ies NATIONAL CENTER FOR SPECIAL EDUCATION RESEARCH
Institute of Education Sciences

The Academic Achievement and Functional Performance of Youth With Disabilities

A Report From the National Longitudinal Transition Study-2 (NLTS2)






NCSER 2006-3000
U.S. DEPARTMENT OF EDUCATION

The Academic Achievement and Functional Performance of Youth With Disabilities

A Report From the National Longitudinal Transition Study-2 (NLTS2)

July 2006

Mary Wagner
Lynn Newman
Renée Cameto
Phyllis Levine
SRI International

Patricia Gonzalez
Project Officer
Institute of Education Sciences

NCSER 2006-3000
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR SPECIAL EDUCATION RESEARCH
Institute of Education Sciences

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Special Education Research
Edward Kame'enui
Commissioner

The National Center for Special Education Research (NCSER) supports a comprehensive research program to promote the highest quality and rigor in research on special education and related services, and to address the full range of issues facing children with disabilities, parents of children with disabilities, school personnel, and others.

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 NCSER product or report, we would like to hear from you. Please direct your comments to

National Center for Special Education Research
Institute of Education Sciences
U.S. Department of Education
555 New Jersey Ave, NW
Washington, DC 20208

The National Longitudinal Transition Study-2 (NLTS2) has been funded by the U.S. Department of Education, Institute of Education Sciences, under contract number ED-01-CO-0003. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

This report was prepared for the Institute of Education Sciences under Contract No. ED-01-CO-0003. The project officer is Patricia Gonzalez in the National Center for Special Education Research.

July 2006

The NCSER World Wide Web Home Page address is http://ncser.ed.gov
The NCSER World Wide Web Electronic Catalog is http://ncser.ed.gov/pubs
The National Longitudinal Transition Study-2 website is http://www.nlts2.org

This publication is only available online. To download, view, and print the report as a PDF file, go to the NCSER World Wide Web Electronic Catalog address shown above.

Suggested Citation

Wagner, M., Newman, L., Cameto, R., and Levine, P. (2006). *The Academic Achievement and Functional Performance of Youth With Disabilities. A Report From the National Longitudinal Transition Study-2 (NLTS2)*. (NCSER 2006-3000). Menlo Park, CA: SRI International.

Content Contact

Patricia Gonzalez
(202) 219-1011
Patricia.Gonzalez@ed.gov

Contents

Appendices

Listing of Tables

Listing of Figures

Executive Summary

To provide a national picture of the academic achievements of American students, the National Center for Education Statistics has administered the National Assessment of Educational Progress (NAEP) periodically since 1969, but there has been no similar national picture of the academic achievement of youth with disabilities. The National Longitudinal Transition Study-2 (NLTS2), funded by the National Center for Special Education Research in the Institute of Education Sciences in the U.S. Department of Education,[1] is filling this gap with information about secondary-school-age students with disabilities. It includes a nationally representative sample of more than 11,000 youth who were ages 13 through 16 and receiving special education services in seventh grade or above in the 2000-2001 school year. NLTS2 is the first national study to include assessments of the academic and functional abilities of youth who receive special education services in secondary school.

One assessment was attempted for each NLTS2 sample member during the biannual data collection cycle in which he or she was in the 16- through 18-year-old age range. The NLTS2 direct assessment uses research editions of subtests of the Woodcock-Johnson III (Woodcock, McGrew, and Mather 2001) that test language arts skills, mathematics abilities, and content knowledge in science and social studies. NLTS2 also includes a functional rating to provide information on youth for whom the direct assessment was reported to be inappropriate. The functional rating is the *Scales of Independent Behavior-Revised* (SIB-R) (Bruininks et al. 1996), a comprehensive measure of adaptive functioning in school, home, employment, and community settings. To determine the form of assessment for which youth qualified, assessors interviewed the school staff person who was most familiar with a youth and his or her school program; information was sought from parents if youth were no longer in school, including any accommodations that a youth required. If a youth did not meet the requirements for the direct assessment, even with accommodations, he or she was eligible for the functional rating, and a rating form was completed by a teacher if a youth was in school or by a parent if he or she was not.

Youth in the direct assessment and functional rating groups do not differ with regard to age, gender, race/ethnicity, or household income. However, the two groups are significantly different in the disability categories they represent, with the category of learning disability having the majority of youth in the direct assessment group, and mental retardation being the most prominent category in the functional rating group. Youth in the functional rating group first were identified as having a disability at a significantly younger age than those of direct assessment participants, and their functional abilities are lower. The functional rating group spends a greater percentage of class time in special education settings and has a higher rate of participation in some kinds of services. In contrast, direct assessment participants spend more time in social activities with friends and in organized groups at school than those in the functional rating group.

---

[1] NLTS2 has been funded with federal funds from the U.S. Department of Education under contract number ED-01-CO-0003. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

This is an executive summary of Wagner, M., Newman, L., Cameto, R., and Levine, P. (2006). *The Academic Achievement and Functional Performance of Youth With Disabilities. A Report From the National Longitudinal Transition Study-2 (NLTS2).* (NCSER 2006-3000). Menlo Park, CA: SRI International.

Results of the NLTS2 direct assessments and functional ratings, the focus of this report, are used to address the following questions:

- How well do youth with disabilities achieve in the areas of language arts, mathematics, science, and social studies?
- How does their achievement compare with the general population of same-age youth?
- What factors related to youths' disability and functioning, individual and household demographics, family support for their education, and previous school experiences are statistically associated with higher academic achievement among youth with disabilities?
- What are the results of the functional ratings of youth's abilities?

Academic Achievement

A considerable gap in achievement in reading, mathematics, science, and social studies exists between youth with disabilities and their peers in the general population.

- Direct assessment results are reported as standard scores, which, for the general population of youth, have a mean of 100 and a standard deviation of 15. In the general population, 50 percent of youth score at the mean of 100 or above and 50 percent score below. In contrast to this distribution for the general population, more than three-quarters of youth with disabilities score below the mean across subtests.
- In the general population, about 2 percent of youth have standard scores that are more than two standard deviations below the mean (i.e., below 70). Among youth with disabilities represented by those who participated in the direct assessment, from 14 percent to 27 percent score more than two standard deviations below the mean across subtests.
- Despite the prevalence of poor scores among youth with disabilities, from 12 percent to 23 percent of youth with disabilities have scores above the mean of 100 for the general population.
- Youth experience the greatest difficulty with reading comprehension; on average, they have a mean score of 79, compared with mean scores that range from 84 for mathematics calculation and social studies content knowledge to 87 for the use of synonyms and antonyms.

Factors Related to the Academic Achievement of Youth With Disabilities

NLTS2 findings reinforce the fact that the academic achievement of youth with disabilities in reading, mathematics, science, and social studies is related to a complex array of factors that characterize youth, their households, and their school experiences. Multivariate analyses demonstrate that several individual factors differentiate youth on the basis of their academic achievement.

- The achievement of youth in several disability categories varies across the academic domains assessed. For example, youth with visual impairments outscore those with learning disabilities (the largest disability category, whose scores dominate the score for youth with disabilities as a whole) on four measures, but are similar on two, independent of other differences between them. Youth with hearing impairments score significantly higher than those with learning disabilities on mathematics calculation but significantly lower on science and social studies content knowledge.
- Youth in the categories of mental retardation and multiple disabilities consistently record low performance scores across the achievement measures.
- Independent of the nature of their disabilities, having higher functional cognitive skills relates strongly and consistently to higher academic achievement. Higher scores on the majority of subtests also are recorded for youth whose disabilities were not manifested until they were older and affect fewer functional domains.

Some demographic and household characteristics also are significantly related to academic achievement, independent of disability-related factors.

- Boys with disabilities score higher than girls on both mathematics and both content knowledge subtests, with differences of 3 or 4 standard score points.
- White youth with disabilities score from 7 to 13 standard score points higher on all academic achievement measures than African American or Hispanic youth with disabilities or those with other racial/ethnic backgrounds.
- Youth with disabilities from low income households (i.e., $25,000 in annual income or less) have lower average scores in all domains relative to youth from moderate income households, independent of racial/ethnic and other differences between them. Differences range from 3 to 5 standard score points.
- Given similar disability, functional, and demographic characteristics, youth with disabilities score from 4 to 6 standard score points higher with each successively higher level of parental expectations regarding their future enrollment in postsecondary school.

Few school experiences of youth with disabilities show statistically significant relationships with youth's academic achievement; students' grades and school mobility and having ever been retained at grade level are not significantly related to academic achievement, independent of other factors considered in the analyses. Two exceptions are:

- Higher absenteeism is associated with lower scores on both mathematics subtests.
- Having had disciplinary problems at school is associated with lower mathematics calculation scores.

In addition, using some kinds of accommodations during the assessment relates to some measures of academic performance, but not in a consistent direction.

- Controlling for other factors, using a calculator provides a 3 or 4 point advantage on the mathematics subtests.

- Using American Sign Language or a sign language interpreter and taking breaks during a session or needing multiple sessions to complete the assessment are associated with lower scores on some subtests.

Functional Ratings of Youth With Disabilities

Youth for whom a functional rating was completed were assessed on four clusters of functional skills (motor skills, social interaction and communication, personal living skills, and community living skills) and on an overall measure of independence.

- Average standard scores for youth with disabilities across the measures range from 43 to 57, compared with a mean of 100 for the general population.
- From 22 percent to 38 percent of youth with disabilities across subtests have scores more than six standard deviations below the mean.
- Across measures, from 11 percent to 15 percent of youth represented by those with a functional rating have scores above the mean for the general population.
- Significantly fewer youth score more than six standard deviations below the mean on personal living skills than on community living skills or the measure of broad independence.
- The few youth with learning disabilities, speech or other health impairments, emotional disturbances, or traumatic brain injuries who have a functional rating together scored higher on the overall measure of broad independence than youth in other disability categories, with a mean of 90.
- The next highest-ranking mean score on the broad independence measure (53) was for youth with hearing impairments; it significantly surpassed the mean scores of all other categories, which ranged from 10 to 23.
- About two-thirds or more of youth with autism, multiple disabilities, visual or orthopedic impairments, or deaf-blindness score more than six standard deviations below the mean on the measure of broad independence.
- Only one statistically significant difference across measures is apparent in the mean standard scores of youth with disabilities who differ in gender, age, household income, or racial/ethnic background, favoring boys over girls on the motor skills measure. Although there were some differences in the percentage of youth in particular standard deviation categories, no consistent patterns were apparent.

Looking Ahead

Future NLTS2 analyses will explore the links between academic performance and both school completion and early postschool outcomes. Those analyses will illuminate the associations between successful learning in school and youth's later ability to continue their education, find employment, and become independent and productive members of their communities, the ultimate goals of secondary education.

1. Assessing the Academic Achievements and Functional Performance of Youth With Disabilities

To provide a national picture of the academic achievements of American students, the National Center for Education Statistics has administered the National Assessment of Educational Progress (NAEP) periodically since 1969. Dubbed "The Nation's Report Card," NAEP involves assessments of the skills of a nationally representative sample of students in reading, mathematics, science, writing, U.S. history, civics, geography, and the arts (National Center for Education Statistics 2005c). The design of NAEP permits trends in achievement to be identified; for example, 2005 NAEP data show that the proportion of eighth-graders who score at or above the level of basic mathematics skills has increased by 17 percentage points since 1990, to 69 percent (National Center for Education Statistics 2005a).

Although valuable in charting the academic performance of the general population of students, NAEP excludes students with disabilities if "the student's IEP team determines that the student cannot participate; or the student's cognitive functioning is so severely impaired that she or he cannot participate; or the student's IEP requires that the student has to be tested with an accommodation or adaptation that NAEP does not allow" (National Center for Education Statistics 2005b). Examples of disallowed accommodations or modifications are Braille materials, tape recorders, use of a calculator or arithmetic tables, or having the test read or responses given in a student's native language (National Center for Education Statistics 2005b). Further, the scores of youth with disabilities who are included in NAEP are not reported separately, so to date, there remains no national picture of the academic achievements of youth with disabilities.

The National Longitudinal Transition Study-2 (NLTS2), funded by the National Center for Special Education Research of the Institute of Education Sciences in the U.S. Department of Education, is filling this important gap in information about secondary-school-age students with disabilities.

Research Questions

In this report, NLTS2 findings are used to address the following questions regarding the academic achievement and functional performance of youth with disabilities:

- How well do youth with disabilities perform in the areas of language arts, mathematics, science, and social studies?
- How does their performance compare with the general population of same-age youth?
- What factors related to youths' disability and functioning, individual and household demographics, family support for their education, and previous school experiences are statistically associated with higher academic performance among youth with disabilities?
- What are the results of the functional ratings of youth's abilities?

As context for interpreting the findings related to these questions, the following sections of this chapter provide a brief overview of the NLTS2 design and sample. The assessments and

other data sources relevant to the report are described briefly, as are the characteristics of the youth for whom findings are reported.

Study Overview

NLTS2 is a 10-year-long study of the characteristics, experiences, and outcomes of a nationally representative sample of youth with disabilities who were ages 13 through 16 and receiving special education services in grade 7 or above on December 1, 2000. The study is designed to collect data on sample members from multiple sources in five waves—i.e., every 2 years from 2001 to 2009.

The NLTS2 sample was constructed in two stages. A stratified random sample of school districts was selected from the universe of approximately 12,000 that serve students receiving special education in at least one grade from 7th through 12th grades. These districts and 77 state-supported special schools that served primarily students with hearing and vision impairments and multiple disabilities were invited to participate in the study, with the intention of recruiting approximately 500 districts and as many special schools as possible from which to select a target sample of about 12,000 students. Recruitment efforts resulted in 501 school districts and 38 special schools agreeing to participate and providing rosters of students receiving special education in the designated age range, from which the student sample was selected.

The roster of all students in the NLTS2 age range who were receiving special education from each district and special school was stratified by primary disability category, as reported by the districts. Students then were selected randomly from each disability category. Sampling fractions were calculated that would produce enough students in each category so that, in the final study year, findings will generalize to most categories individually with an acceptable level of precision, accounting for attrition and for response rates to the parent/youth interview. A total of 11,276 students were selected and eligible to participate in NLTS2.

Data Sources

The five data sources that provide the information reported here are[1]

- a direct assessment of the academic achievement of youth with disabilities;
- an adult-completed rating of the functional performance of youth for whom the direct assessment was reported to be inappropriate;
- a telephone interview with parents of both groups of youth;
- self-administered surveys of school staff serving individual sample members; and
- school districts' reports of the primary disability category in which students were provided special education services when selected for the study.

Each source is described briefly below and discussed in greater detail in appendix A.

[1] Table A-9 in appendix A identifies the data source for each variable included in analyses in this report.

Youth Assessments

Because performance assessments can be labor intensive and costly, the NLTS2 design calls for only one assessment per sample member. An assessment was attempted for each NLTS2 sample member for whom a telephone interview or mail questionnaire had been completed by a parent and parental consent for the assessment had been provided. Youth were eligible for an assessment during the data collection wave in which they were 16 through 18 years old.[2] This age range was selected to limit the variability in performance that could be attributed to differences in the ages of the youth participating and to mesh with the every-2-year data collection cycle of the study. The study design linked the timing of assessments with school data collection (conducted in 2002 and 2004) because most direct assessments took place at school and most functional ratings were completed by teachers. The oldest two single-year age cohorts of youth (i.e., those ages 15 or 16 when sampled) reached the eligible age range in Wave 1 (2002); 5,071 youth met the eligibility criteria for assessment at that time. The younger two cohorts (those ages 13 or 14 when sampled) reached the eligible age range when Wave 2 school data were collected; 4,343 youth met the criteria in 2004.

Two forms of assessment are included in NLTS2 (described below) to accommodate the wide range of abilities among youth with disabilities. For each form of assessment, data are combined across two waves of data collection and reported here for all youth for whom each assessment was completed in either wave. A total of 5,222 youth participated in the NLTS2 direct assessment, and a functional rating was completed for 1,051 youth across the two waves. Taken together, this yields an overall assessment rate of 67 percent of eligible youth.[3]

Direct assessment. The direct assessment included in NLTS2 was selected by a workgroup of assessment and measurement experts over a 6-month period in 2000. The resulting direct assessment uses research editions of subtests of the Woodcock-Johnson III (WJ III); (Woodcock, McGrew, and Mather 2001) that test

- language arts skills (passage comprehension and the use of synonyms and antonyms);
- mathematics abilities (calculation procedures and solving applied problems); and
- content knowledge in science and social studies.

The research editions are shorter versions of the standard WJ III assessment battery and were developed for use in NLTS2 by the original WJ III developers (please see chapter 2 for additional information on the research versions). The WJ III is a comprehensive, norm-referenced, individually administered assessment of the academic skills and knowledge routinely used in school and other settings. The WJ III tests have strong psychometric properties (Cizek 2001) and are appropriate for administration to children as young as 2 years of age and to adults as old as 90. The WJ III subtests are particularly advantageous for NLTS2 because they permit comparisons with a general population norm group assessed in 2000.

---

[2] Wave 1 assessments also included 10 youth whose assessments were not completed until shortly after their 19th birthdays.

[3] Similar direct and alternate assessment procedures were used to assess elementary and middle school students with disabilities in the Special Education Elementary Longitudinal Study (SEELS). NLTS2 response rates were in the mid-range of the SEELS rates, which were 63 percent, 73 percent, and 92 percent across the three waves of SEELS assessments.

Functional rating. NLTS2 includes a functional rating for youth with disabilities for whom the direct assessment was reported to be inappropriate because their sensory, physical, behavioral, or cognitive disabilities made them unable to follow instructions or answer questions reliably in spoken or written English, Braille, or large print. The functional rating instrument is the adult-completed *Scales of Independent Behavior-Revised* (SIB-R) (Bruininks et al. 1996). The SIB-R is a comprehensive measure of adaptive and problem behaviors related to functional independence and adaptive functioning in school, home, employment, and community settings. Its 14 18- to 20-item subtests focus on

- motor skills;
- social interaction and communication skills;
- personal living skills; and
- community living skills.

These four clusters also are combined into an overall scale referred to as "broad independence" (see chapter 5). The SIB-R has norm samples to allow comparison with the general population.

Parent/Guardian Interviews

Unlike the youth assessments, which were administered once for each youth, data are collected repeatedly from parents and, beginning in Wave 2, from youth themselves. Data from both Waves 1 and 2 are included in this report as independent variables in multivariate analyses and as descriptors of the youth in the direct assessment and functional rating groups.

Wave 1 parent/guardian interview/survey. The NLTS2 conceptual framework suggests that a youth's nonschool experiences, such as extracurricular activities and friendships; historical information, such as age when disability was first identified; household characteristics, such as socioeconomic status; and a family's level and type of involvement in school-related areas are crucial to youth outcomes. Parents/guardians are the most knowledgeable about these aspects of youth's lives. They also are important sources of information on outcomes across domains. Thus, parents/guardians of NLTS2 sample members were interviewed by telephone or surveyed by mail[4] in 2001, as part of Wave 1 data collection.

Wave 2 parent/guardian interview and youth interview/survey. NLTS2 sample members for whom working telephone numbers and addresses were available were eligible for the Wave 2 parent/youth telephone interview in 2003. The major distinction between the data collection methods in Waves 1 and 2 is that in the latter wave, interviews were sought both with parents of NLTS2 sample members and with the youth themselves if parents reported they were able to respond to questions by phone or on a mailed questionnaire.

[4] A mail survey was conducted for sample members whose parents could not be reached by telephone.

School Surveys

Some school-related data included in multivariate analyses in chapter 4 come from the following:

- *The NLTS2 Student's School Program Survey*. This mail survey was administered to school staff who were most knowledgeable about the overall school programs of NLTS2 sample members who attended their school.
- *The General Education Teacher Survey*. Respondents to this mail survey were teachers of general education academic classes attended by NLTS2 sample members, if students took such a class.

The surveys collected information about aspects of the classroom experiences of students with disabilities in general education academic classes and in vocational education and special education settings. Both surveys were administered in Waves 1 and 2 for youth still in secondary school at those times. If a youth's direct assessment was conducted in Wave 1, independent variables used school survey data from that wave; similarly, Wave 2 school survey data were used in analyses of youth whose assessment was conducted in that wave.

School- and School-District-Identified Primary Disability Category

Information about the primary disability category of NLTS2 sample members came from rosters of students in the NLTS2 age range receiving special education services in the 2000-01 school year under the auspices of participating school districts and state-supported special schools.[5]

Youth Included in the Report

The two groups of youth who are the focus of this report are distinguished by the form of their assessment. As mentioned above, membership in the two groups was determined from reports by school staff or parents regarding the feasibility and appropriateness of the direct assessment for specific youth. Those for whom it was not considered feasible or appropriate had a functional rating completed by a teacher who was familiar with their abilities if the youth were still in school, or by parents if they were no longer in school. Thus, the groups of youth with disabilities for whom the two forms of assessment were completed are intentionally different with regard to their abilities.[6] Appendix B provides detailed information regarding group differences, examples of which are summarized briefly here.

There are no statistically significant differences between youth with direct assessments and functional ratings with regard to age, gender, race/ethnicity, or household income. However, the two groups are significantly different in the disability categories they represent. For example, almost two-thirds of the direct assessment group are in the learning disability category, youth who comprise only about one-fourth of youth with functional ratings. In contrast, the percentage of youth in the mental retardation category is more than four times higher among youth with

[5] The definitions of the 12 primary disability categories used here are defined by law and presented in table A-10, appendix A.

[6] Note that there is no standardized measure of abilities that is common to the two groups on which their performance can be compared.

functional ratings than with a direct assessment ($p < .001$). The average age at which youth in the functional rating group first were identified as having a disability is significantly lower than those of direct assessment participants ($p < .01$), as are their functional abilities ($p < .001$). Differences in the characteristics of the school programs and the types of related services and supports received by youth in the direct assessment and functional rating groups include greater special education course taking ($p < .001$) and participation in some kinds of services (i.e., $p < .001$ comparing rates of receipt of speech/language pathology, occupational or physical therapy, or transportation services) among youth with a functional rating compared with direct assessment peers. Differences in the social experiences of the two groups entail greater involvement among direct assessment participants in social activities with friends, ($p < .001$) and in organized groups at school ($p < .05$).

Technical Notes

Readers should remember the following issues when interpreting the findings in this report

- **Findings are weighted.** NLTS2 was designed to provide a national picture of the characteristics, experiences, and achievements of youth with disabilities in the NLTS2 age range as they transition to young adulthood. Therefore, all the statistics presented in this report are weighted estimates of the national population of students receiving special education in the NLTS2 age group, and of each disability category individually. Each response for each sample member is weighted to represent the number of youth nationally that are in his or her disability category in the kind of school district (defined by region, student enrollment, and proportion of students in poverty) or special school from which he or she was selected.
- **Standard errors.** For each mean and percentage in this report, a standard error is presented that indicates the precision of the estimate. For example, a variable with a weighted estimated value of 50 percent and a standard error of 2.00 means that the value for the total population, if it had been measured, would, with 95 percent confidence, lie between 46 percent and 54 percent (i.e., within plus or minus 1.96 x 2 percentage points of 50 percent). Thus, smaller standard errors allow for greater confidence to be placed in the estimate, whereas larger ones require caution.
- **Small samples.** Although NLTS2 data are weighted to represent the population, the size of standard errors is influenced heavily by the actual number of youth in a given group (e.g., a disability category). Groups with very small samples have comparatively large standard errors (in fact, findings are not reported separately for groups that do not include at least 35 sample members). For example, because there are relatively few youth with deaf-blindness, estimates for that group have relatively large standard errors. Therefore, readers should be cautious in interpreting results for this group and others with small sample sizes and large standard errors.
- **Significant differences.** In discussions of the descriptive statistics, only differences among groups that reach a level of statistical significance of at least .05 in a two-tailed *F* test are mentioned in the text, accompanied by the specific significance levels.

Organization of the Report

Chapter 2 provides additional information on key aspects of the data and analysis approaches reported in this document. Chapter 3 presents the standard scores of youth with disabilities on the language arts, mathematics, science, and social studies WJ III subtests. Results of multivariate analyses that identify factors associated with variations in academic performance are presented in chapter 4, and chapter 5 describes the performance of youth with disabilities on the seven dimensions included in the functional rating. Chapter 6 summarizes key points from the report. Appendix A details the sample design and sample weighting strategies, sources of data for variables used in the analyses, and analysis approaches. Appendix B reports additional information on the characteristics of the youth participating in the direct assessment and for whom a functional rating was completed. Appendix C supplements the multivariate analysis results reported in chapter 4.

2. Measurement and Analysis Considerations

Although appendix A provides details on many aspects of the NLTS2 design, measurement, and analysis approaches, this chapter highlights the following, which are particularly important in helping readers interpret the findings reported in subsequent chapters:

- the research versions of the direct assessment subtests;
- determining the type of assessment to be administered;
- assessment procedures;
- analysis approaches; and
- the population of youth with disabilities to whom the findings generalize.

Research Versions of the Woodcock-Johnson III Subtests

As mentioned in chapter 1, the NLTS2 direct assessment employed research versions of the Woodcock-Johnson III (WJ III) subtests for reading comprehensions, synonyms and antonyms, mathematics calculation, applied mathematics problem solving, and content knowledge in social studies and science. The research and published (i.e., standard) versions of the subtests share items and administration procedures. The difference between them lies in the larger number of items used in the standard version; the time (and, therefore, expense) of the standard version precluded its use for the large NLTS2 sample.

The research versions were created by the original test developers by reducing the item density from approximately three items per 10 W score points for the published version to one or two items for the research version, depending on the subtests. This is possible without changing the scoring or interpretation of the subtests because the WJ III is based on the Rasch model (Andrich 1988; Wright and Stone 1979), which allows for item-free measurement. Once the pool of items for a subtest is scaled per item response theory, different subsets that differ in item number and content can be used to create different versions of the test, with all subsets based on the same underlying Rasch-scaled measurement scale. Thus, the shorter research versions tests produce scores on the same scale as the full-length test and use the same national norms as those that underlie the published full-length tests. In addition to reducing the number of items, testing time also was reduced by changing the criteria for establishing basal and ceiling points from six consecutive correct items and six consecutive incorrect items, respectively, to three items.

Tests designed with these specifications have an average reliability of .65 and a standard error of measurement (SEM) of 10.0, in contrast to .85 and 5.7 for the publication-length tests. Although the individual SEMs are much larger for the research version, the important statistic for large-scale group analyses is the standard error of the mean, not the SEM, because the results are not used for individual programming decisions (e.g., eligibility for special education services). The standard error of the mean is a function of the standard deviation and sample size. Thus, if a test has a typical standard deviation of 15 W points, in a sample of 1,000, the standard error of the mean would be approximately 0.5, an acceptable level for calculating the group-level statistical estimates used in NLTS2.

Determining the Form of Assessment

Whether an age-eligible youth was administered a direct assessment or an adult was asked to complete a functional rating for him or her was determined through a screening process. For in-school youth, assessors conducted a telephone or in-person screening interview with the school staff person who was most familiar with a youth and his or her school program; in 91 percent of cases, this person was a special educator. Screening interviews were conducted with parents if youth were no longer in school. Screening information was used to determine whether a youth was able to participate in the direct assessment. To do so, a youth needed to be able to understand directions given in spoken English, large print, Braille, or sign language; have a consistent response mode (i.e., the assessor could reliably understand the youth's responses);[1] and the ability to work with an assessor or with someone who was familiar to the youth and who could and would conduct the assessment in the presence of the assessor.[2] If a youth met these criteria, additional questions were asked to determine which components of the assessment the youth could be administered.[3] The screening interview also sought to identify any accommodations that a youth required for the direct assessment. If a youth did not meet the requirements for the direct assessment, even with accommodations, he or she was deemed eligible for the functional rating.

Assessment Procedures

Direct Assessment

Hiring and training assessors. Assessors typically were school psychologists or teachers and were recruited in the geographic areas of eligible youth. Approximately 800 assessors were used in each wave of data collection, with the majority of 2002 assessors returning for the 2004 administration. Potential assessors submitted resumes and participated in a telephone interview to determine that they had experience conducting assessments of students with disabilities. The training of successful applicants consisted of reviewing the Field Assessors Guide, training

---

[1] School staff or parents were told, "The assessment requires that students (or youth) answer questions reliably," and then asked, "Is [YOUTH] able to reliably answer questions?"

