will be sending you a summary of the High School Transcript Study, the Textbook forms, as well as envelope for you to send me your catalog (if applicable). I will also be calling [REGISTRAR'S NAME] to let him/her know what kind of questions to expect from me when I visit your school. Good bye.

[END CALL]

School ______________________________ State ________ ID#________________
Please return this form to:____________________by_____________ or fax to: 240-314-2381

Mathematics Textbook Form

Please list the textbooks used for all mathematics courses in your school and explain how the textbooks are used. If you have an existing list of textbooks, you may send it to us. Please add any information below that is not already included on your list.

Course: ____

ISBN: ____________________
Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ____________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ______________________________

Use of Textbook: ___ Major ___ Supplementary

Course: ____

ISBN: ____________________
Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ____________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ______________________________

Use of Textbook: ___ Major ___ Supplementary

Course: ____

ISBN: ____________________
Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ____________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ______________________________

Use of Textbook: ___ Major ___ Supplementary

Mathematics Textbook Form (continued)

Course: _____

ISBN: _____
Textbook Title _____

Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

Course: _____

ISBN: _____
Textbook Title _____

Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

Course: _____

ISBN: _____

Textbook Title _____

Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

School _____ State _____ ID#_____
Please return this form to:_____by_____ or fax to: 240-314-2381

Science Textbook Form

Please list the textbooks used for all science courses in your school and explain how the textbooks are used. If you have an existing list of textbooks, you may send it to us. Please add any information below that is not already included on your list.

Course: _____
ISBN: _____

Textbook Title _____
Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

Course: _____

ISBN: _____
Textbook Title _____
Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

Course: _____
ISBN: _____
Textbook Title _____
Author(s)/Editor(s) _____

_____

Edition _____Publisher: _____ Copyright: _____

Chapters covered in course: All ___ Chapters Covered: _____

Use of Textbook: ___ Major ___ Supplementary

Science Textbook Form (continued)

Course: ____

ISBN: ____________________

Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ______________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ________________________________

Use of Textbook: ___ Major ___ Supplementary

Course: ____

ISBN: ____________________

Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ______________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ________________________________

Use of Textbook: ___ Major ___ Supplementary

Course: ____

ISBN: ____________________

Textbook Title ____

Author(s)/Editor(s) ____

____

Edition ____________Publisher: ______________________________ Copyright: ___________

Chapters covered in course: All ___ Chapters Covered: ________________________________

Use of Textbook: ___ Major ___ Supplementary

Part 2 of the 2005 HSTS School Information Form (SIF)

Territory: _____ Region:_____ Area:_____
NAEP School ID:_____ NAEP Supervisor:_____ NAEP AC:_____

Materials collected prior to supervisor's call: _____
_____

School Name:_____ School Phone: _____
City, State:_____ Fax: _____
Principal:_____ Email:_____
School Coordinator (SC):_____ Web Site: _____
SC Phone Number:_____ Grade Range of School: _____
Is School Participating in NAEP?_____ Assessment Date: _____

A. Detailed Information on Course Catalogs

1. Which type(s) of catalogs were obtained? [CIRCLE YEARS CATALOG IS IN THIS FORMAT]

   School-level catalogs that provide course names and content descriptions
   2004-2005 2003-2004 2002-2003 2001-2002

   District-level catalogs that provide course names and content descriptions for this particular school clearly marked
   2004-2005 2003-2004 2002-2003 2001-2002

   Course list by department that includes general descriptions of course offerings by department
   2004-2005 2003-2004 2002-2003 2001-2002

   Course lists without content descriptions
   2004-2005 2003-2004 2002-2003 2001-2002

   District-level catalogs without school-level identification
   2004-2005 2003-2004 2002-2003 2001-2002

2. Does this school include 9th grade?

   Yes (GO TO #2b.) No

2a. If no, where do most students attend 9th grade?

___ A single feeder Junior High/Middle School

___ Several Junior High/Middle Schools in the district

___ Other schools not in this district or affiliated with this school

2b. Do the high school catalogs also contain information about the 9th grade?

