U.S. Department of Education
Institute of Education Sciences
NCES 2004–371

The High School Sophomore Class of 2002: A Demographic Description

First Results From the Base Year of the Education Longitudinal Study of 2002

E.D. Tabs

This page is intentionally blank.

U.S. Department of Education
Institute of Education Sciences
NCES 2004–371

The High School Sophomore Class of 2002: A Demographic Description

First Results From the Base Year of the Education Longitudinal Study of 2002

E.D. Tabs

April 2004

Steven J. Ingels
RTI International

Leslie A. Scott
Education Statistics Services Institute

Jeffrey A. Owings
Project Officer
National Center for Education Statistics

U.S. Department of Education
Rod Paige
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Robert Lerner
Commissioner

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

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

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

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006–5651

April 2004

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

Suggested Citation

Ingels, S.J., and Scott, L.A. (2004). *The High School Sophomore Class of 2002: A Demographic Description First Results From the Base Year of the Education Longitudinal Study of 2002* (NCES 2004-371). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

Or call toll free 1-877-4ED-Pubs

Content Contact:
Jeffrey A. Owings
(202) 502–7423
Jeffrey.Owings@ed.gov

Executive Summary

The data for this report, *The High School Sophomore Class of 2002: A Demographic Description*, describe the demographic characteristics and tested achievement of a cohort based on a nationally representative probability sample of 15,362 10th-graders in 752 public, Catholic, and other private schools who were studied in the spring term of the 2001–02 school year. The base-year data collection for the Education Longitudinal Study of 2002 (ELS:2002) is the first wave of a new longitudinal study of high school students that continues a series of nationally representative longitudinal studies conducted by the U.S. Department of Education's National Center for Education Statistics (NCES) in recent decades. Future survey waves will follow both students and high school dropouts and will monitor the transition of the cohort to postsecondary education, the labor force, and family formation. Although the base-year study comprised surveys of parents, teachers, school administrators, and library media specialists, as well as the cohort of high school sophomores, this report draws primarily on data from students, the primary unit of analysis for the study. (Parent, teacher, librarian, and school surveys provide contextual data for better understanding the student cohort.)

This E.D. Tabs report summarizes the sociodemographic and educational characteristics of the cohort. These characteristics are captured in a series of student- and school-level classification variables. At the student level, these variables are sex, age, race/ethnicity, language minority status, family composition, parental education, students' educational expectations, and tested achievement. Also included are three characteristics of each student's school: sector (public, Catholic, or other private), metropolitan status (urban, suburban, or rural), and region in which it is located (Northeast, Midwest, South, or West).

Comparisons drawn in the text of this report have been tested for statistical significance at the .05 level to ensure that the differences are larger than those that might be expected due to sampling variation. Comparisons are tested with the *t*-statistic. Full details of the statistical tests used can be found in appendix A.

Selected Findings

Various background characteristics and differences may influence the educational experiences, achievement, and expectations of students as they progress through high school. Selected characteristics of the high school sophomore class of 2002 are as follows:

- The majority (t=10.52) of sophomores are White (60 percent). Hispanics and Blacks make up 16 percent and 14 percent of the sophomore cohort, respectively; Asian and multiracial sophomores each constitute 4 percent; and American Indians/Alaska Natives constitute 1 percent of the sophomore cohort (figure 2).
- There are differences by racial/ethnic group in the likelihood that English is a sophomore's native language. English is the native language of 94 percent of Black and 97 percent of White sophomores. It is the native language of 37 percent of Asian/Pacific Islander and 48 percent of Hispanic sophomores (table 1) (min t=22.95).

- About 57 percent of sophomores live in a family with both of their biological or adoptive parents. Others live in a single-parent household (22 percent), or with their mother or father and a guardian (17 percent). Still others (4 percent) live in a variety of other arrangements (table 2).

Reading and mathematics achievement were assessed in terms of various levels of skill and content mastery, or proficiency. Selected findings are:

- Overall, in *mathematics* (figure 5), 92 percent of sophomores are able to perform simple arithmetical operations on whole numbers (proficiency level 1).

- Overall, in *reading* (figure 6), 89 percent of sophomores have mastered the skills of simple reading comprehension (proficiency level 1).

Appended Matter

Appendix A includes technical notes on the report. It also provides an overview of the study design and methodology, a summary of the statistical procedures employed in the report, and a glossary of the ELS:2002 variables and measures used in this analysis. Supplementing the technical notes are tables of standard errors of measurement (appendix B) for the estimates contained in the report.

Foreword

This E.D. Tab profiles American high school sophomores in the 2001–02 school year. It is the first publication based on the Education Longitudinal Study of 2002 (ELS:2002), a new longitudinal study of high school students that continues a series of such studies that NCES has conducted since 1972. (A more extensive statistical analysis report—*A Profile of the American High School Sophomore in 2002*—is in preparation at this time.) In the spring term, students completed assessments in reading and mathematics as well as a questionnaire. Their parents, English and mathematics teachers, school principals, and librarians were surveyed as well.

The data analyzed in this report are now available to researchers for their own use in Electronic Codebook (ECB) format on CD-ROM (NCES 2004–404). The report supplies a demographic profile of 2002 sophomores.

We hope that the information provided in this report will be useful to a wide range of interested readers, including policymakers and educators. We further hope that the results reported here will encourage other researchers to use the ELS:2002 data, both now and in the future, as additional waves build upon this baseline.

Robert Lerner, *Commissioner*

Jeffrey A. Owings, *Associate Commissioner*
Elementary/Secondary & Libraries Studies

This page is intentionally blank.

Acknowledgments

A number of people contributed to the production of this E.D. Tab, and to the development of ELS:2002 more generally. In particular, we wish to thank our NCES reviewers, Gail Mulligan, Elaine Kroe, Jeff Owings, and Marilyn Seastrom, for the quality of their input into this document.

We also would like to acknowledge the role of the ELS:2002 Technical Review Panel, whose members reviewed plans for the study and helped refine them and provided important suggestions to help guide development of the instrumentation. The following individuals serve as members of the Technical Review Panel: Marco Clark, Richard Duran, Jeremy Finn, Thomas B. Hoffer, Thomas Kane, Sally Kilgore, Richard Lawrence, Samuel R. Lucas, Aaron Pallas, and Andy Rogers.

Document preparation support for this E.D. Tab was provided by Mariel Escudero at the Education Statistics Services Institute. Sallie Fiore, Lynne Hawley, and Sonja Douglas of RTI International also provided editorial, document preparation, and graphics support, respectively, for this E.D. Tab.

This page is intentionally blank.

Table of Contents

This page is intentionally blank.

List of Tables

This page is intentionally blank.

List of Figures

This page is intentionally blank.

Introduction

The Education Longitudinal Study of 2002 (ELS:2002) is designed to provide trend data about critical transitions experienced by students as they proceed through high school and into postsecondary education and the labor force. The study is intended to produce a general purpose dataset for the study of numerous educational policy issues. Issues that can be addressed with data collected in the high school years include the following:

- students' academic growth in mathematics;
- the process of dropping out of high school (determinants and consequences);
- the role of family background and the home education support system in fostering students' educational success;
- the impact of coursetaking choices on success in the high school years; and
- steps taken to facilitate the transition from high school to postsecondary education or the world of work.

After ELS:2002 students complete or leave high school, a new set of issues can be examined, starting with the second follow-up in 2006. These include:

- the educational and labor market activities of high school dropouts;
- the transition of non-college bound high school graduates to the workplace;
- access to and choice of undergraduate and graduate educational institutions;
- persistence in attaining postsecondary educational goals;
- rate of progress through the postsecondary curriculum;
- degree attainment;
- barriers to persistence and degree attainment;
- entry of new postsecondary graduates into the work force;
- social and economic rates of return on education to both the individual and society; and
- adult roles, such as family formation and civic participation.

The Current Study

The ELS:2002 base-year study was carried out in a nationally representative probability sample of 752 public, Catholic, and other private schools in the spring term of the 2001–02 school year. Of 17,591 eligible selected sophomores, 15,362 completed a base-year questionnaire, as did 13,488 parents, 7,135 teachers, 743 principals, and 718 librarians. Data used in this report assume the student to be the basic unit of analysis and are taken from the student survey (student questionnaire, assessments in reading and mathematics), the parent survey, and the sampling frame (which contained definitive information about school variables

such as urbanicity, region, and sector). The weighted response rate for student questionnaire completion was 87.3 percent. Of the 15,362 student questionnaire completers, 14,543 (95.1 percent, weighted) also had test data. Parent questionnaire coverage was achieved for 13,488 students (or 87.5 percent of the participating sophomores, weighted) and the school frame variables were available for 100 percent of students and schools. The school participation rate (weighted) was approximately 68 percent. A bias analysis was performed to ensure that biases were small and the data could be used with confidence. Further details of the school-level response rate and bias analysis are provided in appendix A.