[2] School staff or parents were asked, "Would [YOUTH] be able to answer questions asked by someone he/she doesn't know?" If the response was "no," the staff person was asked, "Would [YOUTH] be able to answer questions asked by someone he/she doesn't know if someone he/she knew was in the room?" If the response was "no," the person was asked, "Would [YOUTH] be able to answer questions if someone he/she knew asked the questions?" If the response was "yes," the person was asked, "Is there a person [YOUTH] knows available to conduct the assessment?" Across the two waves, 75 youth were reported to need a familiar adult to be present during or to administer the assessment in the presence of the assessment administrator. No statistically significant differences between youth reported to require such support and those not so reported were found on demographic factors; disability category; self-care, social, or functional cognitive skills; or mean standard scores on any assessment subtest. However, youth reported to need the presence of a familiar adult to complete the direct assessment were significantly more likely to have had their disabilities identified at birth than youth who were not reported to need this form of assistance (74 percent vs. 11 percent, $p < .01$) and were less likely to have had their disabilities identified at age 6 or older (7 percent vs. 37 percent, $p < .05$).

[3] To determine participation in the synonyms/antonyms and content knowledge subtests, school staff or parents were asked: "Can [YOUTH] read simple printed [or Braille] words, like 'road' or 'big'?" To determine participation in the mathematics calculation and applied problems subtests, the person was asked, "Can [YOUTH] recognize printed [or Braille] numbers?" To determine participation in the passage comprehension subtest, the person was asked, "Can [YOUTH] read written [or Braille] sentences?"

video, and the testing materials (WJ III direct assessment presentation "easel," test manual, and scoring booklets; functional rating scale; and screening interview questionnaire), and completing with 100 percent accuracy a self-administered test on the information and material presented in the Guide, video, and testing materials.

Each field assessor was assigned to a supervisor, who was available to answer questions about test administration, oversaw the training process, and reviewed and verified the successful completion of the field assessor test. When assessors successfully completed the training, they signed a work agreement and confidentiality pledge and were provided contact information, consent forms, and other assessment materials for the eligible youth in their area.

Conducting assessments. For youth who were able to participate in the direct assessment and who were still in school, the assessments generally were conducted at school when students were not in class. Some out-of-school youth also were assessed at the school they had once attended, but assessments for many out-of-school youth were conducted at youth's homes or in community settings.

Assessors contacted schools and parents to locate youth, identify a staff person who knew the youth well with whom to conduct the screening interview, and arrange for the appropriate assessment to be completed. The screening and direct assessment instruments (i.e., instructions and individual items) were fully scripted to maintain consistency of administration across assessors. Possible response choices for each item and instructions for scoring and for establishing basal and ceiling scores were included in the assessment easel, test manual, and scoring booklet. In the scoring booklet, assessors indicated only whether an item was answered correctly; all other scoring functions were conducted by NLTS2 project staff when booklets were submitted after completion of the assessment.

Use of accommodations. On the basis of recommendations of the assessment design panel and to be consistent with principles[4] underlying the inclusion of students with disabilities in standardized assessments (Thurlow, Quenemoen, Thompson, and Lehr 2001), the NLTS2 direct assessment procedure was designed to mirror students' day-to-day instruction and test participation with regard to the use of accommodations—i.e., a youth participating in the direct assessment was offered the same accommodations called for in his or her IEP for instruction and testing. The screening questionnaire requested information regarding a youth's need to take breaks during testing; use special furniture or lighting; have aides or assistants help with testing; or use American Sign Language (ASL), Braille, large print materials, or an abacus or calculator.

However, the design panel acknowledged that the nature of the WJ III as an untimed, individually administered test would make most accommodations used in state accountability testing (e.g., more time to complete the test) unnecessary. Consistent with this view, fewer accommodations were actually requested for youth in the assessment than they received in classes. Overall, 61 percent of youth received no accommodations, 28 percent received one accommodation, and 11 percent received two or more. The rate of receipt of specific accommodations is as follows: breaks (8 percent), special furniture or lighting (5 percent), an aide or assistant (5 percent), an ASL interpreter (8 percent), Braille (6 percent), and abacus or

[4] "Principle 3. All students with disabilities are included when student scores are publicly reported, in the same frequency and format as all other students, whether they participate with or without accommodations, or in an alternate assessment" (Thurlow, Quenemoen, Thompson, and Lehr 2001, p. 3).

calculator (23 percent). Those who participated in the direct assessment with one or more accommodations do not differ significantly from those who did not in disability-related factors, demographics, or mean standard scores on any direct assessment subtest.

It is important to note that norms for the WJ III were established for the general population, who would not need the accommodations provided to participants in the NLTS2 direct assessment. Thus, the NLTS2 procedures represent a departure from the standard WJ III procedures, but one that was deemed appropriate by the design panel for the population being assessed; without accommodation, some youth would have been unable to demonstrate competencies they in fact possessed, biasing downward measures of true achievement. The actual effects of providing a particular accommodation could be measured only by providing it to some students who required it and withholding it from others. However, analyses reported in chapter 4 attempt to estimate the relationship between provision of accommodations and academic achievement by including measures of their use in multivariate analyses, along with variables intended to control for variations in disability-related factors that could act as a proxy for need for such accommodations.

Functional Rating

If screening information indicated the direct assessment was inappropriate for a youth, a functional rating form and instructions for its completion were sent by the assessor to the youth's teacher if he or she was in school or to a parent if he or she was no longer in school or if the school would not participate in the assessment and rating process.[5] Assessors followed up with recipients to ensure an acceptable response rate. Completed rating forms were returned directly to NLTS2 project staff in postage-paid envelopes provided for that purpose. Respondents were compensated at the rate of $30 for each completed functional rating.

Analysis Approaches

Analyses reported in this document involve simple descriptive statistics (e.g., frequencies, means), correlational methods (i.e., cross-tabulations), and multivariate models (i.e., ordinary least squares regression). With the exception of seven variables that are included in the multivariate models, these analysis approaches exclude cases with missing values; imputation conducted for the seven exceptions is described in appendix A.

Regarding cross-tabulations, statistically significant differences between subgroups (e.g., youth in different disability categories) are identified using *F* tests. This approach has been followed because in all cases, the intent is to identify significant differences between two specific groups (e.g., youth with learning disabilities and those with mental retardation), rather than identifying a more general "disability effect" on the distribution of the variable of interest. In the case of unweighted data, comparing two percentages is usually accomplished using nonparametric statistics, such as the Fisher exact test. In the case of NLTS2, the data are weighted, and the usual nonparametric tests would yield significance levels that are too small,

[5] Approximately 22 percent of youth with a functional rating are estimated to have had it completed by a parent. Only one statistically significant difference between those with ratings that were parent-completed and those with ratings completed by teachers was noted on the variety of demographic and disability-related factors and assessment scores examined. The group with ratings completed by parents had a significantly larger proportion of African American youth than the group with teacher-completed ratings (43 percent vs. 14 percent, $p < .05$).

because the NLTS2 effective sample size is less than the nominal sample size. The p-values for the test statistic used as an alternative approach to determine statistical significance are derived from an *F*(1, infinity) distribution (i.e., a chi-square distribution with one degree of freedom).

Multiple linear regression techniques are used in this report to assess the independent relationships between ordinal measures of academic achievement and characteristics of individual youth, their households, and their school program and experiences.[6] NLTS2 multivariate analyses and correlations are unweighted. Results are reported for analyses that include the full set of individual, household, and school factors simultaneously. This approach allows the modeling of the simultaneous influence of predictor variables on the dependent variable and provide estimates of model fit.

Youth to Whom Findings Generalize

As noted in chapter 1, the universe to which the NLTS2 sample generalizes is a cohort of students who were ages 13 through 16 and received special education services in grade 7 or above in participating schools and school districts as of December 1, 2000. Weights for analyses reported in this document are calculated so that all youth with either a direct assessment or a functional rating, taken together, generalize to that cohort, without regard to when the assessment was done or which form of assessment was done.

To illustrate, consider the following groups:

A = The entire NLTS2 sample.

A1 = The portion of A who are ages 16 through 18 as of the Wave 1 assessment.[7]

A1a = The portion of A1 who would be able to participate in the direct assessment.

A1b = The portion of A1 for whom the functional rating is more appropriate to their abilities.

A2 = The portion of A who are ages 16 through 18 as of the Wave 2 assessment.

A2a = The portion of A1 who would be able to participate in the direct assessment.

A2b = The portion of A1 for whom the functional rating is more appropriate to their abilities.

For each of these sample groups, there is a corresponding group in the universe, which can be denoted with a "B," such that the universe is B, the portion of the universe that is 16 through 18 as of the Wave 1 assessment is denoted B1, the portion of B1 who would be able to participate in the direct assessment is denoted B1a, etc. The sizes of these universe subgroups can be estimated by weighting all youth in A (as if they all had been respondents) up to the entire

[6] Multiple linear regression equations involve a linear combination of a set of independent variables in the following algebraic form: $Y' = a + b_1X_1 + b_2X_2 + ... + b_nX_n$, where Y' is the predicted value of the dependent variable, a is the constant or Y intercept, bs are the partial regression coefficients, and Xs are the values of the independent variables.

[7] This group also includes 10 students who had recently become 19 at the time of their Wave 1 assessment.

universe, B. Then the sum of the weights of youth in A, A1, A1a, A1b, etc. are estimates of the number of youth in B, B1, B1a, B1b, etc.

However, not all students in A1a, A1b, etc. were respondents. Let respondents in each subgroup be denoted by appending an "r" to the label (e.g., A1ar, A1br, etc.). Then weights can be computed (adjusting for various youth and school characteristics used as stratifying or post-stratifying variables) that project A1ar up to B1a, A1br up to B1a, A2ar up to B2a, and A2br up to B2b, that is

- Youth who participated in the direct assessment in Wave 1 represent the portion of the universe who were 16 to 18 as of the Wave 1 assessment and would be able to participate in the direct assessment.
- Youth for whom a functional rating was completed in Wave 1 represent the portion of the universe who were 16 to 18 as of the Wave 1 assessment and whose abilities would make the functional rating appropriate.
- Youth who completed the direct assessment in Wave 2 represent the portion of the universe who were 16 to 18 as of the Wave 2 assessment and would be able to participate in the direct assessment.
- Youth for whom a functional rating was completed in Wave 2 represent the portion of the universe who were 16 to 18 as of the Wave 2 assessment and whose abilities would make the functional rating appropriate.

Additional technical information is presented in appendix A.

3. The Academic Achievement of Youth With Disabilities in Reading, Mathematics, Science, and Social Studies

NLTS2 provides the first nationally representative data on how secondary-age youth with disabilities are faring academically. This chapter presents descriptive findings from standardized assessments of youth with disabilities in reading, mathematics, science, and social studies. It begins with a description of the six assessment subtests in these areas, and then continues with a description of performance on these measures by youth with disabilities as a group and by those who differ in their primary disability category.

Assessment Subtests

As noted in chapter 1, assessments of youths' academic achievement were conducted using six subtests from the research edition of the Woodcock-Johnson III (WJ III); (Woodcock, McGrew, and Mather 2001).[1] Two relate to reading (passage comprehension and synonyms and antonyms), two to mathematics (calculations and applied problems), and one each to science and social studies. These six subtests are described below (Mather and Woodcock 2001).

Passage comprehension. The research version of the WJ III passage comprehension subtest presents youth with a series of items that range in difficulty. The least difficult items present a phrase in conjunction with several graphic representations. Youth point to the appropriate picture that matches the phrase (e.g., two trees). The more difficult items are entirely text-based, address more technical topics, and require both greater vocabulary and the ability to make inferences from context. In this section of the subtest, youth read a short passage and then provide the missing key word that makes sense in the context of that passage. Youth who perform well on this test have well-developed linguistic and cognitive skills, in addition to the ability to notice and use textual information.

Synonyms and antonyms. The research version of this WJ III subtest assesses skills in reading words, understanding vocabulary, and supplying words with similar or opposite meanings. The first part of the subtest requires reading a word and providing a synonym (i.e., a word with the same meaning); the second requires reading a word and providing an antonym (i.e., a word with the opposite meaning).

Mathematics calculation. The research version of the calculation subtest assesses computation skills, ranging in difficulty from elementary (e.g., simple addition) to advanced (e.g., integrating a function). Youth are given a worksheet that presents the mathematics problems. Items are not read to the youth unless that is a typical accommodation received by the youth. Because the calculations are presented in a traditional problem format, including notation that signals the operation that is required to produce the correct result (e.g., a + for addition), youth are not required to decide about what operations to use or what data to include. Youth are required to perform addition, subtraction, multiplication, division, and combinations of these basic operations, and some geometric, trigonometric, logarithmic, and calculus operations. The calculations involve negative numbers, percents, decimals, fractions, and whole numbers. The

[1] Chapter 1 of this report includes a description of the direct assessment data collection process; additional details regarding the assessment instruments and procedures are found in appendix A.

least difficult items are simple, single-digit addition problems; the most difficult require knowledge of calculus.

Applied problems. The research version of the WJ III applied problems subtest requires youth to analyze and solve practical mathematical problems that are read to them. To solve the problems, youth must recognize the procedure to be followed and then perform relatively simple calculations. Because many of the problems include extraneous information, the youth must decide not only the appropriate mathematical operations to use but also which numbers to include in the calculation. Item difficulty increases with complex calculations. All youth are provided with and may use calculators, pencil, and paper.

Science. This subtest assesses knowledge of various areas of biological and physical sciences. The items range in difficulty. Early items require a youth simply to point to the appropriate response; remaining items require a youth to respond orally to questions read to him or her.

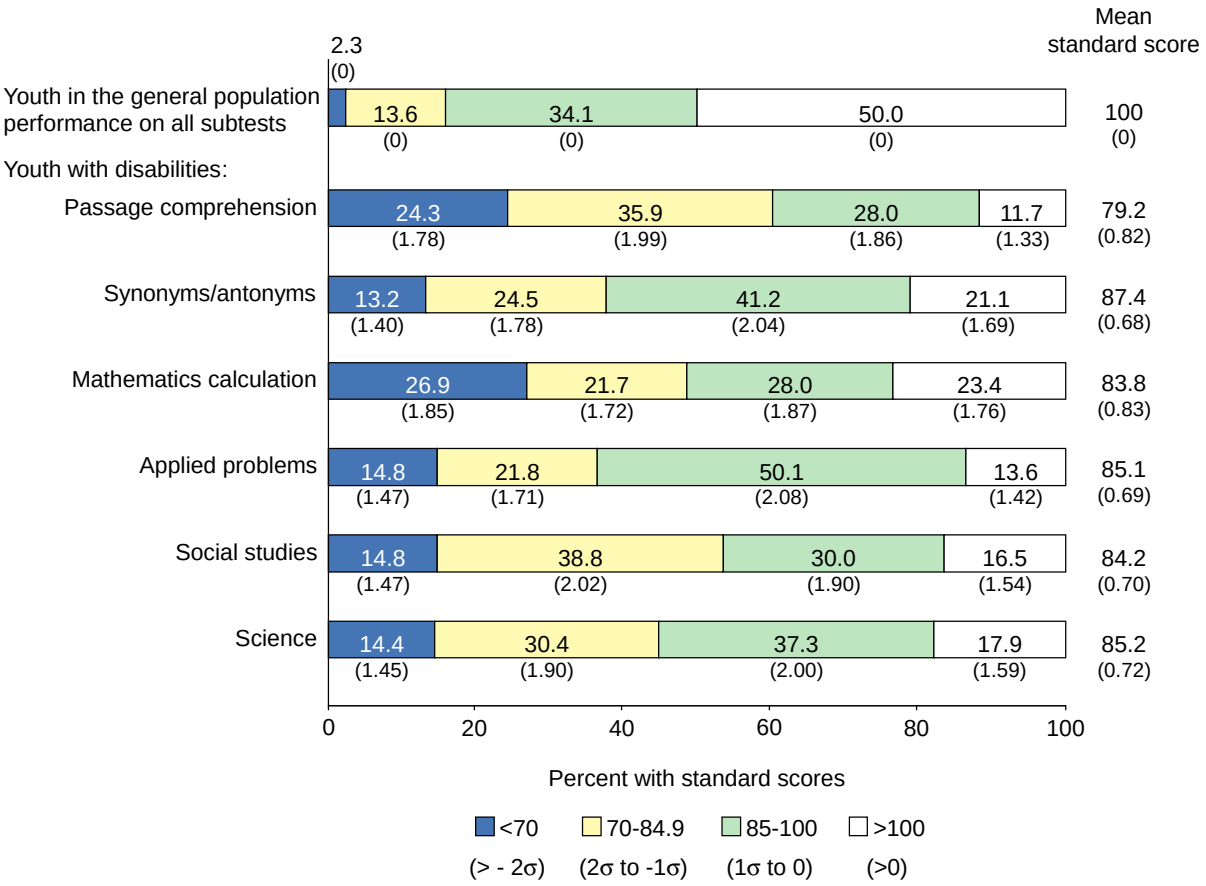
Social studies. The research version of the WJ III social studies subtest assesses knowledge of history, geography, government, economics, and other aspects of social studies. Similar to the science content knowledge subtest, early items require only a pointing response, whereas remaining items require a youth to respond orally to questions read to him or her. Items range in difficulty from early preschool through college.

Youth's Academic Achievement

Scores on the WJ III subtests suggest that many youth with disabilities do not fare nearly as well on these academic assessments as peers in the general population. Figure 1 presents the distribution of standard scores for youth with disabilities and those in the general population. Direct assessment scores are reported as standard scores, which have a mean of 100 and a standard deviation of 15. In the general population, the distribution of test scores on each subtest is equally divided above and below the mean (i.e., 50 percent score above and 50 percent below) (Woodcock and Johnson 1989). In comparison, more than three-quarters of those with disabilities score below the mean across subtests. Compared with the 50 percent of youth in the general population who score 100 or below, 86 percent of youth with disabilities receive standard scores in that range on the applied problems subtest, 83 percent on the passage comprehension and social studies subtests, 82 percent on the science subtest, 79 percent on the synonyms/antonyms subtest, and 77 percent on the mathematics calculation subtest ($p < .001$ for all comparisons).

Approximately 2 percent of youth in the general population receive standard scores that are more than two standard deviations below the mean—i.e., less than 70—a standard score range classified by WJ III as being "very low" (Woodcock and Mather 1990). Across the subtests, between 14 percent and 27 percent of youth with disabilities have scores in this range. Compared with the 2 percent of youth in the general population who score below 70, 27 percent of youth with disabilities do so on the mathematics calculation subtest, as do 24 percent on the passage comprehension subtest, 15 percent on the applied problems and social studies subtests, 14 percent on the science subtest, and 13 percent on the synonyms/antonyms subtest ($p < .001$ for all comparisons).

Figure 1. Performance of youth with disabilities compared with performance of youth in the general population on Woodcock-Johnson III subtests



NOTE: Standard errors are in parentheses.

SOURCE: Woodcock-Johnson, *Tests of Cognitive Ability: Standard and Supplemental Batteries*, Norm Tables, 1989; U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), student assessments, 2002 and 2004.

Despite one out of four to one out of seven youth with disabilities scoring "very low" on these subtests, some youth are performing well. From 12 percent to 23 percent of youth with disabilities score above 100 across subtests. Youth are more likely to score above the mean on the synonyms/antonyms and mathematics calculation subtests, with 21 percent and 23 percent performing above 100 on them. Youth are less likely to perform well on the passage comprehension and applied problems subtests, with 12 percent and 14 percent scoring above the mean ($p < .001$ for all comparisons). Youth experience the greatest difficulty with passage comprehension. On average, they receive a score of 79 on this subtest, compared with scores of 87 on the synonyms/antonyms subtest, 85 on the science and applied problems subtests, and 84 on the social studies and mathematics calculation subtests ($p < .001$ for all comparisons).

Disability Category Differences in Academic Achievement

Academic achievement differs considerably across disability categories (table 1). Mean standard scores range from 56 and 62 on the passage comprehension subtest for youth with mental retardation and multiple disabilities to 93, 94, and 95 on the synonyms/antonyms subtest for those with emotional disturbances, visual impairments, or other health impairments ($p < .001$ for all comparisons). Within disability categories, performance varies across assessment subtests—youth in each category perform better in some academic areas than others.

Other health impairment. Average assessment scores for youth with other health impairments range from 86 to 95. They achieve their highest average score on the synonyms/antonyms subtest (95 vs. 86 to 90 across subtests, $p < .001$ for all comparisons). Youth with other health impairments also receive higher average scores on the science subtest (90) than on the passage comprehension subtest (86, $p < .01$).

Visual impairment. Youth with visual impairments have scores that range from 85 to 94 across subtests. They receive higher scores on the synonyms/antonyms (94) and mathematics calculations subtests (92) than on the passage comprehension subtest (85, $p < .01$ and $p < .05$). They also perform better on the synonyms/antonyms subtest than on the applied problems subtest ($p < .05$).

Emotional disturbance. Youth with emotional disturbances receive average scores from 84 to 93. They score higher on the synonyms/antonyms subtest (93) than on all other tests (84 to 89, $p < .001$ for all comparisons other than $p < .01$ for science subtest). Youth with emotional disturbances also receive higher scores in science (89) and applied problems (88) than passage comprehension (84, $p < .01$ and $p < .05$).

Table 1. Mean standard scores of youth with disabilities on Woodcock-Johnson III subtests, by disability category

Subtests	Other health impairment	Visual impairment	Emotional disturbance	Learning disability	Speech/ language impairment	Orthopedic impairment	Hearing impairment	Traumatic brain injury	Autism	Deaf-blindness	Multiple disabilities	Mental retardation
	Mean standard score / standard error											
Passage comprehension	85.8 (1.10)	84.7 (2.33)	84.2 (1.42)	81.9 (1.00)	81.4 (1.15)	78.8 (1.59)	75.6 (1.73)	74.1 (2.96)	69.6 (2.38)	66.3 (3.81)	61.5 (2.66)	55.7 (1.41)
Synonyms/antonyms	95.0 (0.86)	94.0 (1.89)	93.4 (1.12)	89.5 (0.81)	89.9 (0.93)	88.2 (1.23)	84.1 (1.44)	83.7 (1.95)	81.3 (2.16)	75.5 (2.88)	71.6 (2.11)	65.3 (1.06)
Mathematics calculation	88.2 (1.07)	92.2 (2.41)	86.2 (1.22)	86.1 (1.09)	91.7 (1.14)	82.6 (1.64)	91.5 (1.42)	80.0 (2.65)	80.2 (2.39)	77.7 (3.39)	65.6 (2.89)	61.4 (1.43)
Applied problems	88.4 (0.85)	87.6 (2.23)	88.2 (1.06)	88.3 (0.77)	87.9 (0.98)	79.8 (1.44)	83.9 (1.32)	80.6 (2.23)	71.2 (2.36)	72.8 (3.45)	62.9 (2.42)	63.4 (1.31)
Social studies	87.7 (0.99)	88.4 (2.28)	87.8 (1.23)	86.6 (0.90)	85.6 (1.01)	84.3 (1.27)	80.5 (1.57)	79.1 (2.47)	73.9 (2.42)	73.8 (3.03)	67.5 (1.95)	65.1 (0.98)
Science	90.0 (0.94)	88.8 (2.05)	89.3 (1.25)	87.6 (0.91)	85.6 (1.02)	83.4 (1.28)	75.4 (1.77)	80.0 (2.74)	75.7 (2.21)	68.4 (3.65)	69.3 (2.04)	67.0 (1.15)

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), direct assessments, 2002 and 2004.

Learning disability. Youth with learning disabilities have average scores between 82 and 90 across assessment subtests. They score lowest on the passage comprehension subtest (82) compared with the synonyms/antonyms (90, $p < .001$); mathematics calculation (86, $p < .01$), applied problems (88, $p < .001$), social studies (87, $p < .01$) and science subtests (88, $p < .001$). The ability of youth in this category to use vocabulary skills in the synonyms/antonyms subtest is stronger than their mathematics calculation (90 vs. 86, $p < .05$) or social studies skills (87, $p < .05$).

Speech/language impairment. Youth with speech/language impairments receive their highest average assessment scores on the mathematics calculation subtest (92) and their lowest on the passage comprehension subtest (81, $p < .001$). In comparison, their average score is 90 on the synonyms/antonyms subtest ($p < .001$ compared with passage comprehension), 88 on the applied problems subtest ($p < .05$ and $p < .001$), 86 on the science subtest ($p < .001$ and $p < .01$), and the 86 on the social studies subtest ($p < .001$ and $p < .01$). The performance of youth with speech/language impairments on the synonyms/antonyms subtest is stronger than on the social studies and sciences subtests ($p < .001$ and $p < .01$).

Orthopedic impairment. The achievement of youth with orthopedic impairments ranges from average scores of 79 to 88 across subtests. Those with orthopedic impairments are better able to use vocabulary skills in the synonyms/antonyms subtest (88) than they are skills assessed on all other subtests. Compared with performance on the synonyms/antonyms subtest, their average score is 79 on passage comprehension ($p < .001$), 80 on applied problems ($p < .001$), 83 on science ($p < .01$), 83 on mathematics calculation ($p < .01$), and 84 on social studies ($p < .05$) subtests. They also have stronger social studies than passage comprehension (84 vs. 79, $p < .01$) or applied problem solving skills (84 vs. 80, $p < .05$), and stronger science than passage comprehension skills (83 vs. 79, $p < .05$).

Hearing impairment. Youth with hearing impairments receive average assessment scores that range from 75 to 92. They exhibit stronger mathematics calculation skills (92) than science knowledge (75, $p < .001$), passage comprehension (76, $p < .001$), social studies knowledge (81, $p < .001$), and applied problems solving (84, $p < .001$) and synonyms/antonyms skills (84, $p < .001$). Their applied problem solving skills are stronger than their science knowledge ($p < .001$) or passage comprehension ($p < .001$) and their social studies knowledge is better than their science knowledge ($p < .05$) or passage comprehension ($p < .05$). Similar to those in almost all other disability categories, the use of synonyms/antonyms by youth with hearing impairments is stronger than their passage comprehension skills ($p < .001$).

Traumatic brain injury. Youth with traumatic brain injuries receive average scores between 74 and 84 on assessment subtests, with no significant difference in performance across subtests other than stronger performance on the synonyms/antonyms than the passage comprehension subtest (84 vs. 76, $p < .01$).

Autism. Average assessment scores for youth with autism range from 70 to 81. Youth with autism have better synonyms/antonyms (81) and mathematics calculation skills (80) and science knowledge (76) than passage comprehension (70, $p < .001$, $p < .01$, and $p < .001$). Their use of synonyms/antonyms also is better than their applied problem solving skills (71) and social studies knowledge (74, $p < .01$ and $p < .05$), and their mathematics calculation skills are stronger than their applied problem solving abilities (80 vs. 71, $p < .01$).

Deaf-blindness. Those with deaf-blindness receive average assessments scores that range from 66 to 78. These youth are better able to perform mathematics calculations than understand reading passages (78 vs. 66, $p < .05$). There are no other significant differences in performance across tests.