Yes (GO TO #4) No

3. How can I obtain copies of catalogs with information about the 9th grade courses given by the feeder school in 2001-2002? (NOTE: If 2001-2002 catalogs are not available, obtain the oldest available catalogs, after 2001-2002. If a school has a large number of feeder schools, it may not be practical to obtain all of the 9th grade catalogs. In this case, please try to obtain the district-level catalogs.

_____
_____
_____
_____

4. Do the catalogs obtained cover all the courses available for the class of 2005 during their years at this school. (Include 9th grade courses if taken a junior/middle school.)

Yes No

(If no, obtain any additional catalogs covering the omitted information)

5. Do the catalogs include the following course offerings?

5a. Vocational courses (circle answer)

Yes No

If yes, how are vocational courses indicated in the catalog(s): _____
_____

5b. Remedial courses (circle answer)

Yes No

If yes, how are remedial courses indicated in the catalog(s): _____
_____

5c. Honors courses (circle answer)

Yes No

If yes, how are honors courses indicated in the catalog(s):_____
_____

5d. Special Education courses (circle answer)

Yes No

If yes, are different levels of special education (e.g. resource and self-contained) courses indicated in the catalog(s) and how are they indicated:_____
_____

5e. Off-campus courses (circle answer)

Yes No

If yes, how are off-campus courses indicated in the catalog(s): _____
_____

5f. ESL or bilingual courses (circle answer)

Yes No

If yes, how are ESL or bilingual courses (courses taught in a language other than English indicated in the catalog(s):_____
_____

5g. Courses Offered through Distance Learning (Circle answer)

Yes No

If yes, how are distance-learning courses indicated in the catalog(s)? _____
_____

6. Complete the course catalog checklist. What is the status of the checklist (circle one)?

Complete Incomplete

7. Have there been substantial changes in your course offerings between 2001-2002 and the 2004-2005 school years?

Yes No

8. Who is the best person to contact if HSTS staff have questions about the course catalogs?

School Coordinator Principal Registrar Other

Name: _____ Title: _____ Phone number:_____

B. Other School Information

NOTE to supervisor: fill out as much of this section as possible by reviewing the 2004-2005 catalog and then ask for clarification on the rest.

1. How many class periods does a student typically have per day, not including lunch?

   # of class periods:_____

2. How many credits does a student earn for a year-long course taken for a single period over the school year [2004-2005]?

   2004-2005 # of credits:_____

   2a. Has this changed during the last four school years (circle one)

   Yes No (GO TO #2)

   2b. If yes, how many credits are earned for a year-long course for the following years?

   2003-2004 # of credits:_____

   2002-2003 # of credits:_____

   2001-2002 # of credits:_____

3. What is the maximum number of class periods a student can take per day at this school?

   Maximum # of class periods:_____

4. What is the minimum number of class periods a student can take per day at this school?

   Minimum # of class periods:_____

5. Is the minimum number of courses required different for seniors?

   Yes No

   If yes, what is the minimum number for seniors? _____

6. How long does a typical class period last?

   # of minutes:_____

7. Are credits for honors/AP classes defined the same as in Question #2?

   Yes No

   If no, describe any differences: _____
   _____

8. Are credits for special education classes defined the same as in Question #2?

Yes No

If no, describe any differences : _____
_____

9. Are credits for ESL classes defined the same as in Question #2?

Yes No

If no, describe any differences : _____
_____

10. What type of diplomas does this school offer? (check all that apply)

___ Standard

___ Honors

___ Certificate of Merit

___ Vocational

___ Special Education

___ Certificate of Attendance

___ International Baccalaureate

___ Regents (NY State only)

___ Other (specify): _____

11. Are graduation requirements for all high school diplomas documented in the course catalogs?

Documented Not Documented (GO TO #13)

12. Specify the relevant catalog page number(s) indicating graduation requirement. (Place a paper clip on the corresponding pages)

Graduation requirements recorded on page(s)_____ (GO TO #14)

Comments: _____
_____
_____
_____
_____

13. What are the graduation requirements (diploma type) for the following subject areas? (skip this item if #12 indicates where to locate graduation requirements in the catalog(s)

Diploma type	Standard	Honors	Vocational	Other
*Total credits required for graduation	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)

Write NA on the credit lines if the school does not offer the program.