Focus of This Report

This report provides descriptive information about the nation's high school sophomores in the spring term of the 2001–02 school year. It reports on the sociodemographic characteristics of the cohort. The report profiles the status of America's 10th-graders both overall, and for various distinct subgroups (such as male and female students, students from different racial/ethnic groups, students from different socioeconomic backgrounds, and students from different types of schools [urban, suburban, and rural; public, Catholic, and other private]).

Comparisons in the report have been tested for statistical significance at the .05 level, as is further explained in appendix A.

The report is organized into four sections covering the following: (1) sex, age, race/Hispanic ethnicity, native language; (2) family composition, mothers' and fathers' highest level of education; (3) school type, urbanicity, geographic region; (4) education expectations and tested achievement.

Appendixes A and B provide technical documentation for the findings presented here, as well as information about how to obtain these data.

Selected Findings

Student Sex, Age, Race/Hispanic Ethnicity Group, Native Language

Sex and Age

- Half of the ELS:2002 sophomore cohort is male (50 percent) and half is female (50 percent) (data not shown).

- Most cohort members were born in 1985 and 1986 and were 15 or 16 years old at the time that they were surveyed (figure 1); sample members were surveyed in the first 5 months of 2002.

Race/Hispanic Ethnicity

- Thirty-six percent of 2002 sophomores are from racial or ethnic minority groups (Black, Asian, American Indian, or Hispanic). Some 60 percent are White, and the remainder (4 percent) are Multiracial[1] (figure 2).

- Among America's high school sophomores, Hispanics and Blacks are the largest minority groups (16 percent and 14 percent, respectively)[2] (figure 2).

Native Language

- Sophomores were asked whether English was their native language (i.e., the first language they learned to speak when they were children). Some 14 percent indicated that English was not their native language, and 86 percent indicated that English was their native language (table 1).

- The incidence of English as a native language varies by racial/ethnic group. English was the native language of nearly all White (97 percent) and Black (94 percent) sophomores. However, English was the native language of 48 percent of Hispanics and 37 percent of Asians. In other words, a non-English language was the native language of one-half or more of the Hispanic and Asian sophomores.

[1] For the purpose of convenience, the following shorthand terms are used in the text of this report to refer to racial/ethnic categories: American Indian (includes Alaska Native); White; Black or African American; Asian (includes Pacific Islander and Native Hawaiian); Hispanic or Latino (includes all races); and Multiracial (includes more than one race). The terms "Black" and "African American" are taken to be generally synonymous, as are "Hispanic" and "Latino." Students of Hispanic origin are not included in other racial/ethnic categories.

[2] U.S. Census Bureau figures released in January 2003 for the U.S. population as of July 2001 show Hispanics as 13 percent of the U.S. population, Blacks as 13 percent, Whites as 70 percent, and Asians as 4 percent. Minority proportions are higher, however, in younger age groups; the modal age of the ELS:2002 cohort at the time of being surveyed was about 15.5 years.

Family Composition and Parental Education

Family Composition

- Some 57 percent of sophomores lived in a mother-father household with their biological or adoptive parents. Some 22 percent lived in a single-parent household with either their mother (19 percent of the total) or father (3 percent of the total). Seventeen percent lived with their mother or father and a guardian (13.4 percent lived in a mother and guardian family, and 3.2 percent lived in a father and guardian family). The remaining 4 percent lived in various other arrangements (table 2).

Parental Education

- Approximately 6 out of 10 sophomores (59 percent) have a mother who continued her education beyond high school (table 3), and 56 percent have a father who continued his education beyond high school (table 4).

Students and Their Schools

- The overwhelming majority of sophomores in 2002 attended public schools (92 percent) (t=148.68). The proportion of sophomores attending Catholic schools was 4 percent and the proportion of sophomores attending other private schools was 3 percent (figure 3).

- Some 30 percent of sophomores attended an urban school, 50 percent attended a suburban school, and 20 percent attended a rural school (figure 4).

- Table 5 shows the proportions of sophomores who lived in each of four regions. The four national Census regions are used for geographical reporting: Northeast, Midwest, South, and West. Among 2002 sophomores, 19 percent lived in the Northeast, 24 percent lived in the Midwest, 34 percent lived in the South, and 23 percent lived in the West.

Educational Expectations and Tested Achievement

Educational Expectations

- Some 90 percent of the 2002 sophomore cohort had a definite expectation of how far they expect to go in the education system. The cohort held generally high educational expectations for the future (table 6).[3]

- Some 72 percent of the cohort expected to complete a 4-year college degree or higher. Indeed, 36 percent expected to go beyond a bachelor's degree and to obtain a graduate or professional degree. Eight percent did not expect to go on to postsecondary education in any form.

[3] Some relevant comparison points are the proportion of persons 25 years and over in the United States in 2000 who had college degrees or higher (26 percent) (U.S. Census Bureau 2003) and the proportion of 1988 eighth-graders who had completed a bachelor's degree or higher 12 years later (29 percent) (Ingels et al. 2002, table 2). Some 66 percent of the 1988 eighth-grade cohort indicated that they expected to complete a bachelor's degree or higher (Hafner et al. 1990).

Tested Achievement

ELS:2002 included assessments in reading and mathematics. The two tests were designed to measure the achievement status of 10th-graders at both the individual- and the group-level.

This E.D. Tab employs a criterion-referenced proficiency score so that achievement can be understood in terms of specific levels of skill mastery. Criterion-referenced proficiency scores are based on clusters of items that mark different levels of achievement on the math and reading scales that illustrate the skills that students have. Clusters of four items each mark five hierarchical levels in math and three in reading.

The math levels are: (1) Simple arithmetical operations with whole numbers; (2) Simple operations with decimals, fractions, powers, and roots; (3) Simple problem solving, requiring the understanding of low-level mathematical concepts; (4) Understanding of intermediate-level mathematical concepts and/or multistep solutions to word problems; and (5) Complex multistep word problems and/or advanced mathematics material.

The reading levels are: (1) Simple reading comprehension, including reproduction of detail and/or the author's main thought; (2) Ability to make relatively simple inferences beyond the author's main thought and/or understand and evaluate abstract concepts; and (3) Ability to make complex inferences or evaluative judgments that require piecing together multiple sources of information from the passage.

- About 92 percent of the cohort are proficient in simple arithmetical operations with whole numbers; and 67 percent are proficient in simple operations with decimals, fractions, roots, and powers (figure 5).

- About one-half are capable of simple problem solving in mathematics, and only about one-fifth show proficiency in understanding of intermediate-level mathematical concepts (figure 5).

- The overwhelming majority (nearly 90 percent) of sophomores are proficient in simple reading comprehension. However, when it comes to the ability to make relatively simple inferences beyond the author's main thought or evaluate abstract concepts, less than half of cohort members (46 percent) demonstrate proficiency (figure 6).

- At the highest level of reading proficiency, ability to make complex inferences or judgments based on combining multiple sources of information, only about 8 percent show mastery (figure 6).

References

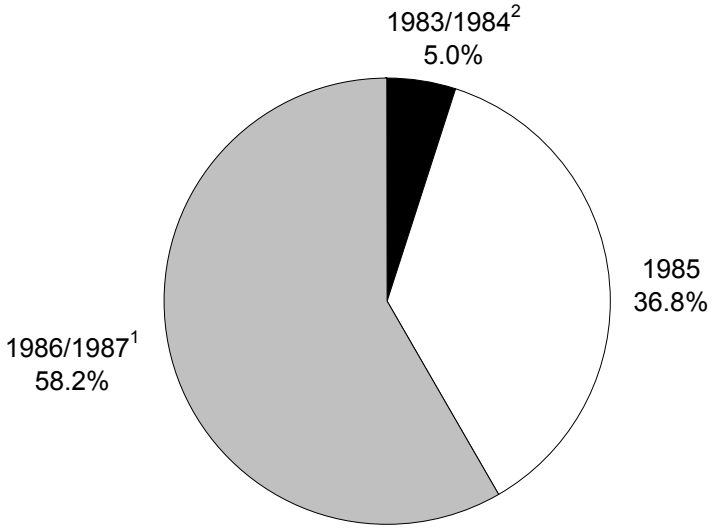
Hafner, A., Ingels, S., Schneider, B., and Stevenson, D. (1990). *A Profile of the American Eighth Grader: NELS:88 Student Descriptive Summary* (NCES 90–458). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

Ingels, S.J., Curtin, T.R., Kaufman, P., Alt, M.N., and Chen, X. (2002). *Coming of Age in the 1990s: The Eighth Grade Class of 1988 12 Years Later* (NCES 2002–321). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office.