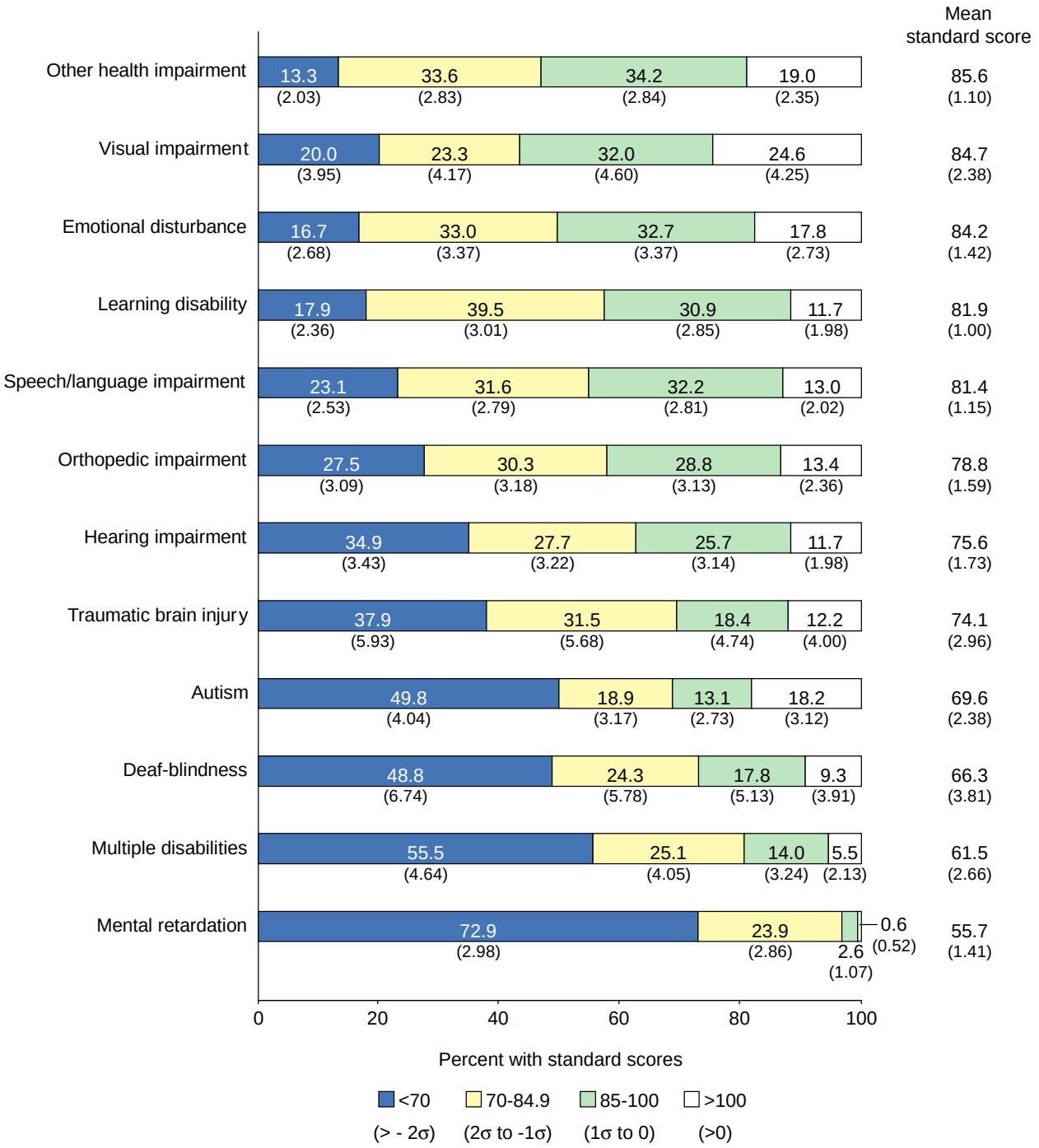
Multiple disabilities. Average scores for youth with multiple disabilities range from 62 to 72. They experience few significant differences in performance across subtests, other than a higher score on the synonyms/antonyms subtest (72) than on the passage comprehension (62) or applied problems (63) subtests ($p < .01$ for both comparisons), and a lower score on the passage comprehension subtest than on the science subtest ($p < .05$).

Mental retardation. Average scores for this group range from 56 to 67. Youth with mental retardation receive lower scores in passage comprehension (56) than in all other subtests. They receive average assessment scores of 67 in science ($p < .001$), 65 in social studies ($p < .001$), 65 in synonyms/antonyms ($p < .001$), 63 in applied problems ($p < .001$), and 61 in mathematics calculations ($p < .01$). They receive higher scores on the science than the applied problems ($p < .05$), and mathematics calculation ($p < .01$) subtests ($p < .01$) and lower scores on the mathematics calculation than the social studies ($p < .05$) or synonyms/antonyms ($p < .05$) subtests.

Although within each disability category, most youth score below the population mean, there is a distribution of scores such that some youth score above the mean. To illustrate, figure 2 depicts achievement on the passage comprehension subtest for youth in each disability category. In every disability category, some youth score across the distribution, including youth who have strong reading skills. For example, one-quarter of those with visual impairments perform above the mean on reading comprehension. Eighteen or 19 percent of those with other health impairments, emotional disturbances, or autism score above 100. One percent of those with mental retardation; 6 percent of those with multiple disabilities; 9 percent of those with deaf-blindness; and from 12 to 14 percent of those with speech, orthopedic, or hearing impairments; learning disabilities; or traumatic brain injuries score above the mean on the passage comprehension subtest. This pattern of variation within disability categories—with some youth in every category performing at every level—is apparent on all of the direct assessment subtests.[2]

[2] Tables presenting within-disability category performance on the five additional subtests are included in appendix C.

Figure 2. Performance on the Woodcock-Johnson III passage comprehension subtest, by disability category



NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), student assessments, 2002 and 2004.

Summary

This chapter reveals that a considerable gap exists between the academic achievement of youth with disabilities and their peers in the general population in reading, mathematics, science, and social studies. Generally, the majority of those with disabilities are scoring below the mean on each subtest, although some are performing well. Chapter 4 examines factors that are independently related to these variations in assessment scores, including disability and functioning, individual and household demographics, past school experiences, and the use of accommodations.

4. What Factors Are Associated With the Academic Achievement of Youth With Disabilities?

Analyses of the original NLTS (Wagner, Blackorby, and Hebbeler 1993), earlier analyses of NLTS2 (Blackorby et al. 2003), and analyses of the Special Education Elementary Longitudinal Study (SEELS) (Blackorby et al. 2004) suggest that the academic achievement of students with disabilities is the result of a complex interplay of many factors. These studies show that some factors are intrinsic to youth themselves (e.g., disability and functioning); some are characteristics of their family environment (e.g., family support for education); and some involve experiences in and outside of school (e.g., having been retained at grade level). We know that many of the factors are interrelated, which complicates the task of understanding how they are associated with the academic achievement of youth with disabilities. For example, many disabilities are defined by specific functional limitations, such as auditory and sight limitations among those with visual or hearing impairments, social skills limitations among youth with emotional disturbances or autism, physical limitations among those with orthopedic impairments, and cognitive challenges among those with mental retardation (Wagner et al. 2003). Yet youth within these disability categories differ in the extent of limitation, with the severity of limitation potentially helping to explain variation in academic achievement beyond the disability category label alone. Interrelationships also are apparent regarding demographic factors; for example, African American and Hispanic youth with disabilities are more likely than White youth to be living in households with incomes below the federal poverty threshold, as are youth with mental retardation relative to youth in most other disability categories (Marder et al. 2003).

Multivariate statistical analyses (i.e., linear regression when dependent measures are continuous, as in the case of standard scores on academic assessments) are required to disentangle these kinds of interrelationships and identify how various factors relate to academic achievement, independent of other factors. Such analyses estimate the magnitude and direction of relationships for numerous explanatory factors, statistically holding constant the other factors included simultaneously in the analyses.[1] The multivariate analysis results reported in this chapter attempt to explain variation in the standard scores of youth with disabilities on three subtests related to reading (passage comprehension), mathematics (calculations), and content knowledge (in science).[2] The factors included as independent variables in these multivariate analyses are drawn from the NLTS2 conceptual framework and are described in the following section, followed by a presentation of the analysis results.

[1] Multivariate analyses identify relationships between a variety of factors and student outcomes, but findings do not imply that the factors cause the outcomes. For example, higher absenteeism is negatively associated with some measures of academic performance, independent of other differences between youth. However, this does not imply that absenteeism causes poorer academic performance; it could be that poor-performing students are less engaged in their schooling and, therefore, absent more often.

[2] One measure was chosen from each of the three academic domains (language arts, mathematics, and content knowledge), for parsimony in presenting the results. Multivariate analysis results for the other three subtests (synonyms/antonyms, applied problem solving, and social studies content knowledge) are quite similar to those presented here and are included in appendix C. Footnotes in this chapter summarize those relationships.

Hypothesized Relationships to the Academic Achievement of Youth With Disabilities

Previous research on students with disabilities (Blackorby et al. 2003; Blackorby et al. 2004; Wagner, Blackorby, and Hebbeler 1993) and other research, noted below, suggest three broad categories of factors that may help explain variations in the academic achievement of youth with disabilities:

- individual characteristics of youth;
- characteristics of their households; and
- school experiences.[3]

Factors within each of these categories that are included in NLTS2 analyses are described below. In addition, the accommodations provided youth with disabilities as part of their individualized education programs (IEPs) that also were provided during the direct assessment may relate to the assessment scores.

Individual Characteristics of Youth

As noted above, the academic achievement of youth with disabilities results from dynamic processes in which youth are active participants. For example, the learning that promotes academic achievement occurs as teachers and students interact with each other and with instructional content and activities. What youth bring to these processes may be related to their success. Three major types of individual characteristics are hypothesized to relate to the academic performance of youth: disability characteristics, functioning, and demographics.

Disability Characteristics

In considering variations in the academic performance of youth, it is important to understand the relationships between performance and of the following aspects of their disabilities.

- **Disability category.** The nature of a particular youth's disability can powerfully condition his or her experiences, as demonstrated in all reports of NLTS2 findings. Dichotomous variables are included in analyses that distinguish youth according to the federally defined special education disability categories in use for secondary-school-age students (see appendix A, table A-10, for definitions of these categories).[4]

  The assignment of youth to a disability category is based on the primary disability designated by the youth's school or district in the 2000-01 school year. Although there are federal definitions for each disability category, criteria and methods for determining eligibility under particular categories vary widely from state to state. Therefore, NLTS2 category designations should be interpreted as describing those reported to have a particular disability, rather than those who have that disability.

---

[3] Multivariate analyses do not include factors related to youth's complete school programs because data available from the Students' School Program Survey regarding course-taking and other key aspects of those programs are for only a single school year. Students' transcripts, which will depict their complete high school career up to the point of their assessment, are expected to yield data that will be more strongly related to overall academic achievement. Transcripts are being collected as youth complete high school and, thus, are not yet available.

[4] For analysis purposes, the deaf-blind category was combined with the multiple disability category because it is too small to be analyzed separately.

Overall, 62 percent of youth receiving special education in the NLTS2 age range are classified as having a learning disability (Marder, Levine, and Wagner 2003). Youth with mental retardation and emotional disturbances make up 12 percent and 11 percent of students, respectively. Another 5 percent of youth are classified as having other health impairments, and 4 percent are identified as having speech/language impairments. The seven remaining disability categories each account for 1 percent or fewer of students and, together, make up about 5 percent of youth with disabilities. The nature of a youth's disability is hypothesized to account for much of the variation in academic achievement, with youth in such categories as visual, hearing, or speech impairment generally experiencing more positive outcomes than, for example, youth in categories such as multiple disabilities or mental retardation (Blackorby et al. 2003).

- **Age at identification of disability.** Early identification of a disability can indicate that the disability affects functioning in obvious, recognizable ways at a young age, compared with disabilities that are recognized in later childhood or adolescence. Thus, on average, youth whose disabilities were identified at an earlier age could have greater challenges to academic performance. Parents reported the age at which youth first exhibited a physical, learning, or other disability or problem for which they eventually were diagnosed. Although the average age is 5.7 years, approximately one in five youth have disabilities that first were diagnosed when they were infants or toddlers, and another 11 percent have disabilities or delays that were identified in their preschool years. School entry, at age 5 or 6, was when almost one-third of youth first had their disabilities identified, whereas 19 percent did not have their disabilities identified until they were at least 9 years old. Age at identification is confounded with type of disability in that sensory and orthopedic impairments tend to be recognized earlier than learning disabilities, for example, which generally are identified when children reach school age (Wagner, Marder, and Cardoso 2003).

Functioning

NLTS2 findings demonstrate the considerable variation in skills across several dimensions among youth who share a primary disability category designation (Cameto et al. 2003; Wagner et al. 2003). To assess the relationship between functioning and academic performance, NLTS2 analyses include variables that distinguish the level of functioning of youth with disabilities in the areas noted below.

- **Number of domains influenced by disability.** The number of functional domains affected by disability indicates the breadth of the potential impact of disability on the outcomes youth may achieve. To assess the breadth of the functional impacts of youth's disabilities, parents were asked to report whether youth experience limitations in seven areas: general health; vision; hearing; use of arms, hands, legs, and feet; expressive language; receptive language; and participation in bidirectional communication. Parents of youth with disabilities report that about half have problems in at least one area, whereas about 1 in 10 have problems in four or more of these areas (Wagner et al. 2003). Having a disability that impacts fewer areas is expected to relate to higher academic performance (Blackorby et al. 2003).

- **Functional cognitive skills.** Parents were asked to use a 4-point scale ranging from "not at all well" to "very well" to evaluate four of their sons' or daughters' skills that often are used in daily activities: reading and understanding common signs, telling time on a clock with hands, counting change, and looking up telephone numbers and using the telephone. These skills are referred to as "functional cognitive skills" because they require the cognitive ability to read, count, and calculate. As such, they suggest much about students' abilities to perform a variety of more complex cognitive tasks. However, they also require sensory and motor skills—for example, to see signs, manipulate a telephone, and so on. Consequently, a high score indicates high functioning in all of these areas, but a low score can result from a deficit in the cognitive, sensory, and/or motor domains. A summative scale of parents' ratings of these functional cognitive skills ranges from 4 (all skills done "not at all well") to 16 (all skills done "very well"). Approximately half of youth with disabilities score in the high range on this scale (15 or 16); almost 6 percent score in the low range (4 to 8). As an indicator of the ability to process information that is important to daily functioning, higher functional cognitive skills are expected to relate strongly to academic achievement (Blackorby et al. 2003; Blackorby et al. 2004).
- **Social skills.** Poor academic performance has been shown to be an "indirect consequence" of poor social skills at school (National Association of School Psychologists 2002); hence, higher social skills are expected to relate to higher academic performance, as measured by standardized assessments. The social skills of youth with disabilities were assessed by asking parents to respond to nine items drawn from the Social Skills Rating System (SSRS), Parent Form (Gresham and Elliott 1990). Items were selected from the assertion and self-control subscales, skill sets considered by the design team to be most relevant to school success. Individual items were selected because they had high factor loadings on the relevant subscale and/or did not duplicate particular skills (e.g., controls temper with children and controls temper with the parent were not both selected).

  Parents were asked whether their adolescent children exhibit each of the following characteristics "never," "sometimes," or "always" (scoring 0, 1, or 2, respectively):

  - makes friends easily;
  - seems confident in social situations, such as parties or group outings;
  - joins group activities without being told to, such as a group having lunch together;
  - starts conversations rather than waiting for others to start;
  - ends disagreements with [parent] calmly;
  - controls temper when arguing with peers other than siblings;
  - speaks in an appropriate tone of voice at home;
  - receives criticism well; and
  - gets into situations that are likely to result in trouble (reverse coded).[5]

[5] The original SSRS item, worded "avoids situations that are likely to result in trouble," was altered in the initial days of interviewing parents because a significant number of them had difficulty understanding the item when not seeing it written in front of them, as is the usual mode of administration for the SSRS.

The first four items come from the assertion subscale, and the last five come from the self-control subscale. Intercorrelations of the assertion subscale items range from .35 ($p < .001$) to .55 ($p < .001$). Intercorrelations of the self-control subscale items range from .13 ($p < .001$) to .33 ($p < .001$).

An overall measure of social skills was created by summing the values across the nine items, producing a scale with raw scores that range from 0 to 18 and have a mean and standard deviation of 11.4 and 3.15.

To standardize the scale for the NLTS2 sample, a standardization sample was requested from American Guidance Services, Inc. (AGS), which distributes the SSRS. AGS graciously provided NLTS2 with the entire 174-case sample that the developers used to standardize the parent form of the SSRS for secondary-level students (7th to 12th graders in spring of 1988). The nine items (with values of 0, 1, and 2) were summed, and the mean and standard deviation were calculated. The mean of the 9-item scale for the 174-case SSRS standardization sample is 12.3, and the standard deviation is 2.79 (the range is 6 to 18). This mean and standard deviation were used to create standardized scores for youth with disabilities as follows:

$$z_i = \frac{x_i - \mu}{\sigma}$$

Where:

- z_i = the z-score of the nine-item social skills scale for student *i* in the NLTS2 sample
- x_i = the sum of the nine items comprising the NLTS2 social skills scale for student *i* in the NLTS2 sample
- μ = the mean of these same nine items in the SSRS standardization sample
- σ = the standard deviation of the nine items in the SSRS standardization sample

Thus, the standardized (z-score) variable expresses each NLTS2 student's score in terms of its distance from general population's mean score, where the metric for the distance is the general population standard deviation. This variable has a range of -4.42 to 2.03, with a weighted mean of -0.34 and a standard deviation of 1.13.

- **Persistence.** Parents were asked how often youth kept "working at something until it is finished, even if it takes a long time." Response categories were "never," "sometimes," or "very often." The ability to persist with tasks to completion is expected to be positively associated with higher academic performance. Parents report that 16 percent of youth "never" persist, and 35 percent do so "very often."

Demographic Characteristics

The factors noted above suggest relationships between the nature of a youth's disability and his or her experiences. However, demographic characteristics also are associated with variations in academic achievement, both for youth with disabilities (Blackorby et al. 2003; Blackorby et al. 2004; Wagner, Blackorby, and Hebbeler 1993), and youth in the general population (Freeman 2004; Kao and Thompson 2003).

- **Age.** The large majority of youth in NLTS2 were ages 16 through 18 when the direct assessments and functional ratings were conducted. Because this is a fairly narrow age range and the standard scores used in the analyses take age into account, the differences in academic achievement for youth who are at the lower and upper ends of the range were expected to be relatively small, though potentially still significant. Differences could result if youth fall increasingly behind or increasingly advance as they age.
- **Gender.** In the general population, differences in the achievement of boys and girls in school are notable, generally favoring girls (National Center for Education Statistics, 2005d). Differences also have been noted for youth with disabilities, although the strength and direction of relationship are less consistent (Blackorby et al. 2003; Blackorby et al. 2004; Wagner, Blackorby, and Hebbeler 1993). Whereas youth in the general population are split about evenly between boys and girls, almost two-thirds of youth with disabilities in the NLTS2 age range are boys. Further, it also is clear that gender is intertwined with the nature of youth's disabilities, with males accounting for a much higher proportion of some disability categories (e.g., autism, emotional disturbances) than others (e.g., hearing or visual impairments) (Marder, Levine, and Wagner 2003). Including both gender and disability in multivariate analyses will enable their independent relationships to academic performance to be identified.
- **Racial/ethnic background.** Research has documented the relative disadvantage minority youth experience in the education domain (National Center for Education Statistics 2005d), as has prior research on youth with disabilities (Blackorby et al. 2003; Blackorby et al. 2004; Wagner, Blackorby, and Hebbeler 1993). Overall, 62 percent of youth with disabilities are White, 21 percent are African American, 14 percent are Hispanic, and 3 percent have other or multiple racial/ethnic backgrounds. However, this distribution varies across disability categories. For example, the category of mental retardation has a particularly large percentage of African Americans (33 percent), and the categories of other health impairment and autism have particularly small percentages of Hispanic students (8 percent and 9 percent, respectively) (Marder, Levine, and Wagner 2003). Again, multivariate analyses permit the relationships of these factors to academic performance for youth with disabilities to be assessed independently.

Household Characteristics

Although the variables described above were expected to do much to help illuminate differences in the academic performance of youth with disabilities, focusing on these variables alone would mistakenly imply that learning is related only to somewhat immutable characteristics that young people bring with them to school and would ignore the important role of household and family context in shaping the experiences of youth (Blackorby et al. 2003;

Henderson and Berla 1994). The following characteristics of the households of youth with disabilities were expected to relate to their academic performance in the ways noted below.

- **Household income.** Poverty has been shown to have serious negative consequences for children and youth as a whole (Duncan and Brooks-Gunn 1997) and for the achievements of youth with disabilities in secondary school (Blackorby et al. 2003; Blackorby et al. 2004; Wagner, Blackorby, and Hebbeler 1993). A similar pattern is predicted for the current NLTS2 analyses. One-fourth of youth with disabilities live in poverty, a higher rate than in the general population (20 percent, $p < .01$) (Marder et al. 2003). A reasonable proxy for poverty is annual household income, for which NLTS2 obtained categorical data in $5,000 increments, with a top category of $75,000 or more. Because the variable is not strictly continuous, it was included in the analyses as two dummy variables: low income (less than $25,000; 32.2 percent of the sample) and high income ($75,000 or more; 12.6 percent of the sample). The moderate income category is the omitted variable. Because low household income often is related to minority racial/ethnic status (Marder et al. 2003), including both household income and the racial/ethnic background of youth with disabilities in analyses will help disentangle their interrelationships.
- **Family support for education.** Parental support for learning is an important contributor to success in school for the general student population (Epstein 1996; Henderson and Berla 1994; Thorkildsen and Stein 1998) and for youth with disabilities (Blackorby et al. 2003). A similar association is expected for the current analysis. Two scales have been constructed to test this expectation. One scale, which assesses family involvement in education at home, is the frequency (on a 4-point scale) with which parents report helping youth with homework and talking with youth, and, a dichotomous variable indicating whether the family provides a computer at home that the student uses for educational purposes; summing responses to these items produces a scale ranging from 0 to 9, with a mean of 6.8. Family involvement at school is assessed with a second scale constructed by summing parents' reports (on a 4-point scale) of the frequency with which they did the following in the 2001-02 school year: "attend a general school meeting, for example, back-to-school night or the meeting of a parent-teacher organization"; "attend a school or class event, such as a play, sports event, or science fair"; or "volunteer at school, for example, chaperoning a class field trip or serving on a committee." The scale ranges from 0 to 12, with a mean of 3.3.
- **Family expectations.** Research has demonstrated that having clear, consistent, and high expectations for academic performance is related to student achievement for the general population (Thorkildsen and Stein 1998). Similar relationships have been found for students with disabilities (Blackorby et al. 2003) and were predicted to emerge in the current NLTS2 analyses. Parents were asked to report their expectations that their adolescent children with disabilities will "attend school after high school." Expectations for youth are generally high. Overall, 62 percent of parents expect youth "definitely" or "probably" to attend postsecondary school.

School Experiences

The analyses include several factors related to students' school experiences that have been shown to relate to academic achievement, as noted below. Because academic achievement is a product of students' cumulative experiences with schooling over time, the school-related factors in the analyses include some measures of past experiences (e.g., cumulative school mobility and grade retention over the school career). They also include measures of current performance at school that research suggests correlate with learning. Instructional experiences (e.g., instructional setting, curriculum modifications) are not included here because NLTS2 currently has only single-point-in-time measures of such factors, rather than data on students' overall school programs, which will come from transcripts collected when students have completed school; subsequent analyses can address relationships between school program factors and academic achievement.

- **Student mobility.** Research has demonstrated relationships between high rates of student mobility and poor school performance (Demie 2002; Rumberger 2002). These negative relationships may relate, at least in part, to the disruption and lack of continuity in students' learning experiences, which, for students with disabilities, may include compromised service coordination, the potential for poor communication between new and old schools and service systems, and inadequate record sharing (Kerbow 1996). For these reasons, parents' reports of the number of times students with disabilities have changed schools, other than because they were moving from one grade level to the next, are included in analyses. Values range from 0 to 8, with a mean of 1. Because this variable is skewed, to ensure that its inclusion as a continuous variable did not mask nonlinear effects, an alternate set of models also was estimated substituting a series of dummy variables for values 1 to 8 (with zero as the omitted category).
- **Grade retention.** The intention in making low-performing students repeat a grade is to provide an opportunity for them to master material missed in their first exposure to it at a given grade level, which may result in improved performance later in school. However, some research indicates that younger students with disabilities who were retained at grade level continue to exhibit lower academic performance than those who were not, independent of other differences between them (Blackorby et al. 2004). The current NLTS2 analyses include a measure of parents' reports of whether youth have ever been retained at grade level (36 percent had done so in Wave 1) (Wagner 2003) to explore this relationship for secondary school youth.
- **Grades.** Although performance on standardized tests receives the greatest attention in discussions of students' academic performance, teachers' evaluations of performance as indicated in course grades represent a common metric of student performance that often is tied more directly to the day-to-day business of teaching and learning than are annual standardized test scores. Grades communicate to students and parents information about students' mastery of course content. In high school, a passing grade also is the criterion for a course's contributing to accumulated credit for graduation, and grades provide information for consideration in college admissions. Parents were asked to report students' overall grades on a 9-point scale (mostly As, mostly As and Bs, mostly Bs, and so on). For youth with no parent interview, teachers were asked to report students' grades in their classes on the same 9-point scale. For students who, according to parents,

received such grades as "excellent," "good," "fair," and "poor" instead of letter grades, grades in this form were converted to correspond to the same scale as letter grades.[6] On this composite measure, 32 percent of students with disabilities receive mostly As and Bs, and 10 percent receive mostly Ds and Fs.

- **Absenteeism.** Absenteeism results in students' missing exposure to curriculum and instruction and can interfere with relationships, resulting in reduced learning and performance (Blackorby et al. 2003). Therefore, the number of days students are absent in a month, excluding suspensions and expulsions (which are accounted for in the behavior variable described below), is included in the analyses. Data for a student's absences, obtained from his or her school through the Student's School Program Survey, indicate students with disabilities miss an average of 2.6 days of schools in a 4-week period (standard error = 0.10); 14 percent miss 6 or more days (Newman, Davies, and Marder 2003).
- **Behavior at school.** The behavior of youth at school is a crucial element in their overall social adjustment. Not only is school the context in which many youth spend most of their day, it also is where they engage in the important activities of gaining academic knowledge; learning and practicing more generalized skills, such as problem solving, being on time, and following directions; and developing formative relationships with peers and adults. Research on poor behavior at school has linked it to poor motivation for learning, which in turn, has been shown to relate to poor academic performance (Anderman and Maehr 1994; Deci et al. 1992; Wiest, Wong, and Kreil 1998). Students whose behavior at school violates school norms typically are subject to disciplinary actions or, in some cases, to suspension or expulsion from school. School staff were asked whether youth had been suspended, expelled, or involved in any other type of disciplinary action, such as a referral to the office or detention, during the current school year, and a dichotomous variable was created indicating whether or not any of these had occurred. More than one-third of youth with disabilities (35 percent) are involved in one or more of these types of disciplinary action in a school year.

Accommodations During Testing

Youth who were reported to need them were given various accommodations during testing. The most frequent accommodation during group-administered testing—more time to complete the test—was not an issue during the NLTS2 assessment because the test was individually administered and not timed. However, as reported in chapter 2, some youth did take the assessment in multiple sessions or with breaks (8 percent); using American Sign Language (ASL) or with an ASL interpreter (8 percent); using Braille or large print materials (6 percent); with special furniture or lighting (5 percent); and/or using a calculator for the mathematics assessments (23 percent). To the extent that accommodations were given to all youth who needed them and "equalized the playing field" for those youth, no relationship with performance would be expected relative to youth who did not receive them because they did not need them. On the other hand, if the accommodations did not fully compensate for the impairment that necessitated the accommodation, a negative association could be found. Conversely, if the accommodations provided an extra advantage, one would expect to see positive associations.

[6] See appendix A for a description of the meshing of these grade measures.

The relationships these characteristics of individual youth with disabilities, their households, and their previous school experiences have with academic performance are presented in the following section.

Factors Related to the Academic Achievement of Youth With Disabilities

Multivariate analyses of the standard scores of youth with disabilities on reading comprehension, mathematics calculation, and content knowledge in science have identified the relationships described below between achievement and specific individual and household characteristics and previous school experiences, independent of other factors in the analyses (table 2).