13a. English/Language Arts	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13b. Mathematics	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13c. Computer Science	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13d. Social Studies/History	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13e. Science	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13f. Foreign Language	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13g. Physical Education/Health	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13h. Other (specify_____)	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)
13i. Other (specify_____)	_____ (credits)	_____ (credits)	_____ (credits)	_____ (credits)

* This number may be larger or smaller than the credits specified for subject areas 13a-13i listed here because of electives and/or overlapping sections.

14. Are there any courses required for graduation that do not receive credits?

Yes No

If yes, please specify: _____
_____

15. Do these graduation requirements cover courses taken in grades 9 through 12?

Yes No

If no, please explain: _____
_____

16. Are there grade point average (GPA) requirements for graduation?

Yes No

If yes, please explain: _____
_____

17. Are there state or district competency tests or performance assessments that are required for graduation?

Yes No

If yes, in what content areas (e.g. Reading, Citizenship, Functional Math): _____
_____

18. Does this school offer any special programs or serve as a Magnet School?

Yes No

18a. What types of special programs are offered? (check all that apply)

___ International Baccalaureate

___ Performing Arts

___ Science/Technology

___ Continuing Education

___ Other (specify) _____

18b. When was this/were these programs established at the school?

Year:_____ Program: _____
Year:_____ Program :_____

19. Does your school use a computerized student information system?

Yes No (GO TO #20)

If Yes:

Products used: _____

Product Name: _____

Publisher or Developer (if developed in your state, district or school, so indicate): _____

Does your system:

19a. Produce electronic transcripts: Yes No

19b. Track attendance: Yes No

19c. Record standardized test scores: Yes No

19d. Record graduation dates: Yes No

19e. Record diploma types: Yes No

20. Who would be the best person to contact if HSTS staff have questions about credits, graduation requirements, special programs, or technology resources?

School Coordinator Principal Registrar Other

Name: _____ Title: _____ Phone number:_____

C. Reviewing the Transcripts – Complete this section after you have received copies of the sample transcripts

1. Sample transcripts obtained include (check all that apply)

___ Regular courses

___ Honors courses

___ Special Education courses

___ Information on other special courses

2. What type of grading system is used (e.g. A, B, C or A+, A, A-, B+, etc.)?

_____A, B, C, etc.

_____A+, A, A-, B+, etc.

_____ Pass/Fail

_____Satisfactory/Unsatisfactory

_____ Other (please specify)_____

_____

3. What do the letter grades or other marks stand for numerically (example, A=90-100, B= 80-90, etc.)?

Letter Grade or Alternate Symbol		Range (or description, if range not possible)
A+		
A		
A-		
B+		
B		
B-		
C+		
C		
C-		
D+		
D		
D-		
F		
Pass		
Fail		
Satisfactory		
Unsatisfactory		

4. Is the grading system the same for all students (including special education, honors, etc.)?

   Yes No

   If no, please explain: _____
   _____

5. Do the course titles or course numbers on the transcript match those in the course catalogs(s)?

   Yes No

   Comments: _____
   _____

6. Are transfer courses identified by the school on the transcripts?

   Yes No

   If yes, please explain: _____
   _____

   If no, is there any way that they can be identified: _____
   _____

7. Are there abbreviations or symbols on the transcripts that are not self-evident?

   Yes No

   7a. If yes, ask the respondent to explain the abbreviations or symbols. Record the explanation on the Transcript Format Checklist

   Explained abbreviations/symbols Explanation not known

8. Are you available at the school in June, July, or August?

   Yes (month available):_____ No

   If No, who can we contact over the summer?