U.S. Census Bureau. (2003). *USA Statistics in Brief*. Available: http://www.census.gov/statab/www/part2.html#education

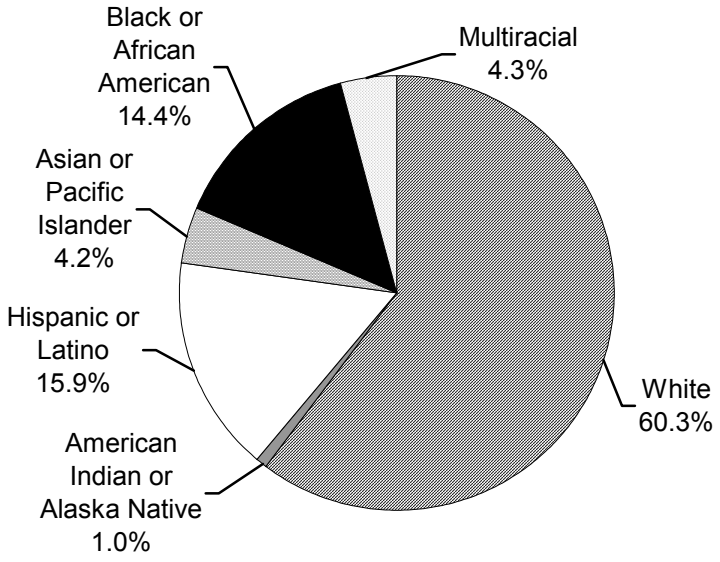
Figures

Figure 1. Percentage of high school sophomores in 2002, by year of birth: 2002



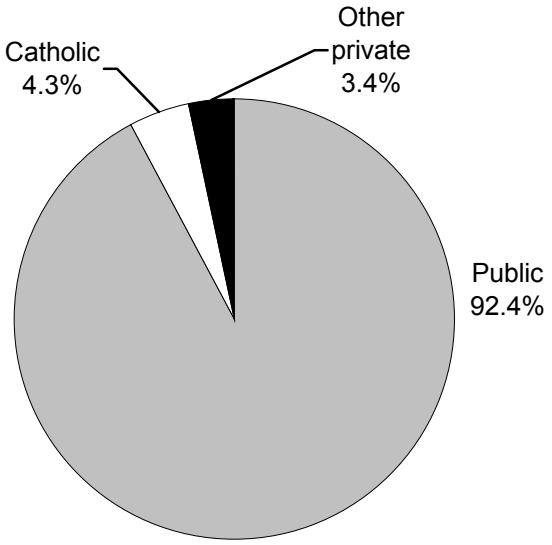
[1] 57.6 percent born in 1986 and an additional 0.5 percent born in 1987 or later.
[2] 4.4 percent born in 1984 and an additional 0.6 percent born in 1983 and earlier.
NOTE: Detail may not sum to totals because of rounding. See appendix A for the weighted response rates of all unimputed variables used in this analysis. Aggregated estimates were derived from unrounded estimates.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Figure 2. Percentage of high school sophomores, by race/ethnicity: 2002



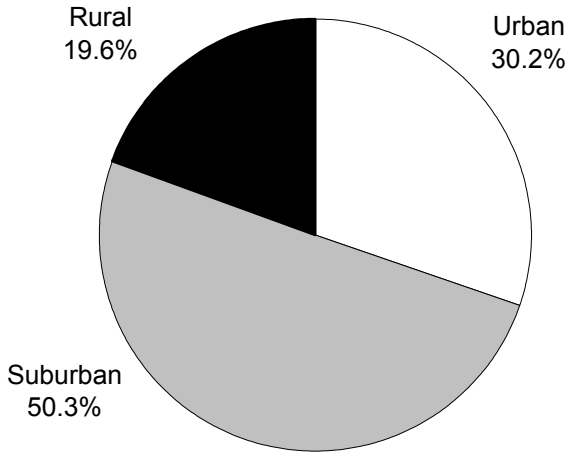
NOTE: Detail may not sum to totals because of rounding. All race categories exclude Hispanic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Figure 3. Percentage of high school sophomores attending various types of schools: 2002



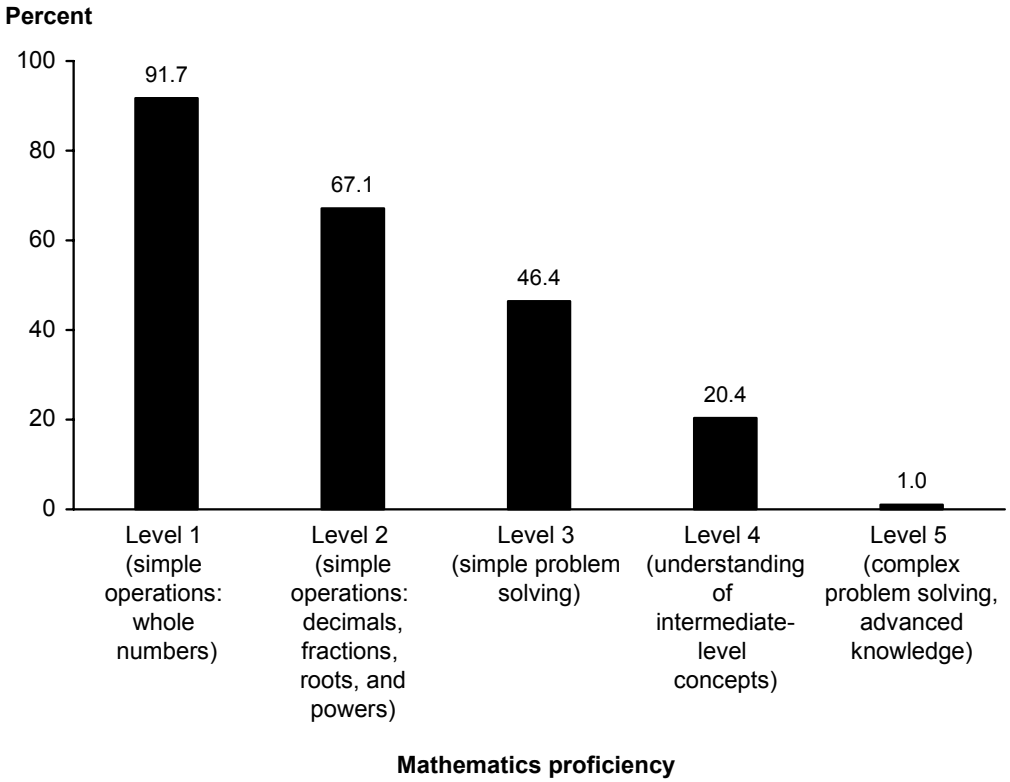
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Figure 4. Percentage of high school sophomores in urban, suburban, and rural schools: 2002



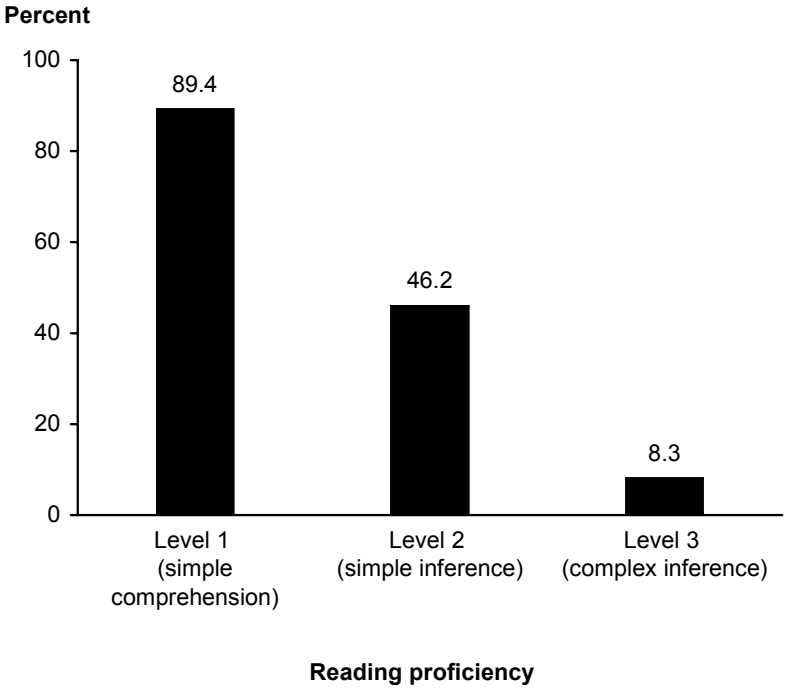
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Figure 5. Percentage of high school sophomores, by demonstrated mathematics proficiency: 2002



SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Figure 6. Percentage of high school sophomores, by demonstrated reading proficiency: 2002



SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Tables

Table 1. Percentage of high school sophomores whose native language is English, by race/ethnicity: 2002

Race/ethnicity	Percent
Total	86.0
American Indian or Alaska Native	83.7
Asian or Pacific Islander	36.9
Black or African American	94.4
Hispanic	47.7
More than one race	92.5
White	97.0

NOTE: All race categories exclude Hispanic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table 2. Percentage of high school sophomores living in various family compositions: 2002

Family composition	Percent
Mother and father	56.8
Mother and guardian	13.4
Father and guardian	3.2
Two guardians	1.9
Mother only	19.0
Father only	3.2
Female guardian only	1.3
Male guardian only	0.2
Parent/guardian lives with student less than half of the time	0.9

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table 3. Percentage of high school sophomores, by mother's highest level of education: 2002

Mother's level of education	Percent
Did not finish high school	13.2
Graduated from high school or received GED	27.9
Attended 2-year school, no degree	13.1
Graduated from 2-year program	11.2
Attended 4-year program, no degree	10.3
Graduated from college	16.6
Completed master's degree or equivalent	6.0
Completed Ph.D., M.D., or other advanced degree	1.7

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table 4. Percentage of high school sophomores, by father's highest level of education: 2002

Father's level of education	Percent
Did not finish high school	13.9
Graduated from high school or received GED	30.1
Attended 2-year school, no degree	10.0
Graduated from 2-year program	8.2
Attended 4-year program, no degree	9.3
Graduated from college	16.7
Completed master's degree or equivalent	7.4
Completed Ph.D., M.D., or other advanced degree	4.4

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table 5. Percentage of high school sophomores in each geographic region: 2002

Region	Percent
Northeast[1]	18.5
Midwest[2]	24.1
South[3]	34.3
West[4]	23.0

[1] Northeast = CT, ME, MA, NH, NJ, NY, PA, RI, VT.
[2] Midwest = IL, IN, IA, KS, MI, MN, MO, ND, NE, OH, SD, WI.
[3] South = AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV.
[4] West = AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY.
NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table 6. Percentage of high school sophomores, by student's highest level of education expected: 2002

Level of education	Percent
Less than high school	0.9
High school completion or GED	7.3
Attend or complete 2-year community college or vocational school	6.4
Attend 4-year program, but not complete degree	3.9
Graduate from college	35.8
Obtain master's degree or equivalent	19.7
Obtain Ph.D., M.D., or other advanced degree	16.1
Don't know	9.8

NOTE: Details may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Appendix A
Technical Notes and Glossary

This page is intentionally blank.

Appendix A
Technical Notes and Glossary

A.1 Overview of the Technical Appendix

The National Center for Education Statistics (NCES) of the U.S. Department of Education has collected longitudinal data for more than 30 years. Starting in 1972 with the National Longitudinal Study of the High School Class of 1972 (NLS-72), and continuing to the most recent study, the Education Longitudinal Study of 2002 (ELS:2002), NCES has provided longitudinal and trend data to education policymakers and researchers that link secondary school educational achievement and experiences with important later outcomes, such as entry into the labor market and postsecondary educational access and attainment.

The base year of ELS:2002 is the first stage of a major effort designed to provide data about critical transitions experienced by students as they proceed through high school and into postsecondary education and the labor force. The 2002 sophomore cohort will be followed, initially at 2-year intervals, to collect policy-relevant data about educational processes and outcomes, especially as such data pertain to student learning, predictors of dropping out, and high school effects on students' access to, and success in, postsecondary education and the work force.

The first section of this appendix details ELS:2002 study objectives; lists some of the major research and policy issues that the study addresses; explains the four kinds of analysis—cross-sectional, longitudinal, cross-cohort, and international comparisons—that can be conducted with ELS:2002 data; and supplies an overview of the base-year study design and methodology.

This section is followed by discussions of base-year sampling, weighting, response rates, quality of estimates, standard errors, and electronic codebooks. Next, an account is provided of the statistical procedures employed. Finally, a glossary is provided that documents the specific variables used in the analyses in this report.

A.2 Overview of ELS:2002

A.2.1 Study Objectives

ELS:2002 is a longitudinal study in which the same units are surveyed repeatedly over time. Individual students will be followed until about age 30; the base-year schools will be surveyed twice (they were surveyed in 2002 and will be surveyed again in 2004). In the high school years, ELS:2002 is an integrated multilevel study, involving multiple respondent populations, including students, their parents, their teachers, and their schools (from which data are collected via questionnaires completed by school administrators and librarians, and by means of an observational facilities checklist completed by survey administrators). This multilevel focus will supply researchers with a comprehensive picture of the home, community, and school

environments and their influences on the student. This multiple-respondent perspective is unified by the fact that, for most purposes, the student is the basic unit of analysis.[1]

Key elements in the ELS:2002 longitudinal design are summarized by wave below.

Base Year (2002)

- Baseline survey of high school sophomores completed in spring term 2002.
- Cognitive test in reading (to serve as a baseline covariate).
- Cognitive test in mathematics (as a baseline for measuring achievement gain over time).
- Surveys of parents, English teachers, and math teachers completed. School administrator questionnaires also collected.
- Additional components for this study included a school facilities checklist and a media center (library) questionnaire.
- Sample sizes of approximately 750 schools and over 15,000 students. Schools were the first-stage unit of selection, with sophomores randomly selected within schools.
- Oversampling of Asians and private schools.
- Test score linkages to the Program for International Student Assessment (PISA), the National Assessment of Educational Progress (NAEP), and prior longitudinal studies (HS&B, NELS:88).

First Follow-up (2004)

- Follow-up in 2004, when most sample members are seniors, but some are dropouts or in other grades.
- Student questionnaire, dropout questionnaire, assessment in mathematics, and school administrator questionnaire to be administered.
- Return to the same schools, but separately follow transfer students.
- Freshening for a 2004 senior cohort[2]

[1] Base-year school administrator, library media center, and facilities data can be used to report on the nation's schools with 10th grades in the 2001–02 school year. However, the primary use of the school-level data (and the purpose of parent and teacher surveys) is to provide further contextual information on the student.

[2] Freshening is a sampling procedure that will bring new students into the study to obtain a nationally representative sample of 12th-graders in 2004. For more information on the procedure of freshening, see the *Education Longitudinal Study of 2002: Base Year Data File User's Manual* (Ingels et al. 2004).

- High school transcript component in 2004 (coursetaking records for grades 9–12 at minimum).

Second Follow-up (2006)

- Post-high-school follow-ups using a single questionnaire with branching of questions to accommodate the diverse pathways followed by the cohort.

- Questionnaire will be available in multiple modalities: web for self-administration, computer-assisted telephone interview, computer-assisted personal interview.

Further Follow-ups

- Number of (and dates for) further follow-ups to be determined.

A.2.2 ELS: 2002 Research and Policy Issues

Apart from helping to describe the characteristics of high school students and their schools, ELS:2002 will provide information to help address a number of key policy and research questions. The study is intended to produce a comprehensive data set for the development and evaluation of educational policy. Part of its aim is to inform decision makers, educational practitioners, and parents about the changes in the operation of the educational system over time and the effects of various elements of the system on the lives of the individuals who pass through it. Issues that can be addressed with data collected in the high school years include the following:

- Students' academic growth in mathematics.
- The process of dropping out of high school—determinants and consequences.
- The role of family background and the home education support system in fostering students' educational success.
- The features of effective schools (e.g., school structural or organizational features or practices associated with higher levels of achievement gain, after controls have been imposed for student background and other factors).
- The impact of coursetaking choices on success in the high school years (and thereafter).
- The distribution of educational opportunities as registered in the distinctive school experiences and performance of students from various policy-relevant subgroups. Such subgroups include:
  - students in public and private high schools;
  - language minority students;
  - students with disabilities;
  - students in urban, suburban, and rural settings;
  - students in different regions of the country;

  - students from upper, middle, and lower socioeconomic status levels;
  - male and female high school students; and
  - students from different racial or ethnic groups.
- Steps taken to facilitate the transition from high school to postsecondary education or the world of work.