Individual Characteristics of Youth

Disability Characteristics

Disability category. Multivariate analyses of categorical variables, such as disability category, require that individual categories be compared with a standard category that is not included as a variable in the analysis. For analyses related to disability category, the standard for comparison is the category of learning disability; it was chosen because it is the largest category and, thus, most closely resembles a comparison with youth with disabilities as a whole. Controlling for other factors, there is a significant amount of variation in the three measures of academic performance related to disability category differences. For example, youth with visual impairments or emotional disturbances score 7 and 5 points higher ($p < .05$) than youth with learning disabilities on passage comprehension, other factors held constant,[7] whereas youth with traumatic brain injuries, autism, multiple disabilities, or mental retardation score from 6 to 13 points lower.[8]

There is somewhat less variability in scores across disability categories with regard to mathematics calculation and science content knowledge, with scores of youth in four categories being significantly different from those of youth with learning disabilities. As with passage comprehension, youth with visual impairments averaged scores about 7 points higher than peers with learning disabilities on mathematics calculation ($p < .01$), although they did not differ in science content knowledge. Youth with hearing impairments also scored significantly higher than youth with learning disabilities on mathematics calculation (5 points, $p < .05$), yet averaged significantly lower scores on science content knowledge (8 points, $p < .01$).[9] Youth with multiple disabilities (including deaf-blindness) or mental retardation score lower than youth with learning disabilities on tests of both mathematics calculation and science content knowledge (ranging from 9 to 11 points, $p < .01$ and $p < .001$); youth with autism also score lower than those with learning disabilities on the measure of science content knowledge (10 points, $p < .001$).

---

[7] A positive relationship also is noted for the use of synonyms/antonyms (appendix C, table C-2).

[8] Youth with autism, multiple disabilities, and mental retardation also score significantly lower than youth with learning disabilities on the three measures reported in appendix C; those with traumatic brain injuries show negative relationships to measures of applied problem solving and social studies content knowledge (appendix C, table C-2).

[9] Youth with hearing impairments also score significantly lower on the measure of social studies content knowledge (appendix C, table C-2)

Table 2. Factors associated with variation in the passage comprehension and mathematics calculation skills and science content knowledge of youth with disabilities

Independent variables	Change in the following scores for each unit change in the independent variable: Passage comprehension	Mathematics calculation	Science content knowledge
Disability characteristics			
Visual impairment (vs. learning disability)	**+6.55***	**+7.06****	+4.00
Emotional disturbance (vs. learning disability)	**+4.50***	+2.14	+0.41
Orthopedic impairment (vs. learning disability)	+3.26	-1.11	-0.01
Other health impairment (vs. learning disability)	+2.32	+1.99	+0.16
Speech/language impairment (vs. learning disability)	-1.27	+2.85	-2.23
Hearing impairment (vs. learning disability)	-3.79	**+4.85***	**-8.19****
Traumatic brain injury (vs. learning disability)	**-5.72***	-0.93	-3.88
Autism (vs. learning disability)	**-7.32*****	-2.21	**-9.77*****
Multiple disabilities/deaf-blindness (vs. learning disability)	**-8.00*****	**-9.29*****	**-9.32****
Mental retardation (vs. learning disability)	**-13.44****	**-11.15****	**-9.74*****
Age at identification of disability	**+0.42****	**+0.41*****	**+0.25***
Functioning			
Number of domains affected	**-0.86***	+0.38	**-0.83****
Functional cognitive skills	**+2.09*****	**+2.60*****	**+1.31*****
Social skills	**-1.16***	-0.79	-0.57
Persistence	-1.20	+0.74	**-1.47***
Demographics			
Age	-0.52	-0.77	-0.75
Gender (boys vs. girls)	+0.05	**+3.23*****	**+2.81*****
African American (vs. White)	**-8.46*****	**-8.33*****	**-10.63*****
Hispanic (vs. White)	**-11.80*****	**-5.21*****	**-12.76*****
Other or multiple race/ethnicity (vs. White)	**-10.48*****	**-8.77*****	**-10.23*****
Household characteristics			
Expectations for postsecondary education	**+6.18*****	**+6.32*****	**+4.64*****
Low income (vs. moderate income)	**-3.36*****	**-2.74****	**-4.74*****
High income (vs. moderate income)	+1.19	+2.03	-0.03
Family involvement at home scale score	-0.35	-0.51	-0.29
Family involvement at school scale score	+0.22	+0.21	+0.11
School experiences			
Ever retained at grade level	-1.33	-1.41	-0.55
Overall grades	-0.21	+0.12	+0.09
Had any suspensions, expulsions, or disciplinary actions in the current school year	+0.96	**-2.09***	-0.17
Days absent per month	-0.31	**-0.52****	-0.16
School mobility other than for grade level changes	+0.59	-0.12	+0.14

See notes at end of table.

Table 2. Factors associated with variation in the passage comprehension and mathematics calculation skills and science content knowledge of youth with disabilities—Continued

Independent variables	Change in following scores for each unit change in the independent variable: Passage comprehension	Mathematics calculation	Science content knowledge
Accommodations			
Breaks or multiple sessions	**-4.00****	-2.80	**-5.74*****
ASL or ASL interpreter	**-8.96*****	**-4.96****	**-8.25*****
Braille or large print materials	+4.50	+3.74	+3.50
Special furniture or lighting	-2.05	+1.56	+0.95
Calculator	†	**+3.94*****	†

† Not applicable; this accommodation was included only in models related to mathematics.

Multivariate analyses require that for categorical variables, such as disability category, each category be compared with another specified category. Learning disability was chosen as the category against which to compare the relationships for other disability categories because it is the largest category and, therefore, most closely resembles the characteristics of youth with disabilities as a whole. Similarly, White youth are the group against which to compare results for other racial/ethnic groups because it is the largest group.

*p < .05, **p < .01, ***p < .001. Note that approximately 100 relationships are included in this table; about 5 would be expected to be statistically significant by chance.

Table reads: The passage comprehension standard score of youth with visual impairments is 6.6 points higher than the score of youth with learning disabilities, other factors being equal. The mathematics calculation score of youth from low-income households is 2.7 points lower than the scores of youth from moderate income households, independent of other factors.

The performance across measures by youth in some categories is quite stable. For example youth with orthopedic, speech/language, or other health impairments consistently have scores that are similar to youth with learning disabilities,[10] and those with mental retardation have the lowest scores on all measures of achievement. In contrast, the performance of youth in some other categories varies markedly across the measures, as noted above for youth with hearing impairments, who score significantly higher and lower than youth with learning disabilities on two measures but are not significantly different on the third. Youth with emotional disturbances or traumatic brain injuries have scores that are not significantly different than youth with learning disabilities on mathematics calculation and science content knowledge,[11] but those with emotional disturbances score significantly higher (4 points, $p < .05$) and those with traumatic brain injuries score lower by 6 points ($p < .05$) on passage comprehension.[12]

Age at identification of disability. Having a disability that was identified at a later age is consistently associated with higher standard scores across the three measures, ranging from one-quarter to about one-half a point for each increasing year.[13] Thus, for example, a youth whose disability was first identified at age 9 is estimated to score about 2 points higher on mathematics

[10] Youth with orthopedic or speech/language impairments also score similarly to peers with learning disabilities on the three measures reported in appendix C; those with other health impairments outscore youth with learning disabilities on the use of synonyms/antonyms (appendix C, table C-2).

[11] Similarly, their scores do not differ from youth with learning disabilities on two of the three measures reported in appendix C.

[12] Youth with emotional disturbances also score higher than those with learning disabilities on the use of synonyms/antonyms, and youth with traumatic brain injuries score lower on social studies content knowledge.

[13] A similar relationship is noted for synonyms/antonyms and applied problem solving (appendix C, table C-2).

calculation and science content knowledge than a youth whose disability was identified at age 4, independent of other factors in the analyses.

Functioning

Number of functional domains affected by disability. The number of functional domains in which youth experience problems related to disability is significantly related to their achievement in reading and science content knowledge;[14] for each additional domain affected by disabilities there is estimated to be a 1-point decline in those two measures ($p < .05$ and $p < .01$, respectively).

Functional cognitive skills. A youth's functional cognitive skills are more consistently and strongly related to academic performance than the other functional abilities included in the analyses. Higher achievement across the measures[15] is apparent for youth with higher functional cognitive skills, ranging from 1 to almost 3 points for each 1-point gain in the scale score ($p < .001$ for all relationships). Thus, compared with youth whose functional cognitive skills scale score is 7, those with a score of 15 are estimated to have achievement scores that are from 11 to 21 points higher, independent of other factors in the analyses.

Other functional skills. In contrast to the consistent pattern of positive relationships between functional cognitive skills and academic achievement, scores measuring social skills and persistence in tasks each relates to only one of the measures of academic achievement in table 2, and both relationships are negative, contrary to expectations.[16] Specifically, there is approximately a 1-point decline in the passage comprehension standard scores associated with each additional 1-point score on the social skills scale and a similar decline in science content knowledge associated with each increasingly higher level of persistence ($p < .05$ for both relationships).

Demographic Characteristics

Age. None of the measures examined varies with the age of youth, all of whom were assessed when they were within a 16- through 18-year-old age range.[17]

Gender. Boys with disabilities outscore girls by about 3 standard score points on measures of both mathematics calculation ability and science content knowledge.[18] There is no significant difference in reading comprehension.

Race/ethnicity. Compared with White youth with disabilities, the academic performance of those in other racial/ethnic groups is markedly lower on all measures.[19] However, the patterns of performance across measures are not uniform for the three racial/ethnic groups. The performance

[14] This relationship also is apparent for synonyms/antonyms and social studies measures (appendix C, table C-2).
[15] Similar consistent and strong relationships are noted for other measures of academic performance (appendix C, table C-2).
[16] Social skills also have a similar relationship to the use of synonyms/antonyms (appendix C, table C-2); persistence is unrelated to measures reported in appendix C.
[17] A similar lack of relationship is apparent with measures reported in appendix C, table C-2.
[18] Similar relationships also are found in analyses of applied problem solving and social studies content knowledge (see appendix C, table C-2).
[19] This includes measures reported in appendix C, table C-2.

gap for African American youth with disabilities relative to White youth is fairly stable, ranging from 8 to 11 standard score points across measures ($p < .001$ for all relationships). In contrast, the gap between Hispanic and White youth with disabilities is about twice as large for reading and science (12 and 13 points, $p < .001$) than for mathematics (5 points, $p < .001$). The pattern for those in other or multiple racial/ethnic groups is similar to those of African American youth with disabilities ($p < .001$ for all relationships).

Household Characteristics

Parents' expectations for education achievements. The strongest relationships between household characteristics and measures of the academic achievement of youth with disabilities are found for parents' expectations of postsecondary education for their adolescent children with disabilities.[20] Youth with disabilities are estimated to score 5 or 6 points higher across the achievement measures ($p < .001$ for all relationships) with each successively higher level of parents' expectations, independent of other differences between them. Thus, for example, parents who report youth "definitely" are expected to attend postsecondary school are estimated to score 13 standard score points higher on mathematics calculation than youth whose parents think they "probably won't" enroll.

Household income. Having a low annual household income (i.e., $25,000 or less) is consistently and negatively related to academic achievement across measures. Independent of other differences between them, youth with disabilities in the low-income group have average scores that are 3 to 5 points lower across measures ($p < .01$ or $p < .001$ for all relationships) than those in the moderate income group (i.e., more than $25,000 to $75,000). In contrast, being in the high-income group (i.e., greater than $75,000) is not related to any of the three measures of academic achievement.[21]

Parents' support for education. Neither the level of parental support for education at home nor at school is associated with academic achievement in reading, mathematics, or science.[22]

School Experiences

Of the variety of school experiences examined in these analyses, only experience with disciplinary problems and absenteeism relate to academic achievement, and only with regard to mathematics calculation.[23] Those who were suspended, expelled, or subject to other disciplinary actions in the current school year average mathematics calculation scores that are 2 points lower than those who have not ($p < .05$). In addition, each day of absenteeism is associated with a half-

[20] Similarly strong and consistent relationships occur with the other measures of achievement (see appendix C, table C-2).

[21] Similar relationships are apparent for the measures reported in appendix C.

[22] A similar lack of relationship also is apparent regarding synonyms/antonyms, applied problems, and social studies content knowledge (table C-2, appendix C).

[23] Regarding measures reported in appendix C, the only relationship of school factors to achievement involves a negative association of absenteeism with the measure of applied problem solving.

point lower mathematics score ($p < .01$). There are no significant relationships between academic achievement and grade retention, grades, or school mobility.[24]

Accommodations Provided During Testing

Several of the accommodations normally provided youth with disabilities as part of their IEPs also were provided them when they participated in the direct assessment; in several cases, that provision is significantly related to variation in the resulting assessment scores. The strongest and most consistent relationships are apparent regarding the use of ASL or an ASL interpreter. Across measures, youth who received this accommodation scored from 5 to 9 points lower than youth who did not ($p < .01$ or $p < .001$ for all relationships), holding constant other differences between them, including the presence of a hearing impairment. Further, youth who took breaks during the assessment session or were assessed in multiple sessions averaged standard scores on reading comprehension and science content knowledge that were 4 and 6 points lower, respectively, than similar youth who completed the assessment without these accommodations ($p < .01$ and $p < .001$). In contrast to these negative relationships, using a calculator for the mathematics calculation subtest is associated with a 4-point higher score on that subtest, independent of other factors included in the models. No associations are apparent between academic achievement and the use of Braille or large print materials or accommodations related to furniture or lighting.

How Much Variation Is Explained by the Models?

The factors examined in multivariate analyses of WJ III direct assessment measures conducted in NLTS2 explain a statistically significant portion of their variation. Across the three measures of academic achievement, the $r2$ values (which measure the proportion of variation explained by the analysis) are .35 for passage comprehension and .38 for both mathematics calculation and .36 for science content knowledge.[25] Nonetheless, about two-thirds of the variation in these measures of academic achievement is attributable to differences between youth on other factors than those included in these analyses. Of the variation explained by the models, the large majority relates to the individual characteristics of youth with disabilities ($r2$ values ranging from .30 to .33). Household characteristics increase the amount of variation explained by .05 for each measure. School experiences and the provision of accommodations add no additional explained variation for any of the three measures beyond what is explained by individual and household characteristics.[26]

[24] In the alternate models, none of the dummy variables created from school mobility showed a significant relationship with any dependent variable.

[25] Values for synonyms/antonyms, applied problem solving, and social studies content knowledge are .46, .46, and .34, respectively.

[26] Added explained variation due to household characteristics is .07, .06, and .06 for synonyms/antonyms, applied problem solving, and social science content knowledge, respectively. Added explained variation associated with school experiences is .03 and.02 for measures of synonyms/antonyms and social studies content knowledge; no additional variation is explained in the measure of applied problem solving.

Summary

NLTS2 findings reinforce the fact that the academic achievement of youth with disabilities in reading, mathematics, science, and social studies is related to a complex array of factors that characterize youth, their households, and their school experiences. Individual factors are the most powerful in differentiating students on the basis of their academic achievement, with the most consistent relationships being apparent for racial/ethnic differences, favoring White youth. Variations in family characteristics also add to an understanding of patterns of academic achievement, with parents' expectations for the postsecondary education of their adolescent children with disabilities being consistently related to higher achievement and having a low household income consistently relating to lower achievement. The school experiences of youth with disabilities examined in these analyses show relatively few significant relationships with youth's academic achievement. Although using a calculator as an accommodation in testing is positively related to mathematics scores, when other accommodations relate significantly to academic achievement measures, the relationships are negative.

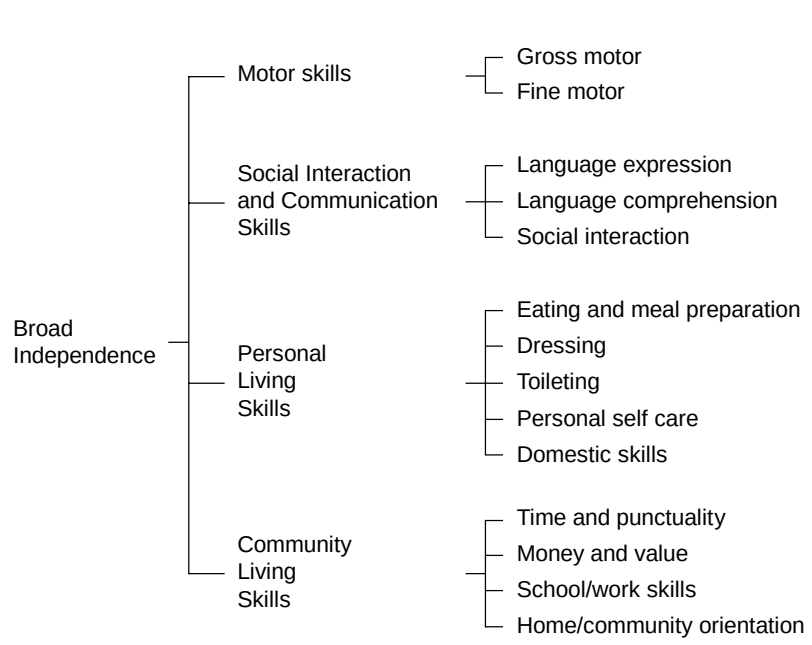
As noted earlier, the assessments analyzed in this chapter do not include all youth with disabilities; those for whom academic assessments were reported not to be feasible or appropriate were assessed using a checklist of abilities that was completed by each youth's teacher or another adult who could assess his or her performance in multiple domains. The results of these alternate assessments are described in the following chapter.

5. The Abilities of Youth Reported on the Functional Rating

This chapter presents findings from the NLTS2 functional rating, the *Scales of Independent Behavior-Revised* (SIB-R) (Bruininks et al. 1996), completed for 16- to 18-year-old youth with disabilities. The SIB-R is designed to assess skills needed to function independently in home, social, school, work, and community settings. The rating scale was completed by a youth's teacher or school staff member most knowledgeable about his or her functional abilities, or by parents, primarily when youth were no longer in school.[1] The SIB-R checklist addresses both academic and functional areas and mirrors the more functional nature of curriculum and instruction for youth with more severe disabilities who were identified by school staff or parents during screening to be appropriate for a functional rating.

As noted in chapter 1, the SIB-R covers a broad range of behaviors that are measured on the 14 subscales identified in figure 3. The subscales are organized into four clusters and an overall rating, broad independence. Each cluster is made up of two to five subscales, with each subscale having up to 20 items;[2] a total of 229 items are included in the overall broad independence rating. The analyses reported in this chapter are at the cluster level. Performance is reported as the percentage of youth whose ratings fall within a standard score range, which allows comparison with the norm sample of same-age youth in the general population. Focusing on findings at the cluster level minimizes the danger of

Figure 3. Content of measures on the SIB-R



[1] As reported in chapter 2, approximately 22 percent of youth with a functional rating are estimated to have had it completed by a parent. Only one statistically significant difference between those with ratings that were parent-completed and those with ratings completed by teachers was noted on the variety of demographic and disability-related factors and assessment scores examined. The group with ratings completed by parents had a significantly larger proportion of African American youth than the group with teacher-completed ratings (43 percent vs. 14 percent, $p < .05$).

[2] Each item is scored on a 4-point scale. On each item, the respondent rated the youth's ability to complete the task without help or supervision. A youth received a rating of 0 if he or she never or rarely performed the task, even if asked; 1 if he or she did the task completely, but not well; 2 if all parts of the task were done fairly well; and 3 if the task was almost always done very well and without prompting. The sum of the item scores is the raw score for each subscale.

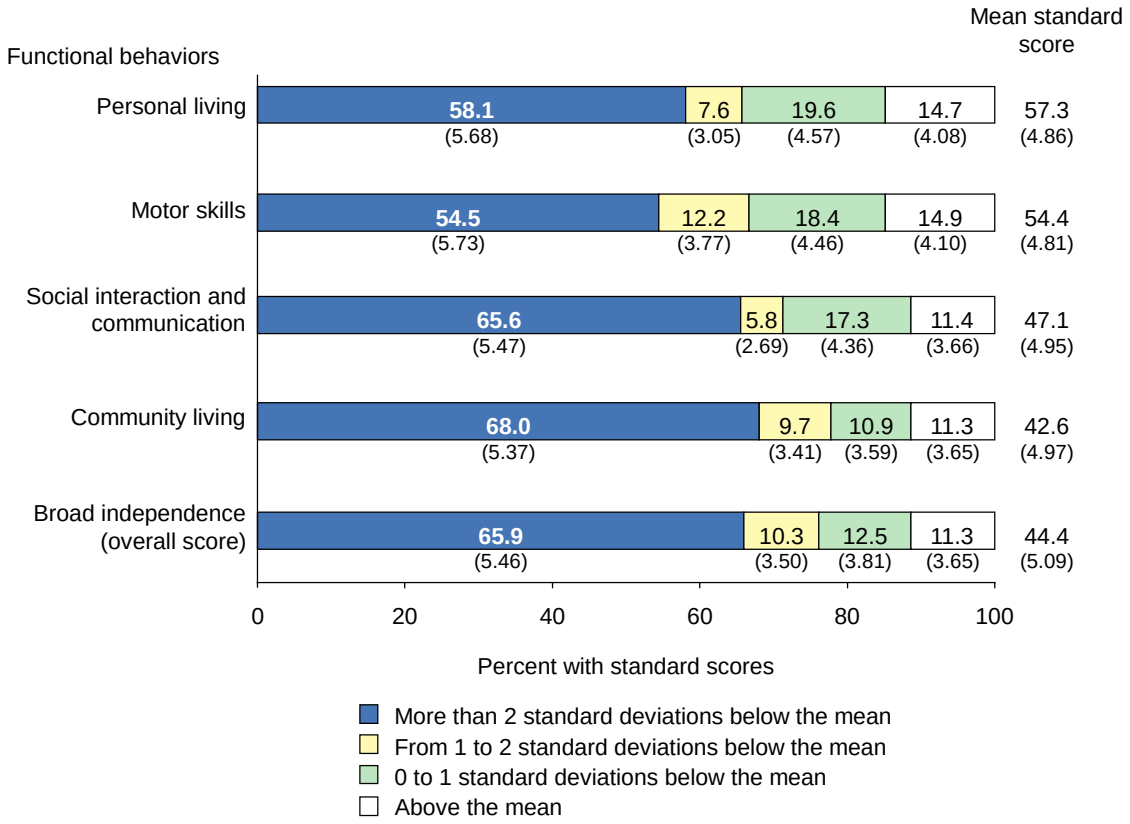
generalizing from the score for a single, narrow behavior (such as time and punctuality) to a broad, multifaceted ability (community living skills).

Functional Skills

Findings regarding functional skills are reported as standard scores, which have a mean of 100 and a standard deviation of 15. In the general population, the distribution of test scores on each cluster is equally divided above and below the mean, and more than 80 percent have skills that are within one standard deviation below the mean or higher. Individuals scoring 100 or above are considered to have "average" to "high-average" and above functional skills for youth their age in the general population (Bruininks et al. 1996). Youth scoring up to one standard deviation below the mean have "average" to "low-average" functional skills, and those scoring from one to two standard deviations below the mean have "low" functional skills. Youth who are rated more than two standard deviations below the mean (about 2 percent of the general population) have "very low" functional skills. Youth with standard scores six or more standard deviations below the mean likely find most age-appropriate functional skills extremely difficult or impossible to perform.

Figure 4 shows the performance of youth with disabilities for whom a functional rating was completed. As described in appendix B, many youth for whom functional ratings were completed have limitations in sensory, physical, and/or communication domains and in their

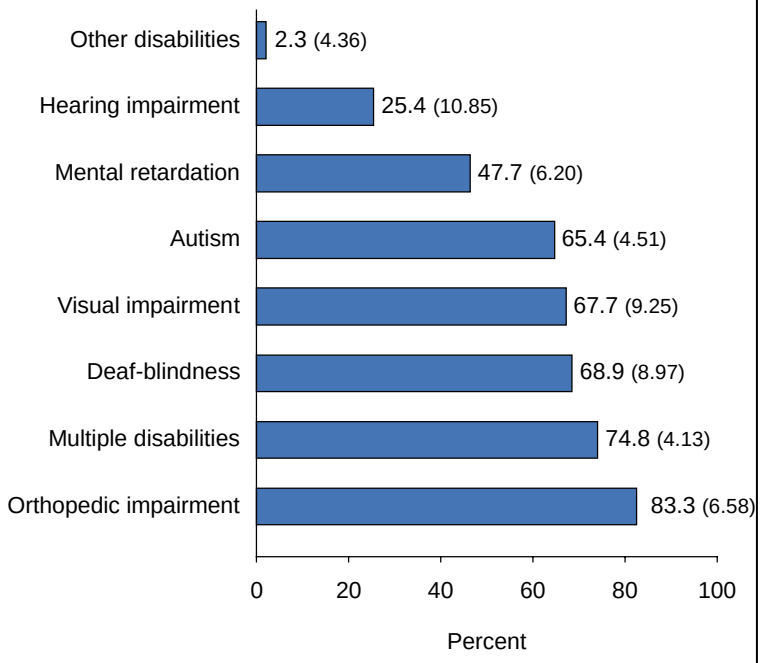
Figure 4. Performance of youth with disabilities on functional rating measures



NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), functional ratings, 2002 and 2004.

Figure 5. Percentage of youth with disabilities performing more than six standard deviations below the mean[1] on functional rating measures



[1] Mean of 100 with standard deviations of .15.

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), functional ratings, 2002 and 2004.

functional cognitive skills. These limitations would be expected to manifest themselves in low scores relative to the general population. Although some of the youth for whom functional ratings were completed perform across the score range, the majority of youth have very low scores. Mean scores range from 43 for community living skills to 57 for personal living skills; across the subtests, only these two means scores differ significantly from each other ($p < .05$). From 22 percent to 38 percent scores more than six standard deviations below the mean (figure 5), with the percentage of youth with these very low scores being significantly larger for the community living skills and broad independence measures than for personal living skills (22 percent vs. 38 percent and 37 percent, $p = .06$). In contrast, from 11 percent to 15 percent score above the mean.

Personal living skills. Five subscales are included in this behavior cluster: eating and meal preparation, dressing, toileting, personal self-care, and domestic skills. The cluster assesses an individual's ability to take care of his or her personal needs primarily within the home. Eating and meal preparation tasks range from simple eating and drinking to meal preparation. Dressing skills range from being able to remove clothing to appropriately selecting and maintaining clothes. Toileting tasks range from using the toilet regularly to selecting and using appropriate toilet facilities outside the home. Personal self-care tasks evaluate performance in basic grooming and health maintenance, with tasks ranging from using a toothbrush to seeking medical care for illness. Domestic skills range from placing dishes in or near the sink to home maintenance tasks. Youth for whom functional ratings were completed have an average standard score on this subscale of 57. More than half (58 percent) are rated more than two standard deviations below the mean, and 22 percent are rated more than six standard deviations below the mean. About one-third of youth with disabilities for whom functional ratings were completed are rated as having personal living skills within one standard deviation below the mean or higher.

Motor skills. The gross and fine motor skills subscales in this cluster assess a range of motor proficiency tasks involving mobility, balance, endurance, and coordination of arms, legs, and the entire body, as well as precise movements ranging from picking up small objects to assembling objects consisting of small parts. The mean standard score on this cluster for youth

who have a functional rating is 54. Fifty-five percent of youth score more than two standard deviations below the mean, with 28 percent scoring more than six standard deviations below the mean. About one-third of youth for whom functional ratings were completed are rated as having low-average to above-average age-appropriate gross and fine motor skills and can perform most motor skill tasks adequately.

Social interaction and communication skills. This skill cluster consists of three subscales: social interaction, language comprehension, and language expression. It measures an individual's interaction with others in various social settings and his or her understanding and communication of information through signs, oral expression, or written symbols. Social interaction skills range from basic interactions, such as handing an object to another person, to more complex interactions involving making plans to attend social activities outside the home. Language comprehension tasks range from recognizing one's name to more complex behaviors, such as securing information through reading or listening. Language expression items evaluate the ability to communicate by talking as well as other forms of expression, including sign and language boards. Tasks range in difficulty from indicating yes or no and repeating common words to preparing and presenting formal reports.

The average standard score for youth with disabilities who have a functional rating is 47. About two-thirds of these youth are rated more than two standard deviations below the mean, with almost one-third scoring more than six standard deviations below the mean. About 28 percent of youth for whom functional ratings were completed are rated as having low-average to above average social and communication skills.