   Name: _____ Title: _____ Phone number:_____

9. When will the final transcripts for the class of 2005 students be available?

   Date:_____

10. When will be a convenient time to return to the school to pick up copies of the transcripts?

    Date:_____ Hours:_____

11. I am leaving a copy of the NAEP administration schedule with you today. Where will this copy be kept so I can retrieve it in the summer?_____

Appendix C

Transcript Request Form (TRF)

APPENDIX C. TRANSCRIPT REQUEST FORM

2005 NAEP High School Transcript Study
School ID:
School Name:
Transcripts Requested:
#Transcripts Received:

National School Lunch Program
1=Student Not Eligible 2=Reduced Price Lunch 3=Free Lunch 4=Information Not Available 5=Refused 6=School Not Participating

Exit Status
A = Standard Diploma B=Honors Diploma C=Diploma with special education adjustments D=Certificate of attendance E=Still enrolled in this school F=Dropped out G=Other (such as transferred, GED, unknown)

2005 High School Transcript Study
Transcript Request Form (TRF)

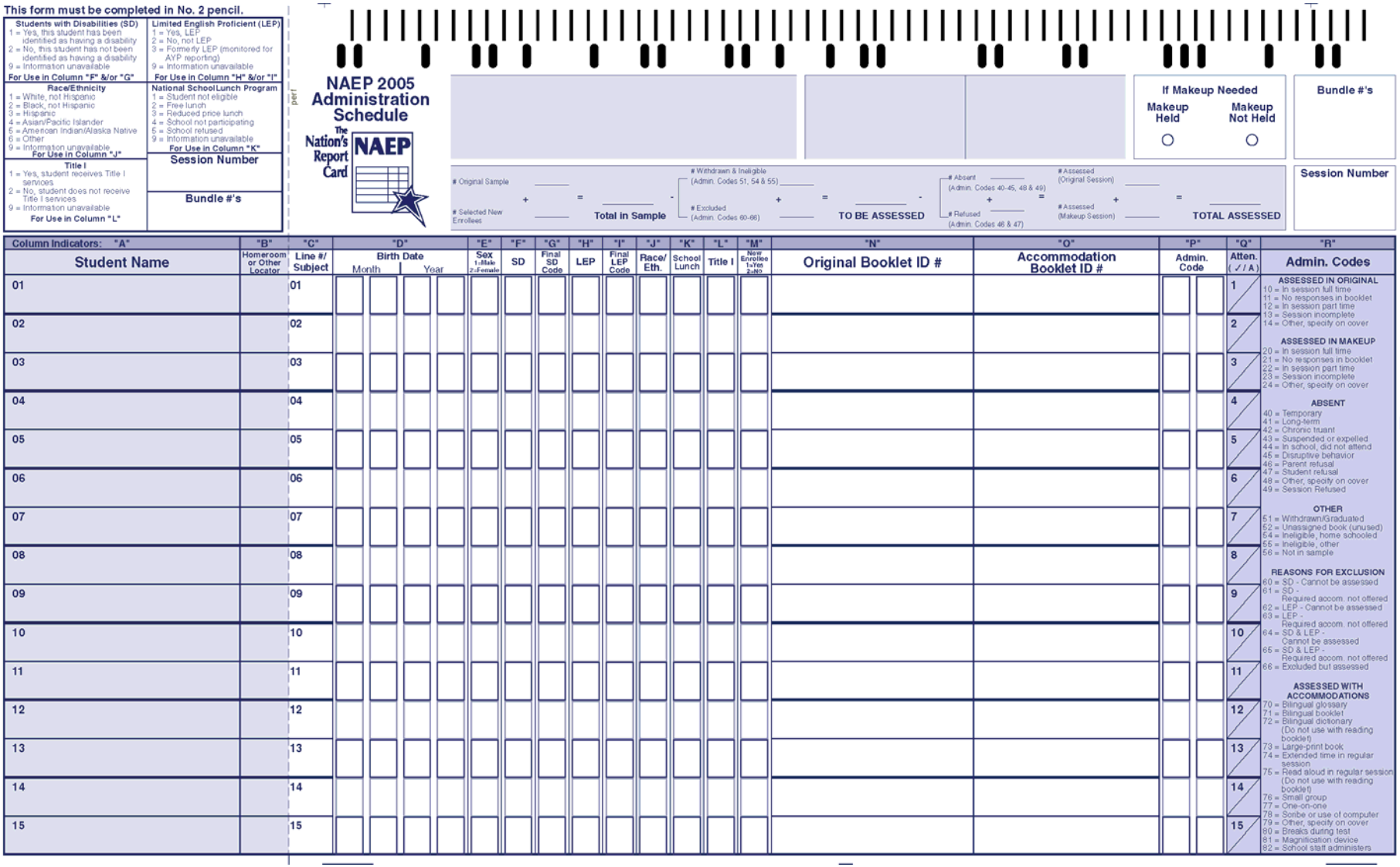
School ID: 123-456-7
School Name: Maple High School