After ELS:2002 students have completed high school, a new set of issues can be examined. These issues include

- the later educational and labor market activities of high school dropouts;
- the transition of those who do not go directly on to postsecondary education or to the world of work;
- access to, and choice of, undergraduate and graduate educational institutions;
- persistence in attaining postsecondary educational goals;
- rate of progress through the postsecondary curriculum;
- degree attainment;
- barriers to persistence and attainment;
- entry of new postsecondary graduates into the work force;
- social and economic rate of return on education to both the individual and society; and
- adult roles, such as family formation and civic participation.

A.2.3 Analytic Levels

These research and policy issues can be investigated at several distinct levels of analysis. The overall scope and design of the study provide for the following four analytical levels:

- cross-sectional profiles of the nation's high school sophomores and seniors (as well as dropouts after spring of the sophomore year);
- longitudinal analysis (including examination of life-course changes);
- intercohort comparisons with American high school students of earlier decades; and
- international comparisons: U.S. 15-year-olds compared to 15-year-olds in other nations.

Cross-sectional Profiles. Cross-sectional data will permit characterization of the nation's high school sophomores in the spring of the 2001–02 school year. Initial cross-sectional findings from the base year are presented in this report. Because of sample freshening, the results 2 years later will provide a basis for profiling the nation's high school seniors in the spring term of the 2003–04 school year. Freshening is a sampling procedure that will bring new students into the study to obtain a nationally representative sample of 12th-graders in 2004.

Longitudinal Analysis. Longitudinal analysis will become possible when data are available from the 2004 first follow-up. The primary research objectives of ELS:2002 are longitudinal in nature. The study provides the basis for within-cohort comparison by following the same individuals over time to measure achievement growth in mathematics, monitor enrollment status over the high school years, and record such key outcomes as postsecondary entry and attainment, labor market experiences, and family formation. These outcomes, in turn, can be related to antecedents identified in earlier rounds, including individual, home, school, and community factors.

Intercohort Comparisons. As part of an important historical series of studies that repeats a core of key items each decade, ELS:2002 offers the opportunity for the analysis of trends in areas of fundamental importance, such as patterns of coursetaking, rates of participation in extracurricular activities, academic performance, and changes in goals and aspirations. A 1980–2002 NCES high school sophomore trend report is currently in preparation. With completion of the first follow-up in 2004, researchers will be able to compare ELS:2002 high school seniors' experience, attitudes, and achievement with those of National Education Longitudinal Study of 1988 (NELS:1988) seniors in 1992, High School and Beyond (HS&B) longitudinal study seniors in 1980 and 1982, and NLS-72 seniors in 1972. Such cross-cohort comparisons are of particular importance to measuring the nation's progress in achieving educational opportunities and in measuring the outcomes of school reform and related initiatives.

Starting with the ELS:2002 first follow-up, trend comparisons can also be made with academic transcript data containing students' high school course histories and sequences, since comparable transcript studies have been conducted, starting with HS&B (1982) and including NELS:88 (1992) and NAEP (1987, 1990, 1994, 1998, and 2000).

International Comparisons. A feature of ELS:2002 that expands the study's power beyond that of the predecessor studies is that it will be used to support international comparisons. Items have been included on the ELS:2002 achievement tests from PISA. The Organization for Economic Cooperation and Development's (OECD's) PISA is an internationally standardized assessment, jointly developed by the 32 participating countries (including the United States) and administered to 15-year-olds in groups in their schools (see Lemke et al. 2001). PISA covers three domains: reading literacy, numeracy, and scientific literacy. A subset of the PISA reading literacy and numeracy items have been included in ELS:2002. PISA aims to define each domain not merely in terms of mastery of the school curriculum, but also in terms of important knowledge and skills needed in adult life. Emphasis is placed on the mastery of processes, the understanding of concepts, and the ability to function in various situations within each domain.

A.2.4 Overview of the Base-year Study Design and Content

ELS:2002 was carried out in a national probability sample of 752 participating (of 1,221 eligible contacted) public, Catholic, and other private schools, in the spring term of the 2001–02 school year. Of 17,591 eligible selected sophomores, 15,362 completed a base-year

questionnaire, as did 13,481 of their parents and 7,135 of their teachers.[3] Of the 752 participating schools, 743 principals and 718 librarians completed questionnaires.

Seven study components comprised the base-year design: assessments of students (achievement tests in mathematics and reading); a survey of students; surveys of parents, teachers, school administrators, and librarians; and a facilities checklist (completed by survey administrators, based on their observations at the school). The student assessments measured achievement in mathematics and reading; the baseline scores can serve as a covariate or control variable for later analyses. Mathematics achievement will be reassessed 2 years hence, so that achievement gain over the last 2 years of high school can be measured and related to school processes and mathematics coursetaking. The student questionnaire gathered information about the student's background, school experiences and activities, plans and goals for the future, employment and out-of-school experiences, language background, and psychological orientation toward learning.

One parent of each participating sophomore was asked to respond to a parent survey. The parent questionnaire was designed to gauge parental aspirations for the child, home background and the home education support system, the child's educational history prior to 10th grade, and parental interactions with and opinions about the student's school. For each student enrolled in English or mathematics, a teacher was also selected to participate in a teacher survey. Teachers typically (but not invariably) reported on multiple ELS:2002 sophomores. The teacher questionnaire collected the teacher's evaluation of the students and provided information about the teacher's background and activities. The head librarian or media center director at each school was asked to complete a library media center questionnaire, which inquired into the school's library media center facility, its staffing, its technological resources, collection and expenditures, and scheduling and transactions. Finally, the facilities checklist was a brief observational form completed for each school. The form collected information about the condition of school buildings and facilities. Information about coursetaking (covering all years of high school and including the sequence in which courses were taken and grades earned) will be collected at the end of high school, through the high school transcript component of the ELS:2002 first follow-up study.

For key classification variables, missing data were replaced with imputed values. (See the *Education Longitudinal Study of 2002: Base Year Data File User's Manual* [Ingels et al. 2004] for a discussion of imputation procedures used.) The data set was also subject to disclosure risk analysis and disclosure avoidance editing, including, among other measures, such perturbation techniques as data swapping.

Further details of the instrumentation, sample design, data collection results, data processing, and the data files available for analysis may be found in the *Education Longitudinal Study of 2002: Base-year Data File User's Manual* (Ingels, et al. 2004).[4]

[3] Note that the participating student sample defines the eligible parent and teacher samples. The 7,135 teacher completions are those linked to student participants. Of the 15,362 student participants, 14,081 had at least one associated teacher-provided student report.

[4] See appendix reference list (section A.6) for full citation. The manual can be downloaded from the NCES web site at http://nces.ed.gov/pubsearch

A.3 Sample Design, Weighting, Response Rates, Quality of Estimates, Standard Errors, and the Electronic Codebook

A.3.1 Sampling

The ELS:2002 base-year sample design began with a nationally representative, two-stage stratified probability sample. The first stage of selection was schools; schools were selected with probability proportional to size (PPS). The public school sample was stratified by the nine U.S. Census divisions, and by urbanicity (metropolitan status of urban, suburban, or rural). Private schools (Catholic and other private) were stratified by four levels of geography (Census region) and urbanicity. Private schools were oversampled. The target sample size was 800 schools. Cooperation was sought from 1,221 eligible schools. The realized sample comprised 752 participating 10th-grade schools. The second stage of selection was students. Of 17,591 sampled students in the schools, 15,362 students participated. Some groups (e.g., Asians) were oversampled. NCES' Common Core of Data (CCD) 1999-2000 and the Private School Survey (PSS) 1999-2000 were used as the sampling frame from which schools were selected for ELS:2002.

A.3.2 School and Student Eligibility

Schools were deemed eligible for ELS:2002, if, at the time of school contacting (July 1999 through March 2000), they were in operation and included a 10^{th} grade with 10^{th} grade student enrollment. Students for whom ELS:2002 surveys would be unsuitable (i.e., students with mental disabilities and students who are not proficient in English) and students whose physical or emotional problems would have made participation unduly difficult were determined to be ineligible to complete the student questionnaire and cognitive tests in the base year. Less than one percent of the sample was deemed unable to complete the ELS:2002 student questionnaire. However, contextual data from their parents, teachers, and school principals were collected, and they will be re-evaluated in the first follow-up to determine whether their previous disability or language barrier changed to the extent to allow participation in the first follow-up. For more information on student eligibility, see the *Education Longitudinal Study of 2002: Base Year Data File User's Manual* (Ingels et al. 2004).