Community living skills. Subscales in this cluster measure the skills needed to use community resources and perform in work and social settings. Tasks involving time and punctuality range from understanding the basic concept of the time of day to the ability to keep appointments. Tasks involving money and value range from selecting particular coins to complex consumer decisions involving investments and credit. Skills required in the workplace or at school range from indicating when a task is finished to vocational skills, such as completing job applications. Orientation skills required in the home and community range from getting around the home and neighborhood to traveling in the community.

The community living skills mean standard score is 43. More than two-thirds are rated more than two standard deviations below the mean, with 38 percent more than six standard deviations below the mean. About 20 percent of youth for whom functional ratings were completed are rated as low-average to above-average on functional skills, being able to perform age-appropriate community living skills at least adequately.

Broad independence. On this overall measure of functional independence, the mean standard score is 44. Fully two-thirds of youth for whom a functional rating was completed have functional independence skills more than two standard deviations below the mean compared with about 2 percent of youth in the population as a whole. Thirty-seven percent of youth for whom functional ratings were completed are rated more than six standard deviations below the mean and experience great difficulty in performing age-appropriate functional skills. About one-fourth of youth for whom functional ratings were completed are rated higher than one standard deviation below the mean and are likely to perform most age appropriate functional living skills at least adequately; this compares with more than 80 percent of youth in the general population who have functional skills in this range.

Disability Differences in Functional Performance

Table 3 shows the range in standard scores on the measure of broad independence for youth across disability categories for whom functional ratings were completed. Few youth in the categories of learning disability, speech/language impairment, emotional disturbance, other health impairment, and traumatic brain injury were assessed using the functional rating, so the scores for youth in those categories cannot be reported separately. However, their combined scores are reported to allow comparison with youth in each of the other disability categories. Among youth with "other disabilities," 21 percent are rated more than two standard deviations below the mean, and about 60 percent are rated in the low-average to above-average range; their mean standard score on broad independence is 90. This score is significantly higher than the mean standard scores for youth in all of the other disability categories; these range from 10 for youth with orthopedic impairments or multiple disabilities to 53 for youth with hearing impairments ($p < .01$ compared with youth with hearing impairments and $p < .001$ for all other comparisons).

Table 3. Standard scores on the broad independence measure, by disability category

Standard scores	Other disabilities[1]	Mental retardation	Hearing impairment	Visual impairment	Orthopedic impairment	Autism	Multiple disabilities	Deaf-blindness
	Percent / standard error							
Percentage with standard scores:								
More than two standard deviations below the mean	20.6 (11.70)	89.3 (3.84)	68.3 (11.59)	80.5 (7.84)	92.4 (4.68)	96.2 (1.81)	95.8 (1.91)	94.9 (4.26)
From 1 to two standard deviations below the mean	18.7 (11.28)	7.3 (3.23)	7.4 (6.52)	8.9 (5.64)	3.2 (3.11)	2.0 (1.33)	2.0 (1.33)	2.3 (2.91)
0 to 1 standard deviations below the mean	31.2 (13.41)	2.1 (1.78)	10.6 (7.67)	10.6 (6.09)	4.3 (3.58)	0 †	1.4 (1.12)	2.8 (3.20)
Above the mean	29.6 (13.22)	1.3 (1.41)	13.8 (8.59)	0 †	0 †	1.9 (1.29)	0.8 (0.85)	0 †
Mean standard score	90.1 (8.82)	22.7 (3.41)	52.7 (10.31)	21.4 (6.86)	9.5 (4.34)	14.2 (2.29)	10.2 (2.10)	13.3 (4.55)

† Not applicable.

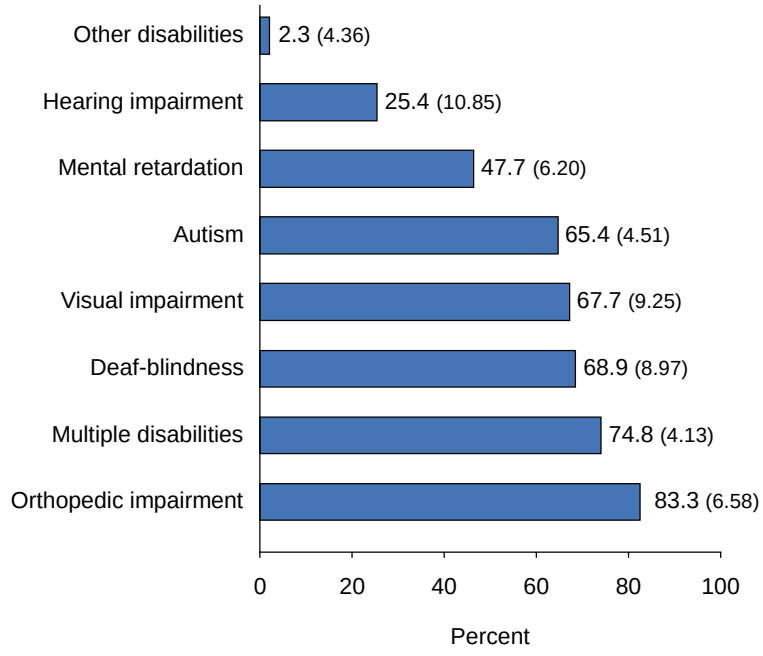
[1] "Other disabilities" includes the categories of learning disability, speech/language impairment, emotional disturbance, other health impairment, and traumatic brain injury.

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), student assessments, 2002 and 2004.

As with the distribution of scores for youth with a functional rating as a whole, and with the exception of youth in the "other disabilities" category, the majority of youth in each disability category are rated more than two standard deviations below the mean. In fact, the majority of youth with autism, visual or orthopedic impairments, deaf-blindness, or multiple disabilities are rated more than six standard deviations below the mean (figure 6). Among youth with hearing

Figure 6. Percentage of youth with disabilities performing more than six standard deviations below the mean on the measure of broad independence



NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), functional ratings, 2002 and 2004.

impairments, about one-fourth are rated in the low-average to above-average range, whereas fewer than 10 percent of youth with visual impairment and 5 percent of youth with mental retardation, orthopedic impairment, autism, multiple disabilities, or deaf-blindness are rated in this range. With the exception of youth with "other disabilities," those with hearing impairments have a significantly higher mean standard score on the broad independence measure than youth in all other categories ($p < .05$ compared with youth with visual impairments, $p < .01$ for youth with mental retardation, and $p < .001$ for youth with orthopedic impairments, autism, multiple disabilities, and deaf-blindness. Youth with mental retardation also are rated higher than their peers in some disability categories, including youth with orthopedic impairments, autism, and multiple disabilities ($p < .05$ compared with youth with orthopedic impairments and autism and $p < .01$ compared with youth with multiple disabilities).

Comparing the percentages of youth for whom a functional rating was completed who are rated more than six standard deviations below the mean shows that 83 percent of youth with orthopedic impairments are rated in this range, a significantly higher rate than for youth with hearing impairments (25 percent, $p < .001$), mental retardation (48 percent, $p < .001$), or autism (65 percent, $p < .05$), whereas youth with hearing impairments are less likely to score in this range (25 percent do so) than youth with autism (65 percent, $p < .001$) or multiple disabilities (75 percent, $p < .01$).

Demographic Differences in Functional Performance

There are few statistically significant differences in performance for youth with a functional rating who differ in gender, age, and racial/ethnic background. Exceptions are presented below.

Gender. Boys with disabilities are less likely than girls to be rated more than two standard deviations below the mean on motor skills (47 percent vs. 71 percent, $p < .05$), personal living skills (51 percent vs. 75 percent, $p < .05$), and community living skills (61 percent vs. 84 percent,

$p < .05$). Instead, boys are more likely than girls to have scores that are one to two standard deviations below the mean on community living skills (13 percent vs. 2 percent, $p < .05$) and are more likely than girls to be rated in the first standard deviation below the mean on motor skills (23 percent vs. 7 percent, $p < .05$), personal living skills (26 percent vs. 4 percent, $p < .01$), and the overall broad independence (17 percent vs. 3 percent, $p < .05$). Only for motor skills are the mean standard scores significantly different between genders, with scores for boys being significantly higher than those for girls (60 vs. 42, $p < .05$).

Age. Greater percentages of 17- and 18-year-olds than 16-year-olds are rated in the low-average to average range, the first standard deviation below the mean (24 percent and 23 percent vs. 5 percent, respectively, $p < .05$). Seventeen-year-olds are more likely than 18-year-olds to be rated in this range on community living skills (18 percent vs. 1 percent, $p < .05$) and are more likely than 16-year-olds to be rated in this range on the measure of broad independence (16 percent vs. 1 percent, $p < .05$).

Race/ethnicity. A greater percentage of White than African American youth are rated above the mean on community living skills (15 percent vs. 2 percent, $p < .05$), and a greater percentage of African American than White or Hispanic youth are rated in the first standard deviation below the mean on broad independence (27 percent vs. 11 percent and <1 percent, $p < .05$ and $p < .01$, respectively).

Summary

This chapter has described the performance of youth with disabilities on four functional dimensions and on an overall measure of broad independence, derived from the SIB-R. Across the measures, the large majority of youth with disabilities for whom the functional rating was completed are rated at the lowest performing level relative to youth in the general population. However, across measures, some youth with disabilities for whom a functional rating was completed were rated as having skills within average performance levels relative to youth in the general population.

6. A Summary of the National Picture of the Academic Achievement and Functional Performance of Youth With Disabilities

Although national information on the academic achievement of the general student population is collected regularly (National Center for Education Statistics 2005c), similar information has not been available for students with disabilities. By administering a uniform direct assessment and functional rating for a nationally representative sample of youth with disabilities, NLTS2 is helping to fill this gap in the knowledge base. This report describes assessment results and identifies selected individual and household characteristics and school experiences that are statistically related to variations in academic achievement. The major findings that have emerged from these analyses are highlighted below.

Academic Achievement

Youth with disabilities were assessed on two measures of language arts abilities, two measures of mathematics skills, and two measures of content knowledge. Results of the NLTS2 direct assessment demonstrate that by the time students who receive special education services reach secondary school, serious academic deficits are apparent for many students. Across the six measures of academic performance included in the NLTS2 direct assessment, from 13 percent to 27 percent of youth with disabilities score below 70—i.e., more than two standard deviations below the mean for the general population—six to eight times as many as youth in the general population. Average standard scores range from 79 to 87 across the measures for youth with disabilities, whereas 100 is the average for the general population.

Youth with disabilities have the greatest difficulty with understanding what they read; the mean passage comprehension standard score of 79 is significantly lower than any of the other scores. Vocabulary, as measured by the use of synonyms and antonyms, appears to pose the fewest challenges, with a mean standard score of 87, significantly higher than all other scores. Despite the low scores overall, on each measure, some youth with disabilities have scores above the mean; across the measures, from 12 percent to 23 percent score above 100.

Low academic achievement scores are pervasive across disability categories. For example, although youth with other health or visual impairments have mean standard scores within one standard deviation of the mean for the general population on all subtests, even within these groups, 13 percent and 20 percent, respectively, score more than two standard deviations below the mean on passage comprehension, for example. Among youth with the cognitive impairment denoted by mental retardation, those who were able to participate in the direct assessment have average standard scores that are all more than two standard deviations below the mean, ranging from 56 to 67 across measures.

Within a disability category, there is considerable variation in scores across the various academic areas. For example, those with hearing impairments have a significantly higher score on mathematics calculation than any other subtest, whereas those with orthopedic impairments have better vocabulary skills than any others tested. Further, in every disability category, some youth have scores that fall above the mean for the general population.

Factors Related to Academic Achievement

Multivariate analyses confirm that the academic achievement of youth with disabilities is related to an array of individual and household factors and school experiences.[1] Independent of the kinds of differences across disability categories noted above, the functional abilities of youth relate strongly to academic achievement. Youths' ability to cognitively process information in performing everyday tasks is the aspect of functioning that is most strongly and consistently related to academic achievement. Regardless of the nature of their disability, youth whose parents report that they are better able to read signs, count change, tell time on an analog clock, and look up telephone numbers demonstrate better performance of the skills that are measured in the direct assessment. In addition, those whose parents report that youth's disabilities affect their functioning broadly and that they were affected by their disabilities earlier in their lives have lower scores on many of the subtests than youth whose disabilities are reported to affect a more limited number of domains or who did not manifest a disability as early, independent of other factors included in the analyses.

Disability-related factors explain much of the variation in measures of youth's academic achievement. But apart from differences in their disabilities and functioning, differences in their racial/ethnic backgrounds and in the incomes of the households in which they live also relate significantly to differences in their academic achievement. Independent of other differences between them on factors included in the analyses, youth who are White score significantly higher on all measures than youth with disabilities who are African American, Hispanic, or members of any of the other racial/ethnic groups. Differences in performance between White youth with disabilities and those in other groups range from 7 to 13 standard score points across measures, with no particular difference in the pattern for different racial/ethnic groups.

Being from a low-income household has a relationship to academic achievement that is independent of race/ethnicity. Across subtests, youth from households with incomes of $25,000 or less a year are estimated to score 3 to 5 standard score points lower than youth from moderate-income households (those earning $25,001 to $75,000 per year), irrespective of other differences between them. Youth from high-income households (those earning more than $75,000) do not differ significantly in their average assessment scores from youth from moderate-income households. Taken together with the racial/ethnic differences described above, on average, a low-income Hispanic youth with disabilities is likely to score 15 points lower on passage comprehension than a White peer from a moderate-income household, holding other factors in the analysis constant.

Gender also has significant relationships with academic achievement on four of the six subtests, favoring boys. On both measures of mathematics abilities and both measures of content knowledge, boys outscore girls by 3 or 4 standard score points. There are no gender differences on the two measures of language arts abilities.

Variations in families' support for the education of their adolescent children with disabilities add significantly to an understanding of their patterns of academic achievement. Of particular note are the significant relationships between achievement on all measures and parents' expectations regarding the postsecondary education of their children. NLTS2 findings

[1] Readers should note that the findings summarized in this section do not imply that the factors discussed have a causal association with academic achievement.

suggest that among youth whose disability, functioning, and demographics are similar, each increasingly higher level of educational expectations is associated with 4 to 6 additional standard score points across the subtests.

Because the NLTS2 database is not yet complete, relationships that course-taking and instructional settings, for example, may have with academic achievement have yet to be explored. However, the associations between academic achievement and several other school experiences have been examined, with few significant relationships emerging. Exceptions are that higher absenteeism is associated with lower scores on both mathematics subtests, and having disciplinary problems at school is associated with lower mathematics calculation abilities. No associations have been identified between academic achievement and grades, grade retention, or school mobility.

Finally, providing youth with some kinds of accommodations during the assessment session is associated with variation in scores, but not in a consistent direction. Specifically, using a calculator is associated with an additional 3 or 4 standard score points on the two mathematics subtests, independent of differences on other factors included in the analyses. In contrast, using American Sign Language or a sign language interpreter is associated with scores that are from 5 to 9 points lower across measures. Similarly, taking breaks during the assessment session or requiring multiple sessions to complete the assessment is associated with lower scores on five of the six subtests, ranging from 4 to 6 points.

The Functional Performance of Youth With Disabilities

Youth for whom a functional rating was completed were assessed on four clusters of functional skills (motor skills, social interaction and communication, personal living skills, and community living skills) and on an overall measure of independence. The mean standard scores for youth with disabilities across the measures range from 43 to 57, compared with 100 for youth in the general population. From 22 percent to 38 percent of youth with disabilities have scores that are more than six standard deviations below the mean for youth in the general population, indicating that functional skills are extremely difficult or impossible to complete for many youth with disabilities. However, across measures, from 11 percent to 15 percent of youth with disabilities score above the mean for the general population. Personal living skills are somewhat stronger than community living skills or overall independence, as indicated by mean standard scores for those subtests.

The few youth with learning disabilities, speech or other health impairments, emotional disturbances, or traumatic brain injuries who have a functional rating together scored higher on the overall measure of broad independence than youth in other disability categories, with a mean score of 90. Youth with hearing impairments, who have a mean score of 53, also outscored other categories, whose scores ranged from 10 (youth with orthopedic impairments) to 23 (youth with mental retardation). Only one statistically significant difference across measures is apparent in the mean standard scores of youth who differ in gender, age, household income, or racial/ethnic background; with the mean motor skills score for boys with disabilities is significantly higher than the score for girls. In addition, some differences in the percentage of youth in particular standard deviation categories are apparent, but no consistent overall patterns have emerged.

The links between academic achievement and both school completion and early postschool outcomes will be a focus of upcoming NLTS2 analyses. Those analyses will illuminate the relationships between learning in school and youth's later ability to pursue their educations, find employment, and become independent and productive members of their communities.

References

Andrich, D. (1988). *Rasch Models for Measurement*. Newbury Park, CA: Sage Publications.

Anderman, E.M., and Maehr, M.L. (1994). Motivation and Schooling in the Middle Grades. *Review of Educational Research, 64*: 287-309.

Blackorby, J., Chorost, M., Garza, N., and Guzman, A. (2003). The Academic Performance of Secondary School Students With Disabilities. In M. Wagner, C. Marder, J. Blackorby, R. Cameto, L. Newman, P. Levine, and E. Davies-Mercier (with M. Chorost, N. Garza, A. Guzman, and C. Sumi), *The Achievements of Youth With Disabilities During Secondary School*. Menlo Park, CA: SRI International.

Blackorby, J., Chorost, M., Garza, N., and Guzman, A. (2004). The Academic Performance of Elementary and Middle School Students With Disabilities. In J. Blackorby, M. Wagner, R. Cameto, E. Davies, P. Levine, L. Newman, C. Marder, and C. Sumi (with M. Chorost, N. Garza, and A. Guzman), *Engagement, Academics, Social Adjustment, and Independence: The Achievements of Elementary and Middle School Students With Disabilities*. Menlo Park, CA: SRI International.

Bruininks, R.H., Woodcock, R.W., Weatherman, R.F., and Hill, B.K. (1996). *Scales of Independent Behavior-Revised.* Chicago: Riverside Publishing.

Cameto, R., Marder, C., Cadwallader, T.W., and Wagner, M. (2003). The Daily Living and Social Skills of Youth With Disabilities. In M. Wagner, C. Marder, P. Levine, R. Cameto, T. W. Cadwallader, and J. Blackorby (with D. Cardoso and L. Newman). *The Individual and Household Characteristics of Youth With Disabilities*. Menlo Park, CA: SRI International.

Cizek, G.C. (2001). Test Review of the Woodcock-Johnson III. In B.S. Plake, J.C. Impara, and R.A. Spies (Eds.), *The Fifteenth Mental Measurements Yearbook* [Electronic version]. Retrieved November 3, 2004 from http://www.unl.edu/buros

Deci, E.L., Hodges, R., Pierson, L., and Tomassone, J. (1992). Motivationally Relevant Predictors of Achievement and Adjustment in Learning Disabled and Emotionally Handicapped Students. *Journal of Learning Disabilities, 25*: 457-471.

Demie, F. (2002). Pupil Mobility and Educational Achievement in Schools: An Empirical Analysis. *Educational Research, 44*(2): 197-215.

Duncan, G.J., and Brooks-Gunn, J. (1997). *Consequences of Growing Up Poor*. New York: Russell Sage Foundation.

Epstein, J. (1996). Perspectives and Previews on Research and Policy for School, Family, and Community Partnerships. In A. Booth and J.F. Dunn (Eds.), *Family-School Links: How Do They Affect Educational Outcomes?* (pp. 209-246). Mahwah, NJ: Lawrence Erlbaum Associates.

Freeman, C. (2004). *Trends in Educational Equity of Girls & Women: 2004*. Washington, DC: National Center for Education Statistics. Retrieved December 12, 2004 from http://nces.ed.gov/pubs2005/2005016.pdf

Gresham, F.M., and Elliott, S.N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Services, Inc.

Henderson, A.T. and Berla, N. (1994). *A New Generation of Evidence. The Family Is Critical to Student Achievement*. Columbia, MD: National Committee for Citizens in Education.

Kao, G., and Thompson, J.S. (2003). Racial and Ethnic Stratification in Educational Achievement and Attainment. *Annual Review of Sociology, 29*: 417-442.

Kerbow, D. (1996). Patterns of Urban Student Mobility and Local School Reform. *Journal of Education of Students Placed at Risk, 1*: 147-169.

Marder, C., Levine, P., Wagner, M., and Cardoso, D. (2003). Household Characteristics of Youth With Disabilities. In M. Wagner, C. Marder, P. Levine, R. Cameto, T.W. Cadwallader, and J. Blackorby (with D. Cardoso and L. Newman). *The Individual and Household Characteristics of Youth With Disabilities*. Menlo Park, CA: SRI International.

Marder, C., Levine, P., and Wagner, M. (2003). Demographic Characteristics of Youth With Disabilities. In M. Wagner, C. Marder, P. Levine, R. Cameto, T.W. Cadwallader, and J. Blackorby (with D. Cardoso and L. Newman). *The Individual and Household Characteristics of Youth With Disabilities*. Menlo Park, CA: SRI International.

Mather, N., and Woodcock, R.W. (2001). Examiners Manual. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

National Association of School Psychologists. (2002). *Social Skills: Promoting Positive Behavior, Academic Success, and School Safety*. Retrieved December 22, 2005 from http://www.naspcenter.org/factsheets/socialskills_fs.html

National Center for Education Statistics. (2005a). *Mathematics Achievement-Level Results for the Nation*. Retrieved November 17, 2005 from http://nces.ed.gov/nationsreportcard/nrc/reading_math_2005/s0019.asp?tab_id=tab2&subtab_id=Tab_1&printver=#chart

National Center for Education Statistics. (2005b). *NAEP Inclusion Policy*. Retrieved November 17, 2005 from http://nces.ed.gov/nationsreportcard/about/inclusion.asp

National Center for Education Statistics. (2005c). *National Assessment of Educational Progress. The Nation's Report Card. Overview*. Retrieved December 14, 2005 from http://nces.ed.gov/nationsreportcard/about/

National Center for Education Statistics. (2005d). Table 110. Average Student Scale Score in Reading, by Age and Selected Student and School Characteristics: Selected Years, 1971 to 2004. *Digest of Education Statistics*, Washington, DC: Author. Retrieved October 18, 2005 from http://nces.ed.gov/programs/digest/d04/tables/dt04_110.asp

Newman, L., Davies, E., and Marder, C. (2003). School Engagement of Youth With Disabilities. In M. Wagner, C. Marder, J. Blackorby, R. Cameto, L. Newman, P. Levine, E. Davies-Mercier, (with M. Chorost, N. Garza, A. Guzman, and C. Sumi), *The Achievements of Youth With Disabilities in Secondary School*. Menlo Park, CA: SRI International. Available at http://www.nlts2.org/reports/achievement_report.html

Rumberger, R.W. (2002). *Student Mobility and Academic Achievement. ERIC Digests*. Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education (ERIC ED466314).

Thorkildsen, R., and Stein, R. (1998). Is Parent Involvement Related to Student Achievement? Exploring the Evidence. *Research Bulletin, 22*: 17–20.

Thurlow, M., Quenemoen, R., Thompson, S., and Lehr, C. (2001). *Principles and Characteristics of Inclusive Assessment and Accountability Systems* (Synthesis Report 40). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Wagner, M. (2003). The Education-Related Histories of Secondary School Students With Disabilities. In M. Wagner, L. Newman, R. Cameto, P. Levine, and C. Marder, *Going to School: Instructional Contexts, Programs, and Participation of Secondary School Students With Disabilities*. Menlo Park, CA: SRI International.

Wagner, M., Blackorby, J., and Hebbeler, K. (1993). *Beyond the Report Card: The Multiple Dimensions of Secondary School Performances of Students With Disabilities*. Menlo Park, CA: SRI International.

Wagner, M., Blackorby, J., Marder, C., and Levine, P. (2003). The Functional Abilities of Youth. In M. Wagner, C. Marder, P. Levine, R. Cameto, T.W. Cadwallader, and J. Blackorby (with D. Cardoso and L. Newman). *The Individual and Household Characteristics of Youth With Disabilities*. Menlo Park, CA: SRI International.

Wagner, M., Marder, C., and Cardoso, D. (2003). Disability Profiles of Youth With Disabilities. In M. Wagner, C. Marder, P. Levine, R. Cameto, T.W. Cadwallader, and J. Blackorby (with D. Cardoso and L. Newman). *The Individual and Household Characteristics of Youth With Disabilities*. Menlo Park, CA: SRI International.

Wiest, D.J., Wong, E.H., and Kreil, D.A. (1998). Predictors of Global Self-Worth and Academic Performance Among Regular Education, Learning Disabled, and Continuation High School Students. *Adolescence, 22*: 601-618.

Woodcock, R.W., and Johnson, M.B. (1989). *Tests of Cognitive Ability: Standard and Supplemental Batteries*. Itasca, IL: Riverside Publishing.

Woodcock, R.W., and Mather, N. (1990). WJ-R Tests of Achievement: Examiner's Manual. In R.W. Woodcock and M.B. Johnson, *Woodcock-Johnson Psycho-Educational Battery – Revised*. Itasca, IL: Riverside Publishing.

Woodcock, R.W, McGrew, K.S., and Mather, N. (2001). *Woodcock-Johnson Tests of Academic Achievement–Research edition.* Itasca, IL: Riverside Publishing.

Wright, B.D., and Stone, M.H. (1979). *Best Test Design*. Chicago: Mesa Press.

Appendix A

NLTS2 Sampling, Data Collection, and Analysis Procedures

Appendix A. NLTS2 Sampling, Data Collection, and Analysis Procedures

Appendix A describes the following aspects of the NLTS2 methodology relevant to the data reported here:

- sampling local education agencies (LEAs) and students;
- data sources and response rates;
- combining data from multiple sources;
- weighting the youth assessment data;
- estimation and use of standard errors;
- unweighted and weighted sample sizes;
- calculating statistical significance; and
- disability classifications.

NLTS2 Sample Overview

The NLTS2 sample was constructed in two stages. A stratified random sample of 3,634 local education agencies (LEAs) was selected from the universe of approximately 12,000 LEAs that serve students receiving special education in at least one grade from 7th through 12th grades. These LEAs and 77 state-supported special schools that served primarily students with hearing and vision impairments and multiple disabilities were invited to participate in the study, with the intention of recruiting 497 LEAs and as many special schools as possible from which to select the target sample of about 12,000 students. The target LEA sample was reached; 501 LEAs and 38 special schools agreed to participate and provided rosters of students receiving special education in the designated age range, from which the student sample was selected.

The roster of all students in the NLTS2 age range who were receiving special education from each LEA[1] and special school was stratified by disability category. Students then were selected randomly from each disability category. Sampling fractions were calculated that would produce enough students in each category so that, in the final study year, findings will generalize to most categories individually with an acceptable level of precision, accounting for attrition and for response rates to the parent/youth interview. A total of 11,276 students were selected and eligible to participate in NLTS2.

Details of the LEA and student samples are provided below.

[1] LEAs were instructed to include on the roster any student for which they were administratively responsible, even if the student was not educated within the LEA (e.g., attended school sponsored by an education cooperative or was sent by the LEA to a private school). Despite these instructions, some LEAs may have underreported students served outside the LEA.

The NLTS2 LEA Sample

Defining the Universe of LEAs

The NLTS2 sample includes only LEAs that have teachers, students, administrators, and operating schools—that is, "operating LEAs." It excludes such units as supervisory unions; Bureau of Indian Affairs schools; public and private agencies (e.g., correctional facilities); LEAs from U.S. territories; and LEAs with 10 or fewer students in the NLTS2 age range, which would be unlikely to have students with disabilities.

The public school universe data file maintained by Quality Education Data (QED 1999) was used to construct the sampling frame because it had more recent information than the alternative list maintained by the National Center for Education Statistics. Correcting for errors and duplications resulted in a master list of 12,435 LEAs that met the selection criteria. These comprised the NLTS2 LEA sampling frame.

Stratification

The NLTS2 LEA sample was stratified to increase the precision of estimates, to ensure that low-frequency types of LEAs (e.g., large urban districts) were adequately represented in the sample, to improve comparisons with the findings of other research, and to make NLTS2 responsive to concerns voiced in policy debate (e.g., differential effects of federal policies in particular regions, LEAs of different sizes). Three stratifying variables were used.