Student Name First, M. Last	NAEP ID	Exit Status	Transcript Received	Complete if Missing						
				Sex	Birth date	Race/ Eth	SD	LEP	Title 1	NSLP
Mary B. Abel	7777777777			F	08/87	White	N	N	N	4
Adam F. Bean	9844555555			M	09/87	Black	N	N	N	4
Susan A Cutter	2222222222			F	05/87	White	N	Y	N	4
Rich S. Danskin	2222222222			M	04/87	White	Y	N	N	4
Stuart L. Fredericks	2222222222			M	12/86	Black	N	N	N	4
Danny M. Guami	2222222222			M	01/87	Hispanic	N	N	N	4
Heather S. Hui	2222222222			F	02/87	Asian	N	Y	N	4

Appendix D

NAEP 2005 Administration Schedule

APPENDIX D. NAEP 2005 ADMINISTRATION SCHEDULE



This form must be completed in No. 2 pencil.

Students with Disabilities (SD)
1 = Yes, this student has been identified as having a disability
2 = No, this student has not been identified as having a disability
9 = Information unavailable
For Use in Column "F" &/or "G"

Limited English Proficient (LEP)
1 = Yes, LEP
2 = No, not LEP
3 = Formerly LEP (monitored for AYP reporting)
9 = Information unavailable
For Use in Column "H" &/or "I"

Race/Ethnicity
1 = White, not Hispanic
2 = Black, not Hispanic
3 = Hispanic
4 = Asian/Pacific Islander
5 = American Indian/Alaska Native
6 = Other
9 = Information unavailable
For Use in Column "J"

National School Lunch Program
1 = Student not eligible
2 = Free lunch
3 = Reduced price lunch
4 = School not participating
5 = School refused
9 = Information unavailable
For Use in Column "K"

Title I
1 = Yes, student receives Title I services
2 = No, student does not receive Title I services
9 = Information unavailable
For Use in Column "L"

Session Number

Bundle #'s

perf

NAEP 2005 Administration Schedule
The Nation's Report Card NAEP

If Makeup Needed
Makeup Held ○ Makeup Not Held ○

Bundle #'s

Original Sample + # Selected New Enrollees = Total in Sample
Total in Sample − # Withdrawn & Ineligible (Admin. Codes 51, 54 & 55) + # Excluded (Admin. Codes 60-66) = TO BE ASSESSED
TO BE ASSESSED − # Absent (Admin. Codes 40-45, 48 & 49) + # Refused (Admin. Codes 46 & 47) = # Assessed (Original Session) + # Assessed (Makeup Session) = TOTAL ASSESSED