A.3.3 Weighting

The general purpose of the weighting scheme was to compensate for unequal probabilities of selection of schools and students into the base-year sample and to adjust for the fact that not all schools and students selected into the sample actually participated. Three sets of weights were computed: a school weight, a weight for student questionnaire completion, and a contextual data weight for the "expanded" sample of questionnaire-eligible and questionnaire-ineligible students.[5] School and student weights were adjusted for nonresponse, and these adjustments were designed to significantly reduce or eliminate nonresponse bias for data elements known for most respondents and nonrespondents. In addition, school weights were

[5] The regular student questionnaire weight (BYSTUWT) generalizes only to the population of students who were eligible to complete the student questionnaire (i.e., those who were not judged incapable of participation by virtue of a severe disability or lack of proficiency in the English language). The expanded sample weight (BYEXPWT) generalizes to the population of all sophomores, regardless of whether they were capable of completing the questionnaire.

poststratified to known population totals. The estimates in this report were produced using BYSTUWT, a cross-sectional weight that generalizes to the population of questionnaire-eligible 10th-graders in regular U.S. high schools in the spring term of the 2001–02 school year.

A.3.4 Response Rates

Of 1,221 eligible contacted schools, 752 participated in the study, for an overall weighted school participation rate of approximately 68 percent (62 percent unweighted). Of 17,591 selected eligible students, 15,362 participated, for an overall weighted student response rate of approximately 87 percent. (School and student weighted response rates reflect use of the base weight [design weight] and do not include nonresponse adjustments.) School and student unit nonresponse bias analyses were performed, as well as an item nonresponse bias analysis for the questionnaires. The school-level bias due to nonresponse prior to computing weights and after computing weights was estimated based on the data collected from both respondents and nonrespondents, as well as sampling frame data. At the unit level (but not the item level), weighting techniques were employed to reduce detected bias, and after final nonresponse adjustments, the remaining relative bias ranged from 0 to 0.2 percent for schools and from 0 to 0.1 for students. For details of the bias analyses, see the *Education Longitudinal Study of 2002: Base-year Data File User's Manual*, NCES 2004-405. Unweighted and weighted school-level response by stratum is summarized in table A-1 (for more details, see table 45 in NCES 2004-405). Second-stage unit response rates by component are summarized in table A-2.[6] Weighted item response rates for all unimputed analysis variables are shown in table A-3. Weighted proportions for missing data that were imputed are shown in table A-4.

Table A-1. Unweighted school sampling and eligibility, and unweighted and weighted participation, by sampling stratum: 2002

School sampling stratum	Sampled schools		Eligible schools		Participating schools		
	Number	Unweighted percent[1]	Number	Unweighted percent[2]	Number	Unweighted percent[3]	Weighted percent
Total	1,268	100.00	1,221	96.29	752	61.59	67.80
Public	953	75.16	926	97.17	580	62.63	69.09
Catholic	140	11.04	140	100.00	95	67.86	74.04
Other private	175	13.80	155	88.57	77	49.68	62.94
Urban	434	34.23	414	95.39	250	60.39	67.27
Suburban	630	49.68	609	96.67	361	59.28	59.81
Rural	204	16.09	198	97.06	141	71.21	79.32

[1] Percent is based on overall total within column. Details may not sum to 100 percent due to rounding.
[2] Percent is based on number sampled within row.
[3] Percent is based on number eligible within row.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

[6] Second stage unit response rate refers to the unit response rate for the second stage of sampling only (e.g., students, teachers, and parents). It does not take into account the cooperation rate for the first stage of sampling, which was schools and which all stage two sampling results are conditional upon. Thus, another measure of unit response is overall unit response, which is the product of the cooperation rate for the first stage of sampling (i.e., schools) and the unit response rate for the second stage of sampling (e.g., students). For example, the overall unit response rate for students is 59 percent—68 percent (the school cooperation rate) x 87 percent (the student unit response rate) (Seastrom 2003).

Table A–2. Summary of ELS:2002 base-year completion and coverage rates: 2002

Instrument	Selected	Participated	Weighted percent	Unweighted percent
Student questionnaire	17,591	15,362	87.28	87.33
Student assessment[1]	15,362	14,543	95.08	94.67
Parent questionnaire[2]	15,362	13,488	87.45	87.80
Teacher ratings of students[3]	15,362	14,081	91.64	91.66
School administrator questionnaire	752	743	98.53	98.80
Library media center questionnaire	752	718	95.93	95.48
Facilities checklist	752	752	100.00	100.00

[1]Percentage of cases for which a student questionnaire and a cognitive test were obtained. Note that test scores have been imputed where missing so that test scores are available for all 15,362 questionnaire completers.
[2]Indicates a coverage rate, the proportion of participating students with a parent report. More parents participated; these completion rates reflect the number of records in the public-use data file, where parent (and teacher) data were excluded for students who did not complete a base-year student questionnaire.
[3]Indicates a coverage rate: ratings obtained from at least one teacher. The number of teachers that completed a teacher questionnaire was 7,135.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table A–3. Weighted response rates for unimputed variables: 2002

Source	Variable label	Variable	Response rate, percent[1]
Student composites	Student's year and month of birth	DOBIRTHP	99.6
Student	How far in school student thinks will get	BYS56	97.5

[1] Weighted item response rates, using the base-year student final weight (BYSTUWT).
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table A–4. ELS:2000 imputation variables: Weighted proportion missing and imputed

Variable	Percent imputed
Student sex	0.06
Student race/ethnicity	0.05
Student language minority status	2.18
Student Hispanic subgroup	3.04
Student Asian subgroup	6.91
School program type	6.47
Student postsecondary educational expectations	2.60
Parental aspirations for student postsecondary achievement	14.25
Family composition	12.80
Mother's educational attainment[1]	4.06
Mother's occupation[1]	5.70
Father's educational attainment[1]	10.00
Father's occupation[1]	14.57
Family income[1]	22.51
Student ability estimates (theta) for reading[2]	6.26
Student ability estimates (theta) for mathematics[2]	5.33

[1]Used to construct socioeconomic status (SES).
[2]Used to construct normative (quartile) and proficiency scores.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

A.3.5 Survey Standard Errors

Because the ELS:2002 sample design involved stratification, the disproportionate sampling of certain strata, and clustered (i.e., multistage) probability sampling, the resulting statistics are more variable than they would have been if they had been based on data from a simple random sample of the same size.

The calculation of exact standard errors for survey estimates can be difficult. Several procedures are available for calculating precise estimates of sampling errors for complex samples. Procedures such as Taylor Series approximations, Balanced Repeated Replication (BRR), and Jackknife Repeated Replication (JRR), which can be found in advanced statistical programs such as SUDAAN, AM, or WESVAR, produce similar results. The ELS:2002 analyses included in this report used SUDAAN and the Taylor Series procedure to calculate standard errors.

A.3.6 Electronic Codebooks

An electronic codebook (ECB)[7] for the ELS:2002 base-year data (NCES 2004–404) is available from NCES. The ECB system is primarily an electronic version of a fully documented survey codebook. It allows the data user to browse through all interview or instrument items (variables) contained in the ELS:2002 data files, to search variable and value labels for key words related to particular research questions, to review the actual wording of these items along with notes and other pertinent information related to them, to examine the definitions and programs used to develop derived variables, and importantly, to output the data for statistical

[7] Information on obtaining electronic codebooks for ELS:2002 and other NCES data collection efforts can be found by reviewing the data products for the study at http://nces.gov/pubsearch

analysis. The ECB also provides an electronic display of the distribution of counts and percentages for each variable in the data set.

Analysts can use the ECB to select or "tag" variables of interest, print hard-copy codebooks that display the distributions of the tagged variables, and generate SAS and SPSS program syntax (including variable and value labels) that can be utilized with the analyst's own statistical software.

Further details of the instrumentation, sample design, data collection results, data processing, and the data files available for analysis may be found in the *Education Longitudinal Study of 2002: Base-year Data File User's Manual* (Ingels, Pratt, Rogers, Siegel, and Stutts 2004).[8]

A.4 Statistical Procedures

Comparisons that have been drawn in this report have been tested for statistical significance to ensure that the differences are larger than those that might be expected due to sampling variation. The statistical comparisons in this report were based on the *t* statistic. Whether the statistical test is considered significant or not is determined by calculating a *t* value for the difference between a pair of means or proportions and comparing this value to published tables of values, called critical values (cv). The alpha level is an *a priori* statement of the probability that a difference exists in fact rather than by chance. Comparisons drawn in the text of this report have been tested for statistical significance at the alpha level of .05.