Region. This variable captures essential political differences and subtle differences in the organization of schools, the economic conditions under which they operate, and the character of public concerns. The regional classification variable selected was used by the Department of Commerce, the Bureau of Economic Analysis, and the National Assessment of Educational Progress (categories are Northeast, Southeast, Midwest, and West).

LEA size (student enrollment). LEAs vary considerably by size, the most useful available measure of which is student enrollment. A host of organizational and contextual variables are associated with size that exert considerable potential influence over the operations and effects of special education and related programs. In addition, total enrollment serves as an initial proxy for the number of students receiving special education served by an LEA. The QED database provides enrollment data from which LEAs were sorted into four categories serving approximately equal numbers of students:

- **very large** (estimated[2] enrollment greater than 14,931 in grades 7 through 12);
- **large** (estimated enrollment from 4,661 to 14,930 in grades 7 through 12);
- **medium** (estimated enrollment from 1,622 to 4,660 in grades 7 through 12); and
- **small** (estimated enrollment from 11 to 1,621 in grades 7 through 12).

[2] Enrollment in grades 7 through 12 was estimated by dividing the total enrollment in all grade levels served by an LEA by the number of grade levels to estimate an enrollment per grade level. This was multiplied by 6 to estimate the enrollment in grades 7 through 12.

LEA/community wealth. As a measure of district wealth, the Orshansky index (the proportion of the student population living below the federal definition of poverty, Employment Policies Institute, 2002) is a well-accepted measure. The distribution of Orshansky index scores was organized into four categories of LEA/community wealth, each containing approximately 25 percent of the student population in grades 7 through 12:

- **high** (0 percent to 13 percent Orshansky);
- **medium** (14 percent to 24 percent Orshansky);
- **low** (25 percent to 43 percent Orshansky); and
- **very low** (more than 43 percent Orshansky).

The three variables generate a 64-cell grid into which the universe of LEAs was arrayed.

LEA Sample Size

On the basis of an analysis of LEAs' estimated enrollment across LEA size, and estimated sampling fractions for each disability category, 497 LEAs (and as many state-sponsored special schools as would participate) was considered sufficient to generate the student sample. Taking into account the rate at which LEAs were expected to refuse to participate, a sample of 3,635 LEAs was invited to participate, from which 497 participating LEAs might be recruited. A total of 501 LEAs actually provided students for the sample, 101 percent of the target number needed and 14 percent of those invited. Analyses of the region, size, and wealth of the LEA sample, both weighted and unweighted, confirmed that the weighted LEA sample closely resembled the LEA universe with respect to those variables.

In addition to ensuring that the LEA sample matched the universe of LEAs on variables used in sampling, it was important to ascertain whether the stratified random sampling approach resulted in skewed distributions on relevant variables not included in the stratification scheme. Several analyses were conducted.

First, three variables from the QED database were chosen to compare the "fit" between the first-stage sample and the population: the LEA's racial/ethnic distribution of students, the proportion who attended college, and the urban/rural status of the LEA. This analysis revealed that the sample of LEAs somewhat underrepresented African American students and college-bound students and overrepresented Hispanic students and LEAs in rural areas. Thus, in addition to accounting for stratification variables, LEA weights were calculated to achieve a distribution on the urbanicity and racial/ethnic distributions of students that matched the universe.

To determine whether the resulting weights, when applied to the participating NLTS2 LEAs, accurately represented the universe of LEAs serving the specified grade levels, data collected from the universe of LEAs by the U.S. Department of Education's Office of Civil Rights (OCR) and additional items from QED were compared for the weighted NLTS2 LEA sample and the universe. Finally, the NLTS2 participating LEAs and a sample of 1,000 LEAs that represented the universe of LEAs were surveyed to assess a variety of policies and practices known to vary among LEAs and to be relevant to secondary-school-age youth with disabilities. Analyses of both the extant databases and the LEA survey data confirm that the weighted NLTS2 LEA sample accurately represents the universe of LEAs.

The NLTS2 Student Sample

Determining the size of the NLTS2 student sample took into account the duration of the study, desired levels of precision, and assumptions regarding attrition and response rates. Analyses determined that approximately three students would need to be sampled for each student who would have a parent/youth interview in Wave 5 of NLTS2 data collection.

The NLTS2 sample design called for findings to be generalizable to students receiving special education as a whole and for the 12 special education disability categories currently in use and reported in this document. Standard errors were to be no more than 3.6 percent, except for the low-incidence categories of traumatic brain injury and deaf-blindness. Thus, by sampling 1,250 students per disability category (with the two exceptions noted) 402 students per category were expected to have a parent or youth interview in year 9. Assuming a 50 percent sampling efficiency (which is likely to be exceeded for most disability categories), 402 students would achieve a standard error of estimate of slightly less than 3.6 percent. All students with traumatic brain injury or with deaf-blindness in participating LEAs and special schools were selected. Students were disproportionately sampled by age to assure that there would be an adequate number of students who were age 24 or older at the conclusion of the study. Among the eligible students, 40.2 percent will be 24 or older as of the final interview.

LEAs and special schools were contacted to obtain their agreement to participate in the study and request rosters of students receiving special education who were ages 13 through 16 on December 1, 2000 and in at least seventh grade.[3] Requests for rosters specified that they contain the names and addresses of students receiving special education under the jurisdiction of the LEA, the disability category of each student, and the students' birthdates or ages. Some LEAs would provide only identification numbers for students, along with the corresponding birthdates and disability categories. When students were sampled in these LEAs, identification numbers of selected students were provided to the LEA, along with materials to mail to their parents/guardians (without revealing their identity).

After estimating the number of students receiving special education in the NLTS2 age range, the appropriate fraction of students in each category was selected randomly from each LEA and special school. In cases in which more than one child in a family was included on a roster, only one was eligible to be selected. LEAs and special schools were notified of the students selected and contact information for their parents/guardians was requested.

Data Sources

Table A-1 identifies the source of data for each variable included in this report. The data that are the primary focus of the report were collected through in-person assessments of youth (i.e., the "direct assessment") or through checklists completed by teachers or parents that rated the functional performance of youth who for whom the direct assessment was reported to be inappropriate (i.e., the "functional rating"). Analyses of factors associated with variations in academic performance measured through the direct assessment also included data drawn from a survey of parents of NLTS2 youth, conducted by telephone and mail, and mail surveys of staff in schools attended by NLTS2 sample members.

[3] Students who were designated as being in ungraded programs also were sampled if they met the age criteria.

Table A-1. Data sources for variables included in this report

	Direct assessment	Functional rating	Parent interview/ survey	Student's school program survey	General education teacher survey	School/ district student roster
Woodcock-Johnson III scores	√					
SIB-R scores		√				
Disability/functioning						
Disability category						√
Ability to understand speech, see, hear, use arms/hands, use legs/feet			√			
Number domains with functional limitation			√			
Skills scales (self-care, cognitive, social, persistence)			√			
Age at disability identification/first special education services			√			
Individual/household characteristics						
Gender			√			
Race/ethnicity			√			
Age at assessment			√			
Household income			√			
Family involvement and expectations			√			
Services received			√			
Social/employment activities			√			
School experiences						
Attends special school			√			
General education participation				√		
Course-taking				√		
School mobility			√			
Days absent in a month				√		
Grades			√	√	√	
Had school disciplinary actions				√		
Ever retained at grade level			√			

Table A-2. Response rates for NLTS2 youth assessments

	Number	Percent
Wave 1		
Eligible sample	5,071	100.0
Completed:		
Direct assessment	2,583	50.9
Functional rating	577	11.4
Total	3,160	62.3
Wave 2		
Eligible sample	4,343	100.0
Completed:		
Direct assessment	2,639	60.7
Functional rating	474	10.9
Total	3,113	71.6
Total		
Eligible sample	9,414	100.0
Completed:		
Direct assessment	5,222	55.4
Functional rating	1,051	11.2
Total	6,273	66.6

Youth Assessments

Chapter 1 provides an overview of the youth assessment procedures, and chapters 3 and 5 give additional information on the content of the direct assessment and functional rating, respectively. Assessments were administered once for each youth, with the timing of the assessment being determined by a youth's age. As noted in chapter 1, assessment data were collected for 6,273 youth, including 3,160 who were 16 through 18 years old and eligible in Wave 1 (the 2001-02 school year) and 3,113 who were age-eligible in Wave 2 (the 2003-04 school year) (table A-2). Data from the two waves have been combined for the analyses included in this report.

Direct assessment. Data for the seven subtests of the direct assessment are reported as percentile scores—i.e., each youth has a score that corresponds to the percentage of same-age youth in the general population who scored lower than the sample member. The number of items in the subtests range from 23 to 31, and are divided into grade-level subsections, with each section testing increasingly advanced skills or content. The assessor begins each subtest at the designated starting point for the grade level of the youth being assessed. If the youth does not answer the first item appropriate for his or her grade level correctly, the assessor moves down the item progression. For all subtests but synonyms and antonyms, the basal score is established when the student answers three consecutive items correctly; four correct answers establish the basal score for the synonyms and antonyms subtest. If a student can answer the first item in the section appropriate to his or her grade level, that item establishes the basal scores, and the assessor continues with increasingly complex items until the student responds to three consecutive items incorrectly (for synonyms and antonyms the ceiling level is set by the incorrect response to four consecutive items) or until the last item is administered. All items below the basal level are considered correct.

Functional rating. The content of and procedures for administering and scoring the functional rating are described in chapters 1 and 5.

Parent/Youth Data Collection

Unlike the youth assessments, which were administered once for each youth, data are collected repeatedly from parents and, beginning in Wave 2, from youth themselves. Data from both Waves 1 and 2 are included in this report as independent variables in the multivariate analyses and as descriptors of the youth in the direct assessment and functional rating groups.

Wave 1 parent interview/survey. The NLTS2 conceptual framework suggests that a youth's nonschool experiences, such as extracurricular activities and friendships; historical information, such as age when disability was first identified; household characteristics, such as

socioeconomic status; and a family's level and type of involvement in school-related areas are related to student outcomes. Parents/guardians are the most knowledgeable about these aspects of students' lives. They also are important sources of information on outcomes across domains. Thus, parents/guardians of NLTS2 sample members were interviewed by telephone or surveyed by mail in 2001, as part of Wave 1 data collection.

Matches of names, addresses, and telephone numbers of NLTS2 parents with existing national locator databases were conducted to maximize the completeness and accuracy of contact information and subsequent response rates. A student was required to have a working telephone number and an accurate address to be eligible for the parent interview sample.

Letters were sent to parents to notify them that their child had been selected for NLTS2 and that an interviewer would attempt to contact them by telephone. The letter included a toll-free telephone number for parents to call to be interviewed if they did not have a telephone number where they could be reached reliably or if they wanted to make an appointment for the interview at a specific time.

Computer-assisted telephone interviewing (CATI) was used for parent interviews, which were conducted between mid-May and late September 2001. Ninety-five percent of interviews were conducted in English and 5 percent in Spanish.

All parents who could not be reached by telephone were mailed a self-administered questionnaire in a survey period that extended from September through December 2001. The questionnaire contained a subset of key items from the telephone interview.

Overall, 91 percent of respondents reported that they were parents of sample members (biological, adoptive, or step), and 1 percent were foster parents. Six percent were relatives other than parents, 2 percent were nonrelative legal guardians, and fewer than 1 percent reported other relationships to sample members.

Wave 2 parent interview and youth interview/survey. NLTS2 sample members for whom working telephone numbers and addresses were available were eligible for the Wave 2 parent/youth telephone interview in 2003. Database matching procedures were used to maximize the eligible sample, as in Wave 1. Contact procedures alerting parents of the interviews also were similar for the two waves. The major distinction between the data collection methods in Waves 1 and 2 is that interviews were sought both with parents of NLTS2 sample members and with the youth themselves if they were able to respond to questions.

The first interview contact was made with parents of eligible sample members. Those who agreed to participate were interviewed using CATI. Items in this portion of the interview, referred to as Parent Part 1, focused on topics for which the parent was considered the most appropriate respondent (e.g., services received, family expectations, and support). At the end of Parent Part 1, the respondent was asked the following:

> *My next questions are about jobs (YOUTH'S NAME) may have had, schools (he/she) may have gone to, and about (his/her) feelings about (him/herself) and (his/her) life. The questions are similar to those I've been asking you, where (he/she) will be asked to answer using scales, like "very well," "pretty well," "not very well," or "not at all well." The interview would probably last about 20 to 30 minutes. Do you think that*

(YOUTH'S NAME) would be able to accurately answer these kinds of questions over the telephone?

If youth could answer questions by phone, they also were told:

I also have some questions about (his/her) involvement in risk behaviors, like smoking, drinking, and sexual activity. Is it all right for me to ask (YOUTH"S NAME) questions like that?

If parents consented, interviewers asked to speak with the youth or asked for contact information to reach the youth in order to complete the youth portion of the interview, referred to as Youth Part 2.

Parents who reported that youth could not answer questions by telephone were asked:

Would (he/she) be able to accurately answer these kinds of questions using a written questionnaire?

If parents indicated youth could complete a written questionnaire, they were asked for the best address to which to send a questionnaire, and a questionnaire was sent. The questionnaire contained a subset of items from the telephone interview that were considered most important for understanding the experiences and perspectives of youth. Multiple follow-up phone or mail contacts were made to maximize the response rate for the mail survey. Data from the mail survey and Youth Part 2 of the telephone interview were merged for analysis purposes.

If parents reported that youth could not answer questions either by telephone or written questionnaire or declined to have youth asked questions related to risk behaviors, interviewers asked them to continue the interview, referred to as Parent Part 2. If youth were reported to be able to complete a telephone interview or a written questionnaire but did not after repeated attempts, parents were contacted again and asked to complete Parent Part 2 in lieu of Youth Part 2.

Table A-3 reports the sample members for whom there are data from the Wave 1 parent interview and mail survey and from the Wave 2 Parent Part 1 and Parent Part 2 telephone interviews and the Youth Part 2 telephone/mail survey.

Table A-3. Response rates for NLTS2 Waves 1 and 2 parent/youth data collection

	Number	Percent
Wave 1		
Total eligible sample	11,276	100.0
Respondents		
Completed telephone interview	8,672	76.9
Partial telephone interview completed	300	2.7
Complete mail questionnaire	258	2.3
Total respondents	9,230	81.9
Total nonrespondents	2,046	18.1
Wave 2		
Total eligible sample	8,210	100.0
Respondents		
Completed Parent Part 1 telephone interview	6,859	83.5
Completed Parent Part 2 telephone interview	2,962	36.1
Complete Youth Part 2 telephone interview or mail questionnaire	3,360	41.9
Total respondents with Part 1 and either Parent or Youth Part 2	6,322	77.0
Total nonrespondents (no parent or youth data)	1,352	16.5

Waves 1 and 2 School Data Collection

Data sources for the findings reported here also include (1) a mail survey of teachers of general education academic classes—if student took such a class, and (2) a mail survey of school staff who were most knowledgeable about students' overall programs; both surveys were administered in Waves 1 and 2 for youth still in secondary school at those times. The NLTS2 conceptual framework holds that classroom context, curriculum, instruction, accommodations, and assessment are related to student outcomes and are most amenable to intervention. Mail surveys collected information about aspects of the classroom experiences of students with disabilities in general education academic classes and in vocational education and special education settings. Further, students' school experiences extend beyond the classroom, so that related services, IEP goals, participation in district/state assessments all have a place in students' experiences and can relate to student progress. These data are best provided by school staff who are most knowledgeable about the student's classroom experiences and school programs.

The first step in the school data collection process was to identify the school attended by NLTS2 students during the 2001-02 school year for Wave 1 and the 2003-04 school year for Wave 2. School attendance data had been collected as part of the parent interview during the summer and fall preceding each of those school years. Parent responses relating to schools were coded (e.g., address, phone) using the QED database. For identified schools not in the QED database or for students for whom there was no parent interview, school district records collected for sampling were used to identify students' schools. Names of students thought to attend each school were sent to schools for verification using the School Enrollment Form. In addition to verification of enrollment, the school enrollment form requested that schools provide the name of a school staff member (i.e., coordinator) who would be willing to oversee the distribution of school surveys for NLTS2 students attending each school. Participation agreements were signed by coordinators, who received reimbursement for their efforts at varying levels, depending on the number of NLTS2 students in the school.

In the spring of the survey school years, packets were sent to each coordinator and to school principals in schools that did not name a coordinator, which included a general education academic teacher questionnaire for each sample member (with instructions to return the

Table A-4. Response rates for NLTS2 Waves 1 and 2 general education academic teacher and school program surveys

	Teacher Survey		School Program Survey	
	Number	Percent	Number	Percent
Wave 1				
Total eligible sample	Unknown		10,517	100.0
Respondents	2,577		5,588	53.1
Wave 2				
Total eligible sample	Unknown		7,815	100.0
Respondents	1,983		4,078	52.2

NOTE: It is not possible to know the eligible population for the general education teacher survey because identifying whether a student took a general education academic course required a response indicating such enrollment from a student's school. Students with no school response cannot be categorized as either eligible or ineligible.

questionnaire if a student did not have such a class) and a school program questionnaire for each sample member.[4] A second packet was sent a month after the first mailing. Additional mailings were conducted to individual teachers in May of the survey years. Response rates for the teacher and school program surveys are reported in table A-4.

Combining Data from Multiple Data Sources

The multivariate analyses reported in chapter 4 combine data from multiple sources (e.g., dependent variables from the direct assessment and independent variables from the parent interviews). Although any single data source has a reasonably high response rate, a smaller number of youth have data from any particular combination of sources. When sample sizes decline markedly from using multiple data sources, statistical power is reduced and it is difficult for relationships to attain statistical significance even when they are quite large. Hence, it is important to maintain the analytic sample size to the maximum extent possible. It also is important to understand the characteristics of students that are omitted from an analysis as the sample declines. NLTS2 approaches to these two issues are described in this section.

Maintaining the Analytic Sample Size

Two approaches are used in NLTS2 to maintain the size of the sample used in analyses that combine data from multiple sources: constructing composite measures and imputing missing values.

Constructing composite measures. Several variables in NLTS2 can be measured using data from more than one source. The measure of students' grades, which is included in the multivariate analyses reported in chapter 4, is such a variable. Parents were asked to describe students' overall grades, and school staff were asked to report students' grades in specific general education academic and special education classes. In understanding the factors that are related to variation in students' grades, parents' reports were the preferred measure because they were considered the broadest indicator of students' overall grades. However, if a student was missing the grades item from the parent interview, the school-reported grade measure was used, with preference given to the setting (general or special education class) in which the student spent the largest part of his or her school day, as indicated on the school program survey. Thus,

[4] A questionnaire soliciting information about the characteristics of the school also was included in the packet in Wave 1; no data are included in the analyses in this report from the School Characteristics Survey.

the grades variable includes students who have either a parent interview, a school program survey (on which grades are reported for a special education class) or a general education teacher survey, which results in a much larger number of youth included in analyses using a measure of grades than would result from including those with a single data source. No other composite variables are included in this report.

Imputing missing values. Missing values for particular variables occur either because an entire data source is missing for a given student (e.g., a student does not have a parent interview) or a respondent refused to answer or did not know the answer to a given item. Multivariate analyses exclude cases for which there is missing data for any variable included in them, resulting in the difficulties associated with reduced sample sizes already discussed.

Thus, it can be beneficial to impute values on key variables for youth who otherwise would be excluded from analyses because of missing data. Imputation procedures involve assigning a value for a youth with missing data that is the best prediction for that youth given what else is known about him or her. Although there are a variety of procedures for imputation, NLTS2 has employed a straightforward assignment of mean values that are calculated for a subset of youth who resemble the youth with missing values on specified dimensions that are relevant to the variable in question. For example, a student who is missing a value for an item that is included in the scale measuring family support for education at home was assigned the mean value on the missing item that was calculated for all other youth who share his or her disability category and whose head of household has the same level of education. These criteria for subsetting youth for purposes of imputation were selected because they relate significantly to variation in family involvement.

Although imputation can be a significant help in maintaining the analytic sample size, it also reduces the amount of variation in the variables chosen for imputation, thus reducing the strength of their relationships to other variables. Therefore, no dependent variables included imputed values. In selecting independent variables for imputation, careful judgment was used in weighing the trade offs between maintaining sample size and maintaining maximum variability and selecting only those that have a fairly limited number of missing values. Table A-5 identifies the independent variables for which missing values were imputed, the criteria for imputation, and the number and percentage of cases across the multivariate analyses that had imputed values for each variable. For a given variable, the models with the smallest number of imputed values are those with a dependent variable that came from the same data source (i.e., missing data resulted from item nonresponse) whereas a larger number of values were imputed for models addressing variables from a different data source.

Table A-5. Imputation of missing values

Variable Name	Criteria for Assigning Mean Values	Number (Percentage) of Cases with Assigned Values Across Multivariate Analyses
Self-care skills scale	Mean value of youth with same disability category and number of domains with functional limitation	1 (< .1%)
Functional cognitive skills scale	Mean value of youth with same disability category and number of domains with functional limitation	1 or 2 (< = .1%)
Household income	Mean value of youth with same disability category, head of household education, and race/ethnicity	63 to 65 (3.6%)
Family involvement at home	Mean value of youth with same disability category and head of household education	110 to 137 (6.2% to 7.5%)
Family involvement at school	Mean value of youth with same disability category and head of household education	12 to 14 (.7% to .8%)
School mobility—number of school changes other than grade-level progression	Mean value of youth with same disability category, student age, and household income	102 to 109 (5.8% to 6.0%)
Absences excluding suspensions and expulsions (used as an independent variable only)	Mean value of youth with same disability category	690 to 716 (39.1% to 39.4%)

Understanding the characteristics of youth included in analyses. As mentioned above, combining data from multiple sources in a given analysis necessarily limits the youth included in it to those who have all data sources. It is important to understand the extent to which the included subset of youth is similar to or differs from the full sample in order to know whether the results of the analysis generalize to all youth or only to those represented in the subset. To address this question, NLTS2 compared means for all dependent and independent variables used in each multivariate model reported in this document with those of the full sample of youth for whom there are data. The number of cases included in each model and the results of the analyses of means are reported in table A-6.

Table A-6. Unweighted means of variables included in the full direct assessment sample and each multivariate analysis

	Dependent variables—percentile rank in:			
	All youth with assessment data	Passage comprehension, synonyms/antonyms, social studies and science	Mathematics calculation	Applied problems
Sample size	2,176 - 5,235	1,823	1,801	1,751
Percentage in category:				
Speech/language impairment	0.11	0.07	0.09	0.08
Mental retardation	0.09	0.10	0.10	0.10
Emotional disturbance	0.08	0.08	0.08	0.09
Hearing impairment	0.11	0.13	0.13	0.13
Visual impairment	0.08	0.09	0.09	0.07
Orthopedic impairment	0.11	0.11	0.11	0.12
Other health impairment	0.11	0.12	0.12	0.13
Autism	0.08	0.08	0.07	0.08
Traumatic brain injury	0.04	0.05	0.05	0.05
Multiple disabilities/deaf-blindness	0.08	0.06	0.06	0.05
Average:				
Age at disability identification	3.78	4.01	4.03	4.07
Number domains with functional limitation	1.90	1.89	1.88	1.90
Functional cognitive skills scale score	13.37	13.50	13.53	13.66
Social skills scale score	20.46	20.38	20.38	20.31
Persistence scale score	2.29	2.29	2.28	2.28
Age	16.90	16.90	16.90	16.90
Household income scale score	8.68	8.91	8.91	8.90
Family involvement at home scale score	6.85	6.98	6.98	6.99
Family involvement at school scale score	3.12	3.41	3.41	3.41
Expectations for postsecondary education scale score	3.02	2.88	2.89	2.87
Number of school changes other than for grade level advancement	0.99	1.96	1.96	1.96
Days absent in a month	1.99	2.02	2.00	2.04
Grade category	6.21	6.21	6.21	6.18
Percentage:				
Male	0.63	0.62	0.62	0.62
African American	0.20	0.19	0.20	0.19
Hispanic	0.13	0.11	0.11	0.11
Multiple/other racial/ethnic category	0.03	0.03	0.03	0.03
With disciplinary actions	0.25	0.25	0.25	0.26
Ever retained at grade level	0.34	0.36	0.36	0.36
Used as accommodation:				
Breaks or multiple sessions	0.08	0.10	0.10	0.10
ASL or an interpreter	0.08	0.09	0.09	0.09
Special furniture or lighting	0.05	0.05	0.05	0.05
Braille or large print	0.06	0.07	0.07	0.04
Calculator (mathematics subtests only)	0.23	†	0.23	0.22

†=Not applicable to the model.

The subset of cases included in the multivariate analyses depicted in table A-5 are significantly different from the full sample of NLTS2 direct assessment participants on only two variables. The multivariate analyses overrepresent students who change schools often for reasons other than grade level progression. Analyses also include youth who are less likely to be expected by parents to attend college.

Weighting the Youth Assessment Data

The percentages and means reported in the data tables throughout this report are estimates of the true values for the population of youth with disabilities in the NLTS2 age range. The estimates are calculated from the results of the direct and indirect assessments of NLTS2 sample members. The response for each sample member is weighted to represent the number of youth in his or her disability category in the kind of LEA (i.e., region, size, and wealth) or special school from which he or she was selected.

Table A-6 illustrates the concept of sample weighting and its effect on values that are calculated for youth with disabilities as a group. In this example, 10 students are included in a sample, 1 from each of 10 disability groups, and each has a hypothetical percentile score on the Woodcock-Johnson III passage comprehension subtest, ranging from 3 to 30 across the disability categories. Summing the values for the 10 youth results in an average of 19.3 for the full group. However, this would not accurately represent the national population of youth with disabilities because many more youth are classified as having a learning disability than orthopedic or other health impairments, for example. Therefore, in calculating a population estimate, weights in the example are applied that correspond to the proportion of youth in the population who are from each disability category (actual NLTS2 weights account for disability category and several aspects of the districts from which they were chosen). The sample weights for this example appear in column C. Using these weights, the weighted population estimate is 21.2 percent. The percentages in all NLTS2 tables are similarly weighted population estimates, whereas the sample sizes are the actual number of cases on which the weighted estimates are based (similar to the 10 cases in table A-7).

The youth in LEAs and state schools with data for each survey were weighted to represent the universe of students in LEAs and state schools using the following process:

- For each of the 64 LEA sampling cells, an LEA student sampling weight was computed. This weight is the ratio of the number of students in participating LEAs in that cell divided by the number of students in all LEAs in that cell in the universe of LEAs. The weight represents the number of students in the universe who are represented by each student in the participating LEAs. For example, if participating LEAs in a particular cell served 4,000 students and the universe of LEAs in the cell served 400,000 students, then the LEA student sampling weight would be 100.

Table A-7. Example of weighted percentage calculation

	A	B	C	D
Disability Category	Number in Sample	Percentage scoring two standard deviations below the mean on passage comprehension subtest	Example Weight for Category	Weighted Value for Category
Learning disability	1	18	5.5	99.0
Speech/language impairment	1	23	2.2	50.6
Mental retardation	1	73	1.1	80.3
Emotional disturbance	1	17	.9	15.3
Hearing impairment	1	35	.2	7.0
Visual impairment	1	20	.1	2.0
Orthopedic impairment	1	28	.1	2.8
Other health impairment	1	13	.6	7.8
Autism	1	50	.2	2.5
Multiple disabilities	1	56	.1	5.6
Total	10	333	10	272.9
	Unweighted sample percentage = 33.3 (Column B total divided by Column A total)		Weighted population estimate = 27.3 (Column D total divided by Column C total)	

- For each of the 64 LEA cells, the number of students in a disability category was estimated by multiplying the number of students with that disability on the rosters of participating LEAs in a cell by the adjusted LEA student sampling weight for that cell. For example, if 350 students with learning disabilities were served by LEAs in a cell, and the LEA student sampling weight for that cell was 100 (that is, each student in the sample of participating LEAs in that cell represented 100 students in the universe), there would be an estimated 35,000 students with learning disabilities in that cell in the universe.
- For the state schools, the number of students in each disability category was estimated by multiplying the number of students with that disability on the rosters by the inverse of the proportion of state schools that submitted rosters.
- Initial student weights were calculated for each cell by disability as the estimated number of students in that cell divided by the number of respondents in that cell.
- Weights were adjusted by disability category using a raking algorithm so that the sum of the weights by geographical region, wealth, LEA size, and ethnicity was equal to the estimated national distribution for that disability. The adjustments were typically small and essentially served as a nonresponse adjustment. However, the adjustments could become substantial when there were relatively few interviewees (as occurred in the small and medium strata for the lowest-incidence disabilities) because in these cases, there might not be any interviewees in some cells, and it was necessary to adjust the weights of other interviewees to compensate. Two constraints were imposed on the adjustments: (1) within each size stratum, the cells' weights could not vary from the average weight by more than a factor of 2, and (2) the average weight within each size strata could not be larger than 4 times the overall average weight. These constraints

substantially increased the efficiency of the sample at the cost of introducing a small amount of weighting bias.