Session Number

Column Indicators: "A"	"B"	"C"	"D"				"E"	"F"	"G"	"H"	"I"	"J"	"K"	"L"	"M"	"N"	"O"	"P"		"Q"	"R"
Student Name	Homeroom or Other Locator	Line #/ Subject	Birth Date Month		Year		Sex 1=Male 2=Female	SD	Final SD Code	LEP	Final LEP Code	Race/ Eth.	School Lunch	Title I	New Enrollee 1=Yes 2=NO	Original Booklet ID #	Accommodation Booklet ID #	Admin. Code		Atten. (✓/A)	Admin. Codes
01		01																		1	
02		02																		2	
03		03																		3	
04		04																		4	
05		05																		5	
06		06																		6	
07		07																		7	
08		08																		8	
09		09																		9	
10		10																		10	
11		11																		11	
12		12																		12	
13		13																		13	
14		14																		14	
15		15																		15	

ASSESSED IN ORIGINAL
10 = In session full time
11 = No responses in booklet
12 = In session part time
13 = Session incomplete
14 = Other, specify on cover

ASSESSED IN MAKEUP
20 = In session full time
21 = No responses in booklet
22 = In session part time
23 = Session incomplete
24 = Other, specify on cover

ABSENT
40 = Temporary
41 = Long-term
42 = Chronic truant
43 = Suspended or expelled
44 = In school, did not attend
45 = Disruptive behavior
46 = Parent refusal
47 = Student refusal
48 = Other, specify on cover
49 = Session Refused

OTHER
51 = Withdrawn/Graduated
52 = Unassigned book (unused)
54 = Ineligible, home schooled
55 = Ineligible, other
56 = Not in sample

REASONS FOR EXCLUSION
60 = SD - Cannot be assessed
61 = SD - Required accom. not offered
62 = LEP - Cannot be assessed
63 = LEP - Required accom. not offered
64 = SD & LEP - Cannot be assessed
65 = SD & LEP - Required accom. not offered
66 = Excluded but assessed

ASSESSED WITH ACCOMMODATIONS
70 = Bilingual glossary
71 = Bilingual booklet
72 = Bilingual dictionary (Do not use with reading booklet)
73 = Large-print book
74 = Extended time in regular session
75 = Read aloud in regular session (Do not use with reading booklet)
76 = Small group
77 = One-on-one
78 = Scribe or use of computer
79 = Other, specify on cover
80 = Breaks during test
81 = Magnification device
82 = School staff administers

Column Indicators: "A"	"B"	"C"	"D"				"E"	"F"	"G"	"H"	"I"	"J"	"K"	"L"	"M"	"N"	"O"	"P"		"Q"	"R"
Student Name	Homeroom or Other Locator	Line #/ Subject	Birth Date Month		Year		Sex 1=Male 2=Female	SD	Final SD Code	LEP	Final LEP Code	Race/ Eth.	School Lunch	Title I	New Enrollee 1=Yes 2=No	Original Booklet ID #	Accommodation Booklet ID #	Admin. Code		Atten. (✓/A)	Admin. Codes
16		16																		16	ASSESSED IN ORIGINAL 10 = In session full time 11 = No responses in booklet 12 = In session part time 13 = Session incomplete 14 = Other, specify on cover ASSESSED IN MAKEUP 20 = In session full time 21 = No responses in booklet 22 = In session part time 23 = Session incomplete 24 = Other, specify on cover ABSENT 40 = Temporary 41 = Long-term 42 = Chronic truant 43 = Suspended or expelled 44 = In school, did not attend 45 = Disruptive behavior 46 = Parent refusal 47 = Student refusal 48 = Other, specify on cover 49 = Session Refused OTHER 51 = Withdrawn/Graduated 52 = Unassigned book (unused) 54 = Ineligible, home schooled 55 = Ineligible, other 56 = Not in sample REASONS FOR EXCLUSION 60 = SD - Cannot be assessed 61 = SD - Required accom. not offered 62 = LEP - Cannot be assessed 63 = LEP - Required accom. not offered 64 = SD & LEP - Cannot be assessed 65 = SD & LEP - Required accom. not offered 66 = Excluded but assessed ASSESSED WITH ACCOMMODATIONS 70 = Bilingual glossary 71 = Bilingual booklet 72 = Bilingual dictionary (Do not use with reading booklet) 73 = Large-print book 74 = Extended time in regular session 75 = Read aloud in regular session (Do not use with reading booklet) 76 = Small group 77 = One-on-one 78 = Scribe or use of computer 79 = Other, specify on cover 80 = Breaks during test 81 = Magnification device 82 = School staff administers
17		17																		17	
18		18																		18	
19		19																		19	
20		20																		20	
21		21																		21	
22		22																		22	
23		23																		23	
24		24																		24	
25		25																		25	
26		26																		26	
27		27																		27	
28		28																		28	
29		29																		29	
30		30																		30	
31		31																		31	
32		32																		32	
33		33																		33	
34		34																		34	Pearson NCS® IM-170038-001:654321 Printed in U.S.A. - 2005 A/S