The *t* statistic between estimates from various subgroups presented in the tables can be computed by using the following formula:

$$t = \frac{x_1 - x_2}{\sqrt{\left(SE_1^2 + SE_2^2\right)}},$$

where x_1 and x_2 are the estimates to be compared (e.g., the means of sample members in two groups), and SE_1 and SE_2 are their corresponding standard errors.

A.5 Glossary—Description of Variables Used

Each variable used in analyses for this report is described below. Variables are alphabetized within topic. The topics are student demographic characteristics; family characteristics; school characteristics; and students' expectations for the future.

When the variable is available in the ELS:2002 base-year data file, the variable name appears in parentheses after the bold entry name. ELS:2002 variables used to construct a variable that is not provided in the ELS:2002 base-year data file are named in all capital letters

[8] See appendix reference list (section A.6) for full citation. The manual can be downloaded from the NCES web site at http://nces.ed.gov/pubsearch

within the descriptive text. For more information on variable construction, please see *The Education Longitudinal Study of 2002: Base-year Data File User's Manual* (Ingels et al. 2004).

STUDENT DEMOGRAPHIC CHARACTERISTICS

NATIVE LANGUAGE/LANGUAGE MINORITY STATUS (STLANG): The data for STLANG are taken directly from the student questionnaire (BYS67) when available. Otherwise, they are imputed.

RACE/ETHNICITY (RACE): The race/ethnicity variable for this report is based on RACE with one simplification: "Hispanic/Latino, race specified" and "Hispanic/Latino, no race specified" are combined into one category, "Hispanic or Latino." The resulting categories are: 1) American Indian or Alaska Native; 2) Asian or Pacific Islander, including Native Hawaiian; 3) Black, including African American; 4) Hispanic or Latino; 5) More than one race or Multiracial; and 6) White. All race categories exclude individuals of Hispanic ethnicity.

RACE reflects new federal standards for collecting race and ethnicity data that allow respondents to mark more than one choice for race. RACE was obtained from the student questionnaire (BYS15, BYS17A, BYS17B, BYS17C, BYS17D, and BYS17E) when available or from (in order of preference) the sampling roster, the parent questionnaire if the parent respondent was a biological parent, or logical imputation based on other questionnaire items (e.g., surname, native language).

SEX (SEX): This variable was constructed from BYS14 on the base-year student questionnaire or, where missing, from (in order of preference) the school roster, logical imputation based on first name, or statistical imputation.

YEAR OF BIRTH (DOBIRTHP): Year of birth was "stripped" from DOBIRTHP, month and year of birth. In the construction of DOBIRTHP, the years 1980, 1981, and 1982 were set to 1983. The years 1988 and 1989 were set to 1987. Dates before 1980 or after 1989 were set to missing. See table A-2 for weighted response rates.

FAMILY CHARACTERISTICS

FAMILY COMPOSITION (BYFCOMP): BYFCOMP is based on parent questionnaire data, or where data were missing, were imputed. BYFCOMP reflects the relationship of the parent questionnaire respondent and his/her spouse/partner to the 10th-grader (BYP01 and BYP04) with one exception; if the parent questionnaire respondent indicated that the 10th-grader lived with him/her less than half time (BYP05) and the 10th-grader did not attend a boarding school (BYA03O), the family was classified as "Lives with student less than half time." Apart from these cases, families were classified into one of eight family types (with a "ninth category "lives with student less than half time"): 1) Mother and father; 2) Mother and guardian; 3) Father and guardian; 4) Two guardians; 5) Mother only; 6) Father only; 7) Female guardian only; and 8) Male guardian only.[9]

[9] "Mother" or "Father" could be either the biological or adoptive mother or father of the ELS:2002 10th-grader. "Guardian" unspecified, as with "Mother and guardian," "Father and guardian," or "Two guardians," could be either a

FATHER'S EDUCATION (FATHED): Father's highest level of education completed is taken from the parent questionnaire (BYP34A or BYP34B, depending on the sex of the respondent) or, where missing, from (in order of preference) the student questionnaire (BYS83B) or imputation. Eight distinct levels of education are identified: 1) Did not finish high school; 2) Graduated from high school or GED; 3) Attended 2-year school, no degree; 4) Graduated from 2-year school; 5) Attended college, no 4-year degree; 6) Graduated from college; 7) Completed master's degree or equivalent; 8) Completed Ph.D., M.D., or other advanced degree.[10]

MOTHER'S EDUCATION (MOTHED): Mother's highest level of education completed is taken from the parent questionnaire (BYP34A or BYP34B, depending on the sex of the respondent) or, where missing, from (in order of preference) the student questionnaire (BYS83A) or imputation. Eight distinct levels of education are identified: 1) Did not finish high school; 2) Graduated from high school or GED; 3) Attended 2-year school, no degree; 4) Graduated from 2-year school; 5) Attended college, no 4-year degree; 6) Graduated from college; 7) Completed master's degree or equivalent; 8) Completed Ph.D., M.D., or other advanced degree.

SCHOOL CHARACTERISTICS

REGION (BYREGION): Geographic region in which the school is located: Northeast (CT, ME, MA, NH, NJ, NY, PA, RI, and VT); Midwest (IL, IN, IA, KS, MI, MN, MO, ND, NE, OH, SD, and WI); South (AL, AR, DC, DE, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, and WV); and West (AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, and WY). This is taken directly from ELS:2002 sampling data.

SECTOR/TYPE (BYSCTRL): Type of school: Public, Catholic, or Other Private. This is taken directly from ELS:2002 sampling data.

URBANICITY/LOCATION (BYURBAN): Metropolitan status of the school: Urban, Suburban or Rural. Urbanicity of school locale was taken from the source data for sampling, which was the Common Core of Data (CCD) 1999-2000 and the Private School Survey (PSS) 1999-2000. CCD contains an 8-level locale variable. For this report, the 8-level CCD variable was collapsed into 3 levels as follows: Urban: large or mid-size central city (i.e., CCD 1 and 2); Suburban: large or small town or urban fringe of a large of mid-size city (i.e., CCD 3, 4, 5, 6); and Rural: school is in a rural area (i.e., CCD 7, 8).

male or female. Approximately 1 percent of the students are in families with a parent and a guardian or two guardians of the same sex.

[10] For about 1 percent of cases, a respondent classified under mother's education could be a male spouse/partner of a 10th-grader's biological or adoptive father, and vice versa, a respondent classified under father's education could be a female spouse/partner of a 10th-grader's biological or adoptive mother.

TEST SCORES

PROBABILITY OF PROFICIENCY SCORES IN READING AND MATHEMATICS (BYTX1RPP, BYTX2RPP, BYTX3RPP, BYTX1MPP, BYTX2MPP, BYTX3MPP, BYTX4MPP, BYTX5MPP)

Criterion-referenced proficiency probability scores are based on clusters of items that mark different levels on the reading and mathematics scales developed in NELS:88. Clusters of four items each were identified in the NELS:88 tests that marked three hierarchical levels in reading and five in mathematics.

Reading Levels:

1. Simple reading comprehension, including reproduction of detail, and/or the author's main thought.
2. Simple inferences beyond the author's main thought and/or understanding and evaluating abstract concepts.
3. Complex inferences or evaluative judgments requiring multiple sources of information.

Mathematics Levels:

1. Simple arithmetical operations with whole numbers.
2. Simple operations with decimals, fractions, powers, and roots.
3. Simple problem solving, requiring the understanding of low-level mathematical concepts.
4. Understanding of intermediate-level mathematical concepts and/or multistep solutions to word problems.
5. Complex multistep word problems and/or advanced mathematics material.

The proficiency levels are hierarchical in the sense that mastery of a higher level typically implies proficiency at lower levels. The proficiency probabilities were computed using IRT-estimated item parameters calibrated in NELS:88. Each proficiency probability represents the likelihood that a student would pass a given proficiency level defined as above in the NELS:88 sample. It should be remembered that probability of proficiency scores are IRT-derived estimates based on overall performance rather than counts of actual item responses. Owing to the two-stage adaptive format of the ELS:2002 assessments, not all sophomores received all items. Nevertheless, the IRT model permits proficiency probabilities to be estimated, even for those sophomores who were not administered a particular proficiency cluster. Table A-5 shows variable names, descriptions, and summary statistics for the ELS:2002 proficiency probability scores.