- In a final step, the weights were adjusted so that they summed to the number of students in each disability category, as reported to OSEP by the states for the 2000-2001 school year (Office of Special Education Programs 2001).

Estimating Standard Errors

Each estimate reported in the data tables is accompanied by a standard error. A standard error acknowledges that any population estimate that is calculated from a sample will only approximate the true value for the population. The true population value will fall within the ranged demarcated by the estimate, plus or minus the standard error 95 percent of the time. For example, if an estimate for youth's current employment rate is 29 percent, with a standard error of 1.8, one can be 95 percent confident that the true current employment rate for the population is between 27.2 percent and 30.8 percent.

Because the NLTS2 sample is both stratified and clustered, calculating standard errors by formula is not straightforward. Standard errors for means and proportions were estimated using pseudoreplication, a procedure that is widely used by the U.S. Census Bureau and other federal agencies involved in fielding complex surveys. To that end, a set of weights was developed for each of 32 balanced half-replicate subsamples. Each half-replicate involved selecting half of the total set of LEAs that provided contact information using a partial factorial balanced design (resulting in about half of the LEAs being selected within each stratum) and then weighting that half to represent the entire universe. The half-replicates were used to estimate the variance of a sample mean by (1) calculating the mean of the variable of interest on the full sample and each half-sample using the appropriate weights; (2) calculating the squares of the deviations of the half-sample estimate from the full sample estimate; and (3) adding the squared deviations and dividing by (n-1) where n is the number of half-replicates.

Although the procedure of pseudoreplication is less unwieldy than development of formulas for calculating standard errors, it is not easily implemented using the Statistical Analysis System (SAS), the analysis program used for NLTS2, and it is computationally expensive. In the past, it was possible to develop straightforward estimates of standard errors using the effective sample size.

When respondents are independent and identically distributed, the effective sample size for a weighted sample of N respondents can be approximated as

$$N_{eff} = N\left(E^2[W] \Big/ E^2[W] + V[W] \right)$$

where N_{eff} is the effective sample size, $E^2[W]$ is the square of the arithmetic average of the weights and $V[W]$ is the variance of the weights. For a variable X, the standard error of estimate can typically be approximated by $\sqrt{V[X]/N_{eff}}$, where $V[X]$ is the weighted variance of X.

NLTS2 respondents are not independent of each other because they are clustered in LEAs, and the intracluster correlation is not zero. Nevertheless, a comparison of standard errors calculated using balance repeated replicate weights and those produced by the formula show that the calculated variances are similar. To be conservative, however, the initial estimate was

multiplied by a "safety factor" that assures that the standard error of estimate is not underestimated.

To determine the adequacy of fit of the variance estimate based on the effective sample size and to estimate the required safety factor, 24 questions with 95 categorical and 2 continuous Wave 1 responses were selected. Standard errors of estimates were calculated for each response category and the mean response to each question for each disability group using both pseudo-replication and the formula involving effective sample size. A safety factor of 1.25 resulted in the effective sample size standard error estimate underestimating the pseudoreplicate standard error estimate for 92 percent of the categorical responses and 89 percent of the mean responses. Because the pseudoreplicate estimates of standard error are themselves estimates of the true standard error, and are therefore subject to sampling variability, this was considered an adequate safety margin.

Unweighted and Weighted Sample Sizes

As indicated above, standard errors accompany all estimates reported in the descriptive data tables. How close an estimate comes to a true population value is influenced by the size of the sample on which the estimate is based. Larger samples yield estimates with smaller standard errors, indicating that those estimates are closer to true population values than estimates with larger standard errors based on smaller samples.

The actual, or "unweighted" sample sizes for each variable reported in the descriptive data tables are included in appendix B. However, some readers may be interested in determining the number of youth in the nation represented by a particular estimate (e.g., if 75 percent of youth with disabilities score in the lowest quartile on the passage comprehension subtest, how many youth in the country perform at that level). A first step in determining these "weighted" sample sizes involves multiplying the weighted estimate for the sample by the actual number of youth in the nation represented by that estimate (see example below). However, 95 percent of the time, the true population value is likely to diverge from that estimate by as much as the amount of the standard error. Therefore, it is more appropriate to use the standard error to calculate a range in the number of youth represented by an estimate, rather than relying on the single value resulting from multiplying the estimate by the size of the population it represents.

Consider the example depicted in table A-8. NLTS2 findings indicate that 11.7 percent of youth with learning disabilities score above the mean on the passage comprehension subtest (see figure 2). The standard error accompanying that estimate is 1.98, indicating that the true rate for the population is likely to fall between 9.7 percent and 13.7 percent. There are 1,130,539 youth with learning disabilities in the NLTS2 age range. Multiplying the percentages by this population size yields a single-point estimate of 132,272 and a range of 109,662 to 154,884, within which the actual population size will fall, with 95 percent confidence.

Table A-8. Example of calculating weighted sample sizes

A	B	C	D	E	F
Percentage Estimate	Standard Error	Range Around Estimate (Column A Plus or Minus Column B)	Population Size	Single-point Weighted Population Affected (Column A x Column D)	Range in Weighted Population Affected (Column C x Column D)
11.7	1.98	9.7 to 13.7	1,130,539	132,273	109,662 to 154,884

Because percentage estimates are provided not only for the full sample of youth with disabilities, but also for youth who differ in primary disability category, readers must have the actual population size for each of these subgroups to calculate weighted sample sizes for some estimates. These population sizes are presented in table A-9 (U.S. Department of Education 2003).

Table A-9. Population sizes of groups represented by NLTS2

Groups	Number
All youth with disabilities	1,838,848
Disability category:	
Learning disability	1,130,539
Speech/language impairment	76,590
Mental retardation	213,552
Emotional disturbance	203,937
Hearing impairment	22,001
Visual impairment	8,013
Orthopedic impairment	21,006
Other health impairment	98,197
Autism	14,637
Traumatic brain injury	6,379
Multiple disabilities	34,865
Deaf-blindness	340

Calculating Significance Levels

In general, references in the text of the report to differences between groups highlight only differences that are statistically significant with at least 95 percent confidence, (denoted as $p < .05$). Beyond the differences highlighted in the text, readers may want to compare percentages or means for specific subgroups to determine, for example, whether the difference in the percentage of students who are male between students with learning disabilities and those with hearing impairments is greater than would be expected to occur by chance. To calculate whether the difference between percentages is statistically significant, the squared difference between the two percentages of interest is divided by the sum of the two squared standard errors. If this product is larger than 3.84, the difference is statistically significant at the .05 level—i.e., it would occur by chance fewer than 5 times in 100. Presented as a formula, a difference in percentages is statistically significant at the .05 level if

$$\frac{(P_1 - P_2)^2}{SE_1^2 + SE_2^2} > 1.96^2$$

where P_1 and SE_1 are the first percentage and its standard error and P_2 and SE_2 are the second percentage and the standard error. If the result of this calculation is at least 6.63, the significance level is .01, products of 10.8 or greater are significant at the .001 level.

Categorizing Students by Primary Disability

Information about the nature of students' disabilities came from rosters of all students in the NLTS2 age range receiving special education services in the 2000-01 school year under the auspices of participating LEAs and state-supported special schools. In analyses in this report, students are assigned to a disability category on the basis of the primary disability designated by the student's school or district. Although there are federal guidelines in making category assignments (table A-10), criteria and methods for assigning students to categories vary from state-to-state and even between districts within states, with the the potential for substantial variation in the nature and severity of disabilities included in the categories (see for example, MacMillan and Siperstein 2002). Therefore, NLTS2 data should not be interpreted as describing students who truly had a particular disability, but rather as describing students who were categorized as having that primary disability.

Table A-10. Definitions of disabilities[1]

Autism. A developmental disability significantly affecting verbal and nonverbal communication and social interaction, generally evident before age 3, that adversely affects a child's educational performance. Other characteristics often associated with autism are engagement in repetitive activities and stereotyped movements, resistance to environmental change or change in daily routines, and unusual responses to sensory experiences. The term does not apply if a child's educational performance is adversely affected primarily because the child has a serious emotional disturbance as defined below.

Deafness. A hearing impairment so severe that the child cannot understand what is being said even with a hearing aid.

Deaf-blindness. A combination of hearing and visual impairments causing such severe communication, developmental, and educational problems that the child cannot be accommodated in either a program specifically for the deaf or a program specifically for the blind.

Emotional disturbance.[2] A condition exhibiting one or more of the following characteristics, displayed over a long period of time and to a marked degree that adversely affects a child's educational performance:

- An inability to learn that cannot be explained by intellectual, sensory, or health factors
- An inability to build or maintain satisfactory interpersonal relationships with peers or teachers
- Inappropriate types of behavior or feelings under normal circumstances
- A general pervasive mood of unhappiness or depression
- A tendency to develop physical symptoms or fears associated with personal or school problems.

This term includes schizophrenia, but does not include students who are socially maladjusted, unless they have a serious emotional disturbance.

Hearing impairment. An impairment in hearing, whether permanent or fluctuating, that adversely affects a child's educational performance but that is not included under the definition of deafness as listed above.

Mental retardation. Significantly subaverage general intellectual functioning existing concurrently with deficits in adaptive behavior and manifested during the developmental period that adversely affects a child's educational performance.

Multiple disabilities. A combination of impairments (such as mental retardation-blindness, or mental retardation-physical disabilities) that causes such severe educational problems that the child cannot be accommodated in a special education program solely for one of the impairments. The term does not include deaf-blindness.

See notes at end of table.

Table A-10. Definitions of disabilities—Continued

Orthopedic impairment. A severe orthopedic impairment that adversely affects educational performance. The term includes impairments such as amputation, absence of a limb, cerebral palsy, poliomyelitis, and bone tuberculosis.

Other health impairment. Having limited strength, vitality, or alertness due to chronic or acute health problems such as a heart condition, rheumatic fever, asthma, hemophilia, and leukemia, which adversely affect educational performance.[3]

Specific learning disability. A disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. This term includes such conditions as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. This term does not include children who have learning problems that are primarily the result of visual, hearing, or motor disabilities; mental retardation; or environmental, cultural or economic disadvantage.

Speech or language impairment. A communication disorder such as stuttering, impaired articulation, language impairment, or a voice impairment that adversely affects a child's educational performance.

Traumatic brain injury. An acquired injury to the brain caused by an external physical force, resulting in total or partial functional disability or psychosocial impairment, or both, that adversely affects a child's educational performance. The term applies to open or closed head injuries resulting in impairments in one or more areas, such as cognition; language; memory; attention; reasoning; abstract thinking; judgment; problem-solving; sensory, perceptual and motor abilities; psychosocial behavior; physical functions; information processing; and speech. The term does not apply to brain injuries that are congenital or degenerative, or brain injuries induced by birth trauma. As with autism, traumatic brain injury (TBI) was added as a separate category of disability in 1990 under P.L. 101-476.

Visual impairment, including blindness. An impairment in vision that, even with correction, adversely affects a child's educational performance. The term includes both partial sight and blindness.

[1] From Knoblauch and Sorenson (1998).

[2.] P.L. 105-17, the Individuals with Disabilities Education Act Amendments of 1997, changed "serious emotional disturbance" to "emotional disturbance." The change has no substantive or legal significance. It is intended strictly to eliminate any negative connotation of the term "serious."

[3.] OSEP guidelines indicate that "children with ADD, where ADD is a chronic or acute health problem resulting in limited alertness, may be considered disabled under Part B solely on the basis of this disorder under the 'other health impaired' category in situations where special education and related services are needed because of the ADD" (Davila 1991).

The exception to reliance on school or district category assignment involves students with deaf-blindness. District variation in assigning students with both hearing and visual impairments to the category of deaf-blindness results in many students with those dual disabilities being assigned to other primary disability categories, most often hearing impairment, visual impairment, and multiple disabilities. Because of these classification differences, national estimates suggest that there were 3,196 students with deaf-blindness who were ages 12 to 17 in 1999 (National Technical Assistance Center 1999), whereas the federal child count indicated that 681 were classified with deaf-blindness as their primary disability (Office of Special Education Programs 2001).

To describe the characteristics and experiences of the larger body of youth with deaf-blindness more accurately and precisely, students who were reported by parents or by schools or

Table A-11. Original primary disability category of youth assigned to deaf-blindness category for NLTS2 reporting purposes

Original Primary Disability Category	Number
Deaf-blindness	24
Visual impairment	46
Hearing impairment	43
Multiple disabilities	31
Orthopedic impairment	7
Mental retardation	6
Traumatic brain injury	4
Other health impairment	3
Speech/language impairment	1
Autism	1
Total	166

school districts[5] as having both a hearing and a visual impairment were assigned to the deaf-blindness category for purposes of NLTS2 reporting, regardless of the primary disability category assigned by the school or school district. This increased the number of youth with deaf-blindness for whom parent data were collected from 24 who were categorized by their school or district as having deaf-blindness as a primary disability to 166. The number of students reassigned to the deaf-blindness category and their original designation of primary disability are indicated in table A-11. Because there still are relatively few members of the deaf-blindness disability category, for purposes of multivariate analyses, they are included with the category of multiple disabilities.

Appendix A References

Davila, R.R. (1991). *Clarification of Policy to Address the Needs of Children With Attention Deficit Disorders Within General and/or Special Education*. Memorandum to Chief State School Officers. Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services.

Employment Policies Institute. (2002). *Measuring Poverty in America: Science or Politics?* Available at http://www.epionline.org/studies/epi_poverty_04-2002.pdf

Knoblauch, B., and Sorenson, B. (1998). *IDEA's Definition of Disabilities*. ERIC Digest #560. Retrieved October 10, 2005 from http://ericec.org/digests/e560.html

MacMillan, D.L., and Siperstein, G.N. (2002). Learning Disabilities as Operationally Defined by Schools. In R. Bradley, L. Danielson, and D.P. Hallahan (Eds.), *Identification of Learning Disabilities: Research to Practice*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

National Technical Assistance Center. (1999). *National Deaf-Blind Child Count Summary.* Monmouth, OR: Teaching Research Division, Western Oregon University.

Office of Special Education Programs. (2001). *Table AD1. Number of Students Age 14 and Older Exiting Special Education During the 1999-2000 School Year*. Available at http://www.ideadata.org/tables24th\ar_ad1.htm

U.S. Department of Education (2003). *Table AA7. Number of Children Served under IDEA Part B by Disability and Age during the 2002-03 School Year*. Retrieved October 10, 2005 from https://www.ideadata.org/tables26th%5Car_aa7.htm

[5] Some special schools and school districts reported secondary disabilities for students. So, for example, a student with visual impairment as his or her primary disability category also could have been reported as having a hearing impairment as a secondary disability.

Appendix B

Characteristics of Youth With Disabilities Participating in the Direct Assessment and Functional Rating

Appendix B. Characteristics of Youth With Disabilities Participating in the Direct Assessment and Functional Rating

It is important to understand the characteristics of the youth with disabilities who participated in the NLTS2 direct assessment and functional ratings as context for interpreting assessment results. As noted in chapter 1, school staff or parents reported their views on the feasibility and appropriateness of the direct assessment for individual youth; those for whom it was not considered feasible or appropriate had a functional rating completed by a teacher who was familiar with their abilities if the youth was still in school, or by a parent if he or she was no longer in school. Thus, the groups of youth with disabilities for whom the two forms of assessment were completed are intentionally different. But how different? And in what ways?

The following sections address such questions by describing the 94 percent of assessed youth with disabilities who participated in the direct assessment and the 6 percent for whom a functional rating was completed in terms of their

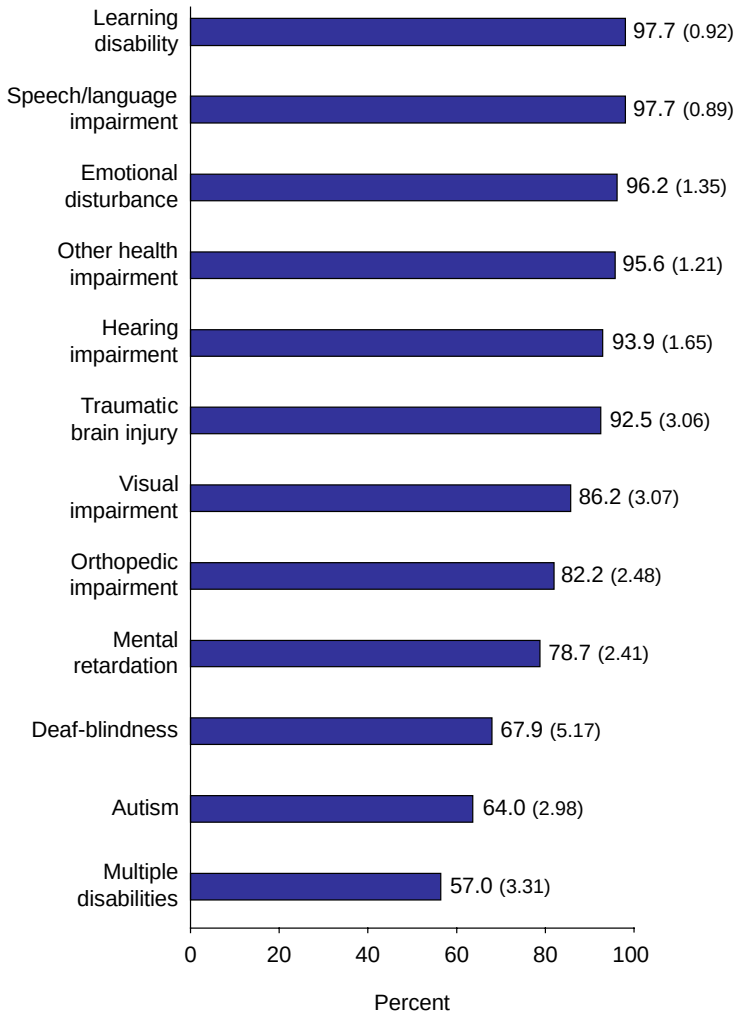
- disability category;[1]
- age at identification of and first service for a disability;
- functional abilities;
- schools and school programs;
- academic and social performance; and
- parents' expectations for the future of their adolescent children with disabilities.

The chapter does not address the demographic characteristics of youth with direct assessments and functional ratings because there are no statistically significant differences between the groups in age, gender, race/ethnicity, or household income. In the direct assessment group, 31 percent of youth are 16 years old, 49 percent are 17, and 21 percent are 18,[2] compared with 23 percent, 54 percent, and 23 percent among youth for whom a functional rating was completed. Almost two-thirds (65 percent) of direct assessment participants are male, as are 71 percent of the functional rating group. Among direct assessment participants, 62 percent are White, 21 percent are African American, and 14 percent are Hispanic, compared with 64 percent, 20 percent, and 13 percent of the functional rating group who are White, African American, and Hispanic, respectively. With regard to the three household income categories used in NLTS2 analyses ($25,000 or less, $25,001 to $50,000, and more than $50,000), 37 percent, 30 percent, and 33 percent of direct assessment participants fall in the three groups, respectively; comparable figures for youth for whom a functional rating was completed are 32 percent, 32 percent, and 36 percent.

[1] The terminology for classifying students who receive special education services that is used throughout this report is specified in federal regulations for the implementation of the Individuals with Disabilities Education Act Amendments of 1997 (P.L. 105-17), 111 Stat. 37-157 (Knoblauch and Sorenson 1998).

[2] This group also includes 10 youth from Wave 1 who had their 19th birthdays shortly before the assessment.

Figure B-1. Percentage participating in the direct assessment, by disability category



NOTE: Standard errors are in parentheses.

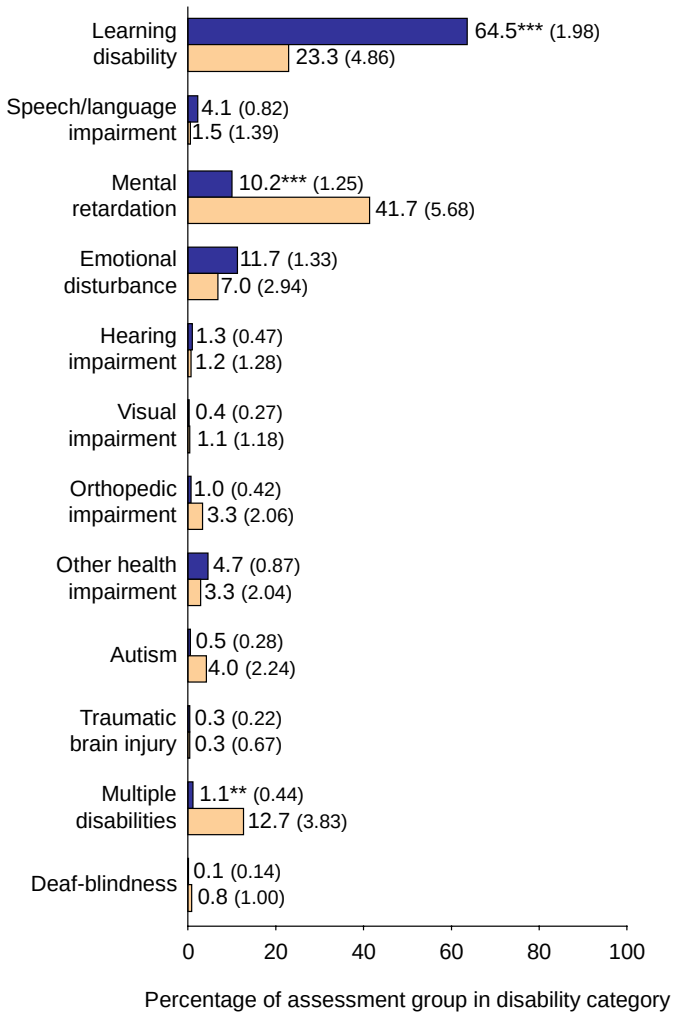
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), student rosters for participating local education agencies and special schools.

Disability Categories of Assessment Participants

Youth in the various federal special education disability categories differ markedly from each other in the extent to which they participated in the direct assessment or had a functional rating (figure B-1). Whereas almost all youth (98 percent) in the categories of learning disabilities and speech/language impairment participated in the direct assessment, participation rates were 86 percent and 82 percent for youth with visual or orthopedic impairments ($p < .01$ compared with learning disability and speech/language impairment). Almost 8 in 10 youth with mental retardation (79 percent) took part in the direct assessment, as did about two-thirds of youth with deaf-blindness or autism (68 percent and 64 percent) and 57 percent of youth with multiple disabilities.

These different rates of participation in the direct assessment vs. having a functional rating resulted in quite different disability category compositions of the two groups (figure B-2). Sixty-five percent of youth with disabilities who participated in the direct assessment are identified with learning disabilities, whereas they are 23 percent of youth with a functional rating ($p < .001$). Because a significant cognitive impairment was a frequent reason for recommending a student for participation in the functional rating rather than the direct assessment, it is not surprising that the proportion of the functional rating group that is comprised of youth with mental retardation is about four times higher than their proportion among direct assessment participants (42 percent vs. 10 percent, $p < .001$); nonetheless, 79 percent of youth with mental retardation took the direct assessment. Similarly, youth with multiple disabilities also comprise a larger proportion of the functional rating than the direct assessment group (13 percent vs. 1 percent, $p < .001$); they are the group with the highest likelihood of have a

Figure B-2. Disability category composition of direct assessment and functional rating groups



**p < .01, ***p < .001

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), student rosters for participating local education agencies and special schools.

functional rating completed (43 percent). Rates also are relatively high for youth with autism (36 percent) and deaf-blindness (32 percent).

Age at Disability Identification and First Receipt of Special Education Services

The difference in the disability category distribution of the direct assessment and functional rating groups would be expected to be accompanied by differences in the ages at which their disabilities first were identified and they received services for them. For example, because the functional rating group has a larger proportion of youth with mental retardation than the direct assessment group, a disability that can become apparent in early childhood, the average age at first identification of a disability would be expected to be lower for youth with a functional rating than with a direct assessment. In contrast, the latter group is heavily dominated by youth with learning disabilities, which generally are not identified until children reach school age.

These expectations are born out in parents' reports of the ages at which youth first were identified with and received special education services for a disability (table B-1). On average, youth for whom a functional rating was completed were identified as having a disability at about age 3, whereas the average age for youth who participated in the direct assessment is almost 6 (p < .001). Half of youth with a functional rating had their disability identified before their first birthday, compared with 14 percent of direct assessment participants (p < .001). In contrast, significantly larger shares of the direct assessment than the functional rating group had their disabilities identified between ages 5 and 7 (43 percent vs. 27 percent, p < .01) or between ages 8 and 10 (19 percent vs. 9 percent, p < .01). On average, youth with a functional rating also began receiving special education services at an earlier age than direct assessment participants (average ages of 6.5 and 8.3 respectively, p < .001). A large majority of those with a functional rating (79 percent) started receiving special education services between ages 5 and 7, a significantly larger proportion than their direct assessment peers (46 percent, p < .001). In contrast, significantly more of the direct

Table B-1. Age at identification of and first services for disabilities of youth in the direct assessment and functional rating groups

Youth's age	Youth with a: Direct assessment	Youth with a: Functional rating
	Percent / standard error	
Disability first identified at age:		
Birth-1	13.9 (1.60)	**49.9*** ** (6.31)
2-4	15.7 (1.68)	11.0 (3.95)
5-7	43.4 (2.29)	**27.3** ** (5.63)
8-10	19.1 (1.82)	**8.6** ** (3.53)
11 or older	8.0 (1.39)	3.2 (2.11)
Average age	5.7 (0.15)	**3.2*** ** (0.47)
Special education services in school first received at age:		
5-7	45.8 (2.31)	**78.6*** ** (4.30)
8-10	30.6 (2.13)	**16.0** ** (4.73)
11-13	19.5 (1.83)	**4.7*** ** (2.72)
14-17	4.3 (0.94)	**0.7*** (1.11)
Average age	8.3 (0.13)	**6.5*** ** (0.27)

*p < .05, **p < .01, ***p < .001

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

assessment than the functional rating group started receiving special education services between ages 8 and 10 (31 percent vs. 16 percent, $p < .01$) or between ages 11 and 13 (20 percent vs. 5 percent, $p < .001$).

Functional Abilities

Differences in the disability classifications of the direct assessment and functional rating groups are reflected in differences in their functional abilities (table B-2). Compared with their peers who participated in the direct assessment, significantly larger shares of youth with a functional rating have sensory, physical, or cognitive impairments. For example, according to parents, approximately twice as many youth in the functional rating as the direct assessment group have at least "some trouble" understanding speech (57 percent vs. 30 percent, $p < .05$) and/or communicating with others (63 percent vs. 29 percent, $p < .001$). Youth with a functional rating also are more likely than direct assessment participants to have at least some trouble using their arms or hands for gross motor (24 percent vs. 4 percent, $p < .001$) or fine motor activities (31 percent vs. 5 percent, $p < .001$), and using their legs or feet (28 percent vs. 6 percent, $p < .001$). Although about 1 in 10 youth in both assessment groups have some trouble with their hearing, more than twice as many of those with a functional rating are reported to have at least some trouble seeing (27 percent vs. 12 percent, $p < .05$).