This form must be completed in No. 2 pencil.

Appendix E

Documentation of Missing Transcripts

APPENDIX E. DOCUMENTATION OF MISSING TRANSCRIPTS

NAEP School ID: _____

Supervisor: _____

DOCUMENTATION OF MISSING TRANSCRIPTS

School Name: _____ Date: _____

School ID#: _____

Supervisor: _____

Number of Transcripts Requested: _____

Number of Transcripts Received: _____

Please enter the Student IDs for each missing transcript and the reason given for missing:

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Appendix F

Summary of School Activities

NAEP2005

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

2005 High School Transcript Study

Summary of School Activities

What is the NAEP High School Transcript Study?

The NAEP High School Transcript Study (HSTS), sponsored by the National Center for Education Statistics (NCES), is a periodic survey of transcripts of high school graduates. These studies serve as a barometer for changes in high school student coursetaking patterns, which, in combination with school course offerings, provide valuable information about the rigor of high school curricula across the nation. In addition, this study provides an opportunity to examine the relationship between coursetaking patterns and educational achievement through the link to National Assessment of Educational Progress (NAEP) proficiency data.

The 2005 HSTS is the sixth NAEP transcript study. Additional transcript studies were conducted by NCES with longitudinal studies (i.e., High School and Beyond and the National Education Longitudinal Study of 1988). The HSTS school sample includes public and nonpublic schools in the NAEP 2005 sample. As with the previous studies, no student or teacher time is involved.

What is the Schedule for HSTS 2005 Activities?

Phase 1: September 2004 – March 2005, Preliminary Activities

- A NAEP representative will call each sampled school to discuss the details of the HSTS.
- A NAEP representative will go to sampled high schools on the NAEP assessment day and will place a disclosure notice in the sampled students' files to assist with identifying the student population for whom transcripts will be obtained. This notice will include the provisions from the Family Education Rights and Privacy Act (FERPA), which explain the disclosure safeguards that grant NAEP the authority to obtain transcript information.
- Course catalogs and/or course lists will be requested for the current school year (2004-2005) and for the preceding 3 years (2003-2004, 2002-2003, and 2001-2002).
- Three (3) sample transcripts will be requested. One should include honors courses, one special education courses, and one with regular courses.
- The NAEP representative will review all materials and will ask questions necessary to understand the transcripts and course catalogs.

Phase 2: June – October 2005, Collection of Transcripts

- A NAEP representative will return to each sampled school to collect copies of transcripts for the sampled students.

What are the Key Aspects of the 2005 High School Transcript Study?

- NO STUDENT OR TEACHER TIME IS INVOLVED (NAEP staff will work with school personnel to minimize as much of the burden as possible).
- Confidentiality (students' names and all other identifying information will be removed or masked before removing copies of transcripts from the school).
- No cost to schools (NAEP will pay the school's usual charge for providing transcripts).
- Parental notification (not required by NAEP or No Child Left Behind).

Where Can I Find More Information about the High School Transcript Study?

More information about the NAEP high school transcript studies can be found by visiting the NAEP web site at http://nces.ed.gov/nationsreportcard/hsts/.

NCES National Center for Education Statistics

NAEP is administered by the National Center for Education Statistics
U.S. Department of Education • Institute of Education Sciences