Table A–5. Reading and mathematics probability of proficiency scores

Variable name	Description	Range	Weighted mean	Weighted standard deviation
BYTX1RPP	Reading – Level 1	0–1	0.89	0.26
BYTX2RPP	Reading – Level 2	0–1	0.46	0.40
BYTX3RPP	Reading – Level 3	0–1	0.08	0.21
BYTX1MPP	Math – Level 1	0–1	0.92	0.20
BYTX2MPP	Math – Level 2	0–1	0.67	0.42
BYTX3MPP	Math – Level 3	0–1	0.46	0.46
BYTX4MPP	Math – Level 4	0–1	0.21	0.33
BYTX5MPP	Math – Level 5	0–1	0.01	0.07

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

This report illustrates a cross-sectional use of the probability of proficiency scores: proficiency probabilities are averaged to produce estimates of mastery rates both overall and within population subgroups. (Note that dichotomous proficiency scores [as appeared on the NELS:88 data set], indicating in yes/no fashion whether a given student is proficient at a particular level, have not been produced for the ELS:2002 data.) Since the range of the scores is zero to one, means can be expressed in percentage form. For example, the weighted mean for mastery of math level 1 is 0.92, which is equivalent to saying that 92 percent of the sophomore cohort had achieved mastery at this level (simple arithmetical operations on whole numbers). While the continuous probability of proficiency scores can be used to measure status, they are perhaps most useful for measuring change. A sophomore trend report (currently in preparation) will illustrate the use of the proficiency probabilities in measuring intercohort change (essentially, since NELS:88 and ELS:2002 have been equated and are on the same scale, mean gain or loss across cohorts at any proficiency level can be measured by subtracting the NELS:88 score from the ELS:2002 score). With the addition of the ELS:2002 first follow-up data, the probability of proficiency scores can also be used longitudinally to measure achievement gain. Since base year and first follow-up will be on the same vertical scale, mean gain (or loss) can be determined by subtracting the base-year probability score from the first follow-up probability score. Measuring gains in probability of proficiency at each mastery level permits researchers to investigate not only the amount of gain in total scale score points but also where (that is, what proficiency level) along the score scale different students are making their largest gains in achievement between sophomore and senior year. In turn, it is possible to relate gains in specific skills to specific school processes or curricular experiences.

EXPECTATIONS FOR THE FUTURE

EDUCATIONAL EXPECTATIONS (STEXPECT): This variable is taken directly from the student questionnaire (BYS56) when available and imputed otherwise. Students were asked, "As things stand now, how far in school do you think you will get?" The eight response options were: 1) Less than high school graduation; 2) High school graduation or GED only; 3) Attend or complete a 2-year program in a community college or vocational school; 4) Attend college, but not complete a 4-year degree; 5) Graduate from college; 6) Obtain a master's degree or equivalent; 7) Obtain a Ph.D., M.D., or other advanced degree; 8) Don't know.

A.6 Appendix A References

Ingels, S.J., Pratt, D.J., Rogers, J., Siegel, P.H., and Stutts, E.S. (2004). *Education Longitudinal Study of 2002: Base-year Data File User's Manual* (NCES 2004–405). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Available: http://nces.ed.gov/pubsearch

Lemke, M., Lippman, L., Bairu, G., Calsyn, C., Kruger, T., Jocelyn, L., Kastberg, D., Liu, Y., Roey, S., and Williams, T. (2001). *Outcomes of Learning: Results from the 2000 Program for International Student Assessment of 15-Year-Olds in Reading, Mathematics, and Science Literacy* (NCES 2002–115). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Available: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2002115

Seastrom, M. (2003). *NCES Statistical Standards* (NCES 2003-601). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Available: http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003601

Appendix B
Standard Error Tables

This page is intentionally blank.

Appendix B
Standard Error Tables

Table B–1. Standard errors for table 1 estimates (percentage of high school sophomores whose native language is English, by race/ethnicity: 2002)

Race/ethnicity	Standard error
Total	0.60
American Indian or Alaska Native	4.46
Asian or Pacific Islander	2.01
Black or African American	0.64
Hispanic	1.93
More than one race	1.04
White	0.28

NOTE: All race categories exclude Hispanic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–2. Standard errors for table 2 estimates (percentage of high school sophomores living in various family compositions: 2002)

Family composition	Standard error
Mother and father	0.57
Mother and guardian	0.36
Father and guardian	0.16
Two guardians	0.13
Mother only	0.44
Father only	0.20
Female guardian only	0.11
Male guardian only	0.04
Parent/guardian lives with student less than half of the time	0.09

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–3. Standard errors for table 3 estimates (percentage of high school sophomores, by mother's highest level of education: 2002)

Mother's level of education	Standard error
Did not finish high school	0.54
Graduated from high school or received GED	0.49
Attended 2-year school, no degree	0.35
Graduated from 2-year program	0.33
Attended 4-year program, no degree	0.29
Graduated from college	0.46
Completed master's degree or equivalent	0.27
Completed Ph.D., M.D., or other advanced degree	0.15

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–4. Standard errors for table 4 estimates (percentage of high school sophomores, by father's highest level of education: 2002)

Father's level of education	Standard error
Did not finish high school	0.54
Graduated from high school or received GED	0.53
Attended 2-year school, no degree	0.31
Graduated from 2-year program	0.31
Attended 4-year program, no degree	0.30
Graduated from college	0.43
Completed master's degree or equivalent	0.30
Completed Ph.D., M.D., or other advanced degree	0.26

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–5. Standard errors for table 5 estimates (percentage of high school sophomores in each geographic region: 2002)

Region	Standard error
Northeast[1]	0.65
Midwest[2]	0.65
South[3]	0.66
West[4]	0.81

[1] Northeast = CT, ME, MA, NH, NJ, NY, PA, RI, VT.
[2] Midwest = IL, IN, IA, KS, MI, MN, MO, NE, ND, OH, SD, WI.
[3] South = AL, AR, DE, DC, FL, GA, KY, LA, MD, MS, NC, OK, SC, TN, TX, VA, WV.
[4] West = AK, AZ, CA, CO, HI, ID, MT, NV, NM, OR, UT, WA, WY.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–6. Standard errors for table 6 estimates (percentage of high school sophomores, by student's highest education expected: 2002)

Level of education	Standard error
Less than high school	0.10
High school completion or GED	0.30
Attend or complete 2-year community college or vocational school	0.29
Attend 4-year program, but not complete degree	0.18
Graduate from college	0.46
Obtain master's degree or equivalent	0.44
Obtain Ph.D., M.D., or other advanced degree	0.40
Don't know	0.30

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–7. Standard errors for figure 1 estimates (percentage of high school sophomores in 2002, by year of birth: 2002)

Year	Standard error
1983/1984[1]	0.27
1985	0.48
1986/1987[2]	0.54

[1] 4.4 percent born in 1984 and an additional 0.6 percent born in 1983 and earlier.
[2] Includes 0.5 percent students who were born in 1987 or later.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–8. Standard errors for figure 2 estimates (percentage of high school sophomores, by race/ethnicity: 2002)

Race/ethnicity	Standard error
American Indian or Alaska Native	0.20
Asian or Pacific Islander	0.26
Black or African American	0.66
Hispanic or Latino	0.87
More than one race	0.23
White	0.98

NOTE: All race categories exclude Hispanic.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–9. Standard errors for figure 3 estimates (percentage of high school sophomores attending various types of schools: 2002)

School type	Standard error
Catholic	0.16
Other private	0.23
Public	0.29

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–10. Standard errors for figure 4 estimates (percentage of high school sophomores in urban, suburban, and rural schools: 2002)

School location	Standard error
Urban	0.75
Suburban	0.80
Rural	0.63

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–11. Standard errors for figure 5 estimates (percentage of high school sophomores, by demonstrated mathematics proficiency: 2002)

Mathematics proficiency	Standard error
Level 1 (simple operations: whole numbers)	0.30
Level 2 (simple operations: decimals, fractions, roots, and powers	0.77
Level 3 (simple problem solving)	0.81
Level 4 (understanding of intermediate concepts)	0.54
Level 5 (complex problem solving, advanced knowledge)	0.08

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).

Table B–12. Standard errors for figure 6 estimates (percentage of high school sophomores, by demonstrated reading proficiency: 2002)

Reading proficiency	Standard error
Level 1 (simple comprehension)	0.39
Level 2 (simple inference)	0.70
Level 3 (complex inference)	0.28

SOURCE: U.S. Department of Education, National Center for Education Statistics, Education Longitudinal Study of 2002 (ELS:2002).