A summary measure of functional abilities—the number of domains[3] that are affected by disability—shows that youth who participated in the direct assessment are reported to have an average of 1.3 domains affected by their disability, compared with 2.6 domains reported for youth for whom a functional rating was completed, ($p < .001$). More specifically,

[3] The domains included in this measure are vision, hearing, use of appendages, expressive language, receptive language, bidirectional communication, and general health.

Table B-2. Functional characteristics of youth with disabilities in the direct assessment and functional rating groups

Functional characteristics	Youth with a: Direct assessment	Youth with a: Functional rating
	Percent / standard error	
Youth has at least "some trouble":		
Seeing	12.4 (1.47)	**27.4*** (5.60)
Hearing	9.2 (1.28)	11.9 (4.00)
Understanding speech	30.3 (2.04)	**57.0**** (6.14)
Communicating with others	28.5 (2.01)	**62.5***** (6.02)
Using arms/hands for gross motor activities	4.1 (0.88)	**24.0***** (5.28)
Using arms/hands for fine motor activities	4.7 (0.94)	**30.7***** (5.70)
Using legs/feet	6.2 (1.07)	**27.5***** (5.53)
Number of functional domains affected by disability		
None	40.1 (2.03)	**18.3***** (4.45)
1 or 2	39.7 (2.02)	**22.7**** (4.82)
3 or 4	17.4 (1.57)	**47.4***** (5.75)
5 through 7	2.7 (0.67)	**11.8*** (3.98)
Youth's general health is excellent	41.9 (2.19)	**24.1**** (5.31)

*p < .05, **p < .01, ***p < .001

NOTE: The six functional domains in the scale of domains affected include vision, hearing, expressive language, receptive language, participation in bidirectional communication, use of arms/hands, and use of legs/feet.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

4 in 10 direct assessment participants are reported to be unaffected in the sensory, physical, or communication functions included in this measure; almost one in five youth with a functional rating also are reported to be unaffected in these functional areas (p < .001 compared with direct assessment participants). Conversely, almost half of youth with a functional rating are reported to be affected in three or four of the sensory, physical, or communication domains, compared with fewer than one in five direct assessment peers (47 percent vs. 17 percent, p < .001). Furthermore, youth with disabilities who participated in the direct assessment appear to be healthier than their peers in the functional rating group; parents of 42 percent of direct assessment participants report their sons and daughters are in "excellent" health, compared with 24 percent of youth with a functional rating (p < .01).

The functional characteristics of youth in direct assessment and functional ratings groups play out in their ability to perform a variety of activities of daily living (table B-3). NTLS2 analyses include scales[4] that distinguish the level of functioning of youth with disabilities with regard to self-care (independently taking care of feeding and dressing); functional cognitive abilities (processing information that is important to daily living, such as counting change, telling time, or using a telephone); and social competence (interacting appropriately with others). On the self-care skills scale, parents of 95 percent of direct assessment participants report their sons and daughters are highly skilled, compared with 60 percent of those with a functional rating (p < .001). Conversely, 17 percent of youth with a functional rating are reported to have low self-care skills, significantly more than the 1 percent of direct assessment participants reported to have low skills (p < .001).

More than 50 percentage points separate high ratings on the functional cognitive skills scale for the two assessment groups; parents of 64 percent of direct assessment participants report their

[4] Chapter 4 includes details on the construction of these scales.

Table B-3. Daily living scale scores of youth with disabilities in the direct assessment and functional rating groups

Functional skills scale scores	Youth with a: Direct assessment	Youth with a: Functional rating
	Percent / standard error	
Self-care skills scale score:		
Low (2-4)	0.4 (0.27)	**17.0*** (4.40)
Medium (5-7)	5.1 (0.92)	**22.8*** (4.92)
High (8)	94.5 (0.98)	**60.2*** (5.74)
Mean	7.9 (0.02)	**6.7*** (0.24)
Functional cognitive skills scale score:		
Low (4-7)	1.5 (0.54)	**33.4*** (5.99)
Medium (8-12)	34.24 (2.11)	45.4 (6.23)
High (13-16)	64.3 (2.13)	**10.2*** (4.93)
Mean	13.9 (0.10)	**9.6*** (0.52)
Social skills scale score:		
Low (0-10)	17.1 (1.60)	21.7 (5.10)
Medium (11-16)	61.2 (2.08)	55.6 (6.14)
High (17-22)	21.7 (1.76)	22.7 (5.18)
Mean	13.7 (0.16)	13.2 (0.49)

***$p < .001$

NOTE: The category "medium" is omitted from the table.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

sons and daughters display high functional cognitive skills, compared with 10 percent of youth with a functional rating ($p < .001$). Moreover, low functional cognitive scores are significantly more common for the functional rating than the direct assessment group (33 percent vs. 2 percent, $p < .001$), consistent with the larger proportion of youth in the mental retardation category among those with a functional rating. Despite these differences in self-care and functional cognitive skill levels between youth in the two groups, there are no discernable differences in the proportions of youth receiving low or high scores on the social skills scale.

School and Program Characteristics

The individualized education program (IEP) to be provided each student receiving special education services under the Individuals with Disabilities Education Improvement Act of 2004 (IDEA) Section 614(3) would be expected to reflect the differences in disabilities and functioning between the direct assessment and functional rating groups in differences in the nature of the educational services provided and in their instructional settings, for example. Several such differences are apparent (table B-4). A small share of youth with disabilities in both groups attend or have attended a special school that serves only students with disabilities, although the experience is about four times more common for youth with a functional rating than with a direct assessment (13 percent vs. 3 percent, $p < .05$).

Overall, most youth with disabilities take classes in both general and special education settings; however, participation differs considerably by assessment group. On average, direct assessment participants take 63 percent of their classes in general education settings, more than twice the rate of those with a functional rating (29 percent, $p < .001$). In fact, half of the functional rating group take no general education classes, compared with 1 in 10 direct assessment participants ($p < .001$). Further, more than twice as many in the direct assessment

Table B-4. School program characteristics of youth with disabilities who participated in the direct assessment and functional ratings

School program characteristics	Youth participated in: Direct assessment (Percent / standard error)	Youth participated in: Functional rating (Percent / standard error)
Youth attends or attended a special school serving only students with disabilities	3.2 (0.78)	**12.8*** (4.10)
Percentage of classes in general education		
None	10.0 (1.58)	**50.3***** (7.49)
1% to 25%	9.6 (1.55)	13.3 (5.09)
26% to 50%	19.5 (2.08)	14.2 (5.22)
51% to 75%	16.4 (1.95)	**2.6***** (2.38)
More than 75%	44.5 (2.61)	**19.6***** (5.94)
Average percentage of classes in taken in general education	62.9 (1.80)	**28.9***** (5.74)
Youth takes		
Any academic class	98.9 (0.56)	**85.2*** (5.32)
Any vocational class	62.4 (2.55)	**84.1***** (5.48)
Any nonacademic, nonvocational class	83.1 (1.97)	**95.7***** (3.03)

*$p < .05$, ***$p < .001$

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001, and Student's School Program Survey, 2002.

group as the functional rating group take three-quarters or more of their classes in general education settings (45 percent vs. 20 percent, $p < .001$).

NLTS2 findings demonstrate that the course-taking pattern of youth with disabilities who participated in the direct assessment is quite similar to that of secondary students with disabilities overall (National Center for Education Statistics 2001) but, as with instructional setting, the pattern differs substantially from their peers for whom functional ratings were completed. On average, almost all direct assessment participants (99 percent) take at least one academic class (i.e., language arts, mathematics, science, social studies, or a foreign language) during a semester, compared with 85 percent of youth with a functional rating ($p < .05$). Conversely, significantly larger shares of youth with a functional rating than with a direct assessment take at least one vocational class (84 percent vs. 62 percent, $p < .001$) and at least one nonacademic class (i.e., physical education, fine arts, study skills, life skills, or social skills) (96 percent vs. 83 percent, $p < .001$).

Differences in parents' reports of the related services and supports provided students with disabilities as part of their school programs also differ in some cases between youth who participated in the direct assessment rather those with a functional rating (table B-5). Youth with a functional rating are significantly more likely to be reported to receive services and supports of various kinds than their peers who participated in the direct assessment. The greatest differences are in the proportions of youth reported to receive speech/ language pathology services (51 percent vs. 20 percent), occupational or physical therapy (43 percent vs. 12 percent and 36 percent vs. 7 percent), or transportation services (41 percent vs. 9 percent). Youth with a functional rating also are more likely than their direct assessment peers to be reported to receive medical diagnostic services (40 percent vs. 22 percent), respite care (15 percent vs. 2 percent), SSI benefits (29 percent vs. 14 percent), or nursing care (8 percent vs. 1 percent). On the other hand, youth with disabilities in both assessment groups are equally likely to be reported to receive audiology, vocational, social work, orientation/mobility, or assistive technology services or personal or psychological counseling or to have a reader/interpreter or tutor.

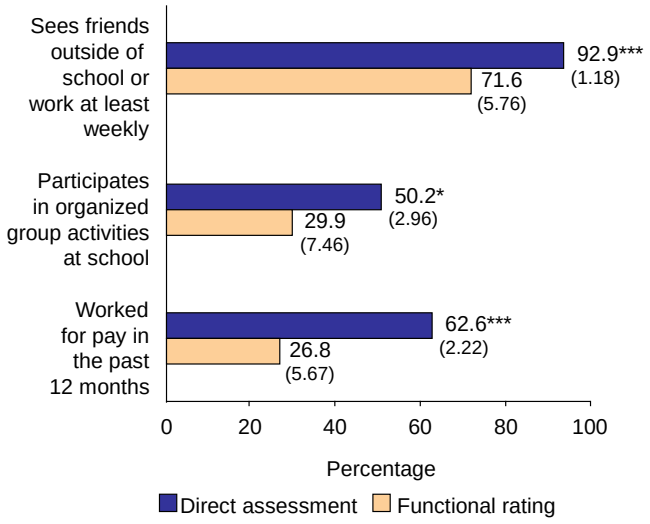
Table B-5. Services and supports received by youth with disabilities in the direct assessment and functional rating groups

Services and supports received	Youth with a: Direct assessment	Youth with a: Functional rating
	Percent / standard error	
In the past 12 months, youth received:		
Speech/language pathology services	19.7 (1.80)	**51.4*** ** (6.34)
Occupational therapy	12.4 (1.49)	**43.1*** ** (6.28)
Transportation services	8.5 (1.26)	**41.0*** ** (6.21)
Physical therapy	6.5 (1.11)	**35.6*** ** (6.05)
Medical diagnostic services	21.9 (1.83)	**40.1** ** (6.20)
Respite care	1.6 (0.56)	**15.3** ** (4.53)
SSI	13.8 (1.56)	**28.9*** (5.76)
Nursing care	1.1 (0.47)	**8.3*** (3.47)

*p < .05, **p < .01, ***p < .001

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

Figure B-3. Social activities of youth with disabilities who participated in the direct assessment and functional ratings



*p < .05, ***p < .001

NOTE: Standard errors are in parentheses.

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

Academic and Social Outcomes and Experiences

Despite their differences in disability, functioning, and school programs, youth in the two assessment groups do not differ on indicators of academic performance, such as course grades or a history of grade retention. About 3 in 10 youth in both groups are reported by parents to get "mostly As or As and Bs" (30 percent and 27 percent of direct assessment and functional rating groups, respectively), and 40 percent and 37 percent of direct assessment and functional rating groups, respectively, have been held back one or more grades. Differences in the social experiences of the two groups are apparent, however (figure B-3). The majority (72 percent) of youth with a functional rating are reported to be socially engaged, seeing friends outside of school or work at least weekly; however, this is a significantly smaller group than the socially active youth participating in the direct assessment (93 percent, p < .001). Similarly, youth with a functional rating are less likely than direct assessment participants to take part in organized group activities at school (30 percent vs. 50 percent, p < .05), and they are less than half as likely to have had some paid work experience in the preceding year (27 percent vs. 63 percent, p < .001).

Table B-6. Parents' expectations for youth with disabilities who participated in the direct assessment and functional ratings

Parents' expectations	Youth participated in: Direct assessment	Youth participated in: Functional rating
	Percent / standard error	
Youth expected to:		
Graduate from high school with a regular diploma		
Definitely will	57.3 (2.41)	**25.6*** ** (5.67)
Probably will	29.4 (2.22)	29.4 (5.92)
Definitely/probably won't	13.3 (1.65)	**45.1*** ** (6.47)
Attend postsecondary school		
Definitely will	31.3 (2.15)	**8.7*** ** (4.12)
Probably will	38.6 (2.26)	27.3 (6.51)
Definitely/probably won't	30.1 (2.13)	**64.1*** ** (7.02)
Get a paid job		
Definitely will	85.9 (1.82)	**51.2*** ** (6.64)
Probably will	12.4 (1.73)	22.5 (5.55)
Definitely/probably won't	1.7 (0.68)	**26.3*** ** (5.84)
Get a driver's license		
Definitely will	52.8 (2.71)	**17.1*** ** (4.22)
Probably will	34.1 (2.57)	**14.8*** ** (3.97)
Definitely/probably won't	13.1 (1.83)	**68.2*** ** (5.22)
Live independently		
Definitely will	56.8 (2.26)	**28.4*** ** (5.8)
Probably will	30.9 (2.11)	**14.7** ** (4.52)
Definitely/probably won't	12.3 (1.50)	**57.0*** ** (6.32)
Be financially self-supporting		
Definitely will	48.6 (2.30)	**26.9*** ** (5.70)
Probably will	36.8 (2.22)	**11.9*** ** (4.16)
Definitely/probably won't	14.4 (1.68)	**61.3*** ** (6.23)

**p < .01, ***p < .001

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), Wave 1 parent interviews, 2001.

Parents' Expectations for Their Children's Futures

The marked differences in the characteristics and attributes of high school youth with disabilities who participate in direct assessment or who have a functional ratings are illustrated by parents' visions of their children's futures. Parents reported their expectations for their sons and daughters with regard to high school graduation, postsecondary school attendance, securing paid employment, living independently, and becoming financially self-supporting (table B-6).[5]

More than half of direct assessment youth (57 percent) are expected "definitely" to graduate from high school with a regular diploma (table B-6). In contrast, parents of one-quarter of youth with a functional rating (26 percent) expect their son or daughter "definitely" to receive such a diploma ($p < .001$). Consistent with this difference, parents of fewer youth in the functional rating than direct assessment group expect their child to attend postsecondary school (9 percent vs. 31 percent are expected "definitely" to attend postsecondary school, $p < .001$).

Although about half (51 percent) of youth with a functional rating have parents who expect they "definitely" will get a paid job, this is significantly fewer than the

[5] Possible responses were "definitely will," "probably will," "probably won't," and "definitely won't."

86 percent of direct assessment participants who are expected to secure paid employment ($p < .001$). Residential and financial independence are considered definite possibilities for about half of direct assessment participants (57 percent and 49 percent, respectively), compared with just over one-quarter of youth with a functional rating (28 percent and 27 percent respectively, $p < .001$ for both comparisons).

Summary

Demographically, youth with disabilities who participated in the direct assessment and those for whom a functional rating was completed are similar, but they are markedly different in the disability categories they represent, the average ages at which they first were identified as having and received services for a disability, and in their functional abilities and health. Differences in the characteristics of the school programs and the types of related services and supports received by youth in the direct assessment and functional rating groups include greater special education and service participation among youth with a functional rating compared with direct assessment peers. Differences in the social experiences of the two groups involve greater social engagement among direct assessment participants, the group for whom higher parental expectations for the future also are reported.

Appendix B References

Knoblauch, B., and Sorenson, B. (1998). *IDEA's Definition of Disabilities*. ERIC Digest #560. Retrieved October 10, 2005 from http://ericec.org/digests/e560.html

National Center for Education Statistics. (2001). *The 1998 High School Transcript Study Tabulations: Comparative Data on Credits Earned and Demographics for 1998, 1994, 1900, 1987, and 1982 High School Graduates*. Washington, DC: Author. Retrieved October 10, 2005 from http://nces.ed.gov/pubs2001/2001498_1.pdf

Appendix C

Supplemental Analysis Results

Table C-1. Performance of youth with disabilities on Woodcock-Johnson III subtests, by disability category

Score categories	Learning disability	Speech/ language impair-ment	Mental retar-dation	Emo-tional distur-bance	Hearing impair-ment	Visual impair-ment	Ortho-pedic impair-ment	Other Health impair-ment	Autism	Trau-matic brain injury	Multiple disabili-ties	Deaf-blind-ness
	Percent / standard error											
Synonym/antonym subtest												
More than 2 standard deviations below the mean	6.9 (1.56)	8.4 (1.67)	58.7 (3.29)	7.9 (1.94)	22.4 (3.00)	8.5 (2.75)	13.9 (2.39)	5.1 (1.32)	36.0 (3.89)	17.5 (4.65)	45.6 (4.63)	34.1 (6.39)
From 1 to 2 standard deviations below the mean	25.9 (2.70)	23.1 (2.53)	30.3 (3.07)	16.1 (2.64)	25.6 (3.14)	16.4 (3.64)	24.8 (2.98)	15.3 (2.15)	16.6 (3.02)	29.8 (5.59)	21.3 (3.81)	27.2 (6.00)
0 to 1 standard deviations below the mean	47.1 (3.08)	45.2 (2.99)	10.8 (2.07)	38.5 (3.49)	33.6 (3.40)	31.7 (4.58)	35.4 (3.31)	41.1 (2.94)	17.1 (3.05)	41.8 (6.03)	24.8 (4.02)	28.8 (6.10)
Above the mean	20.0 (2.47)	23.3 (2.54)	0.3 (0.37)	37.5 (3.47)	18.3 (2.78)	43.4 (4.88)	26.0 (3.03)	38.6 (2.91)	30.2 (3.72)	10.9 (3.81)	8.3 (2.57)	9.9 (4.03)
Mean standard score	89.5 (0.81)	89.9 (0.93)	65.3 (1.06)	93.4 (1.12)	84.1 (1.44)	94.0 (1.89)	88.2 (1.23)	95.0 (0.86)	81.3 (2.16)	83.7 (1.95)	71.6 (2.11)	75.5 (2.88)
Mathematics calculation subtest												
More than 2 standard deviations below the mean	23.0 (2.60)	15.0 (2.15)	65.0 (3.24)	20.8 (2.92)	17.6 (2.72)	16.0 (3.64)	26.6 (3.07)	19.2 (2.36)	39.0 (4.03)	36.5 (5.97)	59.0 (4.75)	34.6 (6.52)
From 1 to 2 standard deviations below the mean	21.7 (2.55)	17.3 (2.28)	26.4 (2.99)	22.7 (3.02)	14.8 (2.54)	14.9 (3.53)	22.8 (2.92)	17.5 (2.28)	14.3 (2.89)	14.8 (4.40)	14.1 (3.36)	22.2 (5.70)
0 to 1 standard deviations below the mean	29.7 (2.82)	31.4 (2.80)	7.2 (1.76)	34.2 (3.42)	27.7 (3.20)	24.0 (4.24)	27.2 (3.09)	35.8 (2.88)	17.2 (3.12)	27.0 (5.50)	15.4 (3.49)	23.3 (5.80)
Above the mean	25.6 (2.70)	36.2 (2.90)	1.4 (0.80)	22.2 (2.99)	39.8 (3.50)	45.1 (4.94)	23.4 (2.94)	27.5 (2.68)	29.5 (3.77)	21.6 (5.10)	11.5 (3.08)	19.9 (5.48)
Mean standard score	86.1 (1.09)	91.7 (1.14)	61.4 (1.43)	86.2 (1.22)	91.5 (1.42)	92.2 (2.41)	82.6 (1.64)	88.2 (1.07)	80.2 (2.39)	80.0 (2.65)	65.6 (2.89)	77.7 (3.39)
Applied problems subtest												
More than 2 standard deviations below the mean	8.3 (1.70)	11.9 (1.96)	56.5 (3.32)	9.4 (2.10)	17.3 (2.70)	16.4 (4.14)	27.5 (3.10)	10.6 (1.85)	43.6 (4.03)	26.0 (5.43)	56.3 (4.66)	39.4 (7.24)
From 1 to 2 standard deviations below the mean	20.8 (2.50)	20.5 (2.44)	30.0 (3.07)	20.6 (2.91)	21.6 (2.93)	13.1 (3.77)	21.2 (2.83)	18.1 (2.31)	14.5 (2.86)	23.9 (5.28)	17.4 (3.56)	23.5 (6.28)
0 to 1 standard deviations below the mean	56.3 (3.06)	45.4 (3.01)	12.9 (2.24)	54.4 (3.58)	47.6 (3.56)	44.9 (5.56)	40.2 (3.40)	54.3 (2.99)	24.8 (3.51)	40.0 (6.07)	20.1 (3.77)	26.4 (6.53)
Above the mean	14.6 (2.18)	22.1 (2.51)	0.6 (0.52)	15.6 (2.61)	13.4 (2.43)	25.6 (4.87)	11.1 (2.18)	16.9 (2.25)	17.1 (3.06)	10.1 (3.73)	6.1 (2.25)	10.8 (4.60)
Mean standard score	88.3 (0.77)	87.9 (0.98)	63.4 (1.31)	88.2 (1.06)	83.9 (1.32)	87.6 (2.23)	79.8 (1.44)	88.4 (0.85)	71.2 (2.36)	80.6 (2.23)	62.9 (2.42)	72.8 (3.45)

See notes at end of table.

Table C-1. Performance of youth with disabilities on Woodcock-Johnson III subtests, by disability category—Continued

Score categories	Learning disability	Speech/ language impair-ment	Mental retar-dation	Emo-tional distur-bance	Hearing impair-ment	Visual impair-ment	Ortho-pedic impair-ment	Other Health impair-ment	Autism	Trau-matic brain injury	Multiple disabili-ties	Deaf-blind-ness
	Percent / standard error											
Social studies subtest												
More than 2 standard deviations below the mean	8.2 (1.69)	13.1 (2.03)	55.9 (3.32)	11.5 (2.29)	26.5 (3.18)	13.0 (3.31)	16.6 (2.57)	8.6 (1.68)	44.8 (4.03)	22.7 (5.12)	50.2 (4.65)	37.3 (6.52)
From 1 to 2 standard deviations below the mean	41.4 (3.04)	36.4 (2.89)	39.0 (3.26)	29.4 (3.27)	27.8 (3.22)	21.8 (4.07)	33.4 (3.26)	37.2 (2.89)	17.3 (3.07)	40.6 (6.01)	31.5 (4.32)	29.3 (6.13)
0 to 1 standard deviations below the mean	32.4 (2.89)	34.5 (2.86)	4.7 (1.41)	36.9 (3.46)	31.6 (3.35)	31.1 (4.56)	34.0 (3.27)	35.3 (2.86)	18.4 (3.14)	25.8 (5.35)	13.5 (3.18)	26.4 (5.94)
Above the mean	18.0 (2.37)	16.0 (2.20)	0.3 (0.37)	22.2 (2.98)	14.1 (2.50)	34.1 (4.67)	15.9 (2.53)	19.0 (2.35)	19.5 (3.21)	10.8 (3.80)	4.8 (1.99)	6.9 (3.42)
Mean standard score	86.6 (0.90)	85.6 (1.01)	65.1 (0.98)	87.8 (1.23)	80.5 (1.57)	88.4 (2.28)	84.3 (1.27)	87.7 (0.99)	73.9 (2.42)	79.1 (2.47)	67.5 (1.95)	73.8 (3.03)
Science subtest												
More than 2 standard deviations below the mean	9.5 (1.81)	13.7 (2.07)	48.2 (3.34)	8.5 (2.00)	35.0 (3.43)	10.3 (2.99)	16.3 (2.55)	7.0 (1.53)	38.1 (3.94)	20.8 (4.96)	45.8 (4.63)	42.1 (6.65)
From 1 to 2 standard deviations below the mean	30.4 (2.84)	36.4 (2.89)	37.7 (3.24)	24.6 (3.09)	26.5 (3.18)	22.9 (4.14)	33.6 (3.26)	26.3 (2.63)	21.4 (3.33)	32.3 (5.72)	29.8 (4.25)	28.5 (6.08)
0 to 1 standard deviations below the mean	41.0 (3.03)	30.4 (2.76)	13.5 (2.28)	42.4 (3.55)	24.1 (3.08)	41.2 (4.85)	35.1 (3.30)	42.9 (2.96)	18.6 (3.16)	30.0 (5.60)	19.3 (3.67)	17.9 (5.17)
Above the mean	19.2 (2.43)	19.5 (2.38)	0.6 (0.52)	24.5 (3.09)	14.4 (2.53)	25.6 (4.30)	15.1 (2.47)	23.9 (2.55)	21.9 (3.35)	16.9 (4.58)	5.1 (2.05)	11.5 (4.30)
Mean standard score	87.6 (0.91)	85.6 (1.02)	67.0 (1.15)	89.3 (1.25)	75.4 (1.77)	88.8 (2.05)	83.4 (1.28)	90.0 (0.94)	75.7 (2.21)	80.0 (2.74)	69.3 (2.04)	68.4 (3.65)

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Special Education Research, National Longitudinal Transition Study-2 (NLTS2), direct assessments, 2002 and 2004.

Table C-2. Factors associated with variation in the use of synonyms and antonyms, applied problem solving skills, and social studies content knowledge of youth with disabilities

Independent variables	Change in the following scores for each unit change in the independent variable: Synonyms and antonyms	Applied problem solving	Social studies content knowledge
Disability characteristics			
Visual impairment (vs. learning disability)	**+6.29****	+4.81	+4.34
Emotional disturbance (vs. learning disability)	**+5.67*****	+0.66	+3.67
Other health impairment (vs. learning disability)	**+5.14*****	-0.96	-1.53
Orthopedic impairment (vs. learning disability)	+2.10	-2.07	+0.28
Speech/language impairment (vs. learning disability)	-0.40	-0.83	-2.33
Hearing impairment (vs. learning disability)	-2.76	-1.73	**-5.63****
Traumatic brain injury (vs. learning disability)	+0.05	**+4.98****	**-5.90****
Autism (vs. learning disability)	**-3.82***	**-10.14*****	**-10. 27*****
Multiple disabilities/deaf-blindness (vs. learning disability)	**-8.80*****	**-11.90*****	**-9.43****
Mental retardation (vs. learning disability)	**-13.50*****	**-11.54*****	**-10.27*****
Age at identification of disability	**+0.26***	**+0.41*****	+0.18
Functioning			
Number of domains affected	**-1.00*****	-0.17	**-0.64***
Functional cognitive skills	**+1.52*****	**+2.54*****	**+1.81*****
Social skills	**-1.13****	-0.65	-0.48
Persistence	-0.60	-0.29	-0.84
Demographics			
Age	-0.55	-0.37	-0.96
Gender (boys vs. girls)	+1.27	**+2.71*****	**+3.98*****
African American (vs. White)	**-7.06*****	**-7.70*****	**-8.86*****
Hispanic (vs. White)	**-8.93*****	**-6.15*****	**-7.94*****
Other or multiple race/ethnicity (vs. White)	**-6.97*****	**-7.48*****	**-7.89*****
Household characteristics			
Expectations for postsecondary education	**+5.57*****	**+3.70*****	**+4.65*****
Low income (vs. moderate income)	**-4.50*****	**-2.47****	**-3.67*****
High income (vs. moderate income)	+1.57	+1.18	+1.60
Family involvement at home scale score	-0.44	-0.26	-0.09
Family involvement at school scale score	+0.06	+0.19	+0.14
School experiences			
Ever retained at grade level	-1.23	+0.69	-1.25
Overall grades	-0.35	-0.19	-0.12
Had any suspensions, expulsions, or disciplinary actions in the current school year	-0.97	+0.57	+0.60
Days absent per month	-0.21	**-0.29***	-0.19
School mobility other than for grade level changes	+0.32	+0.09	+0.26

See notes at end of table.

Table C-2. Factors associated with variation in the use of synonyms and antonyms, applied problem solving skills, and social studies content knowledge of youth with disabilities—Continued

Independent variables	Change in the following scores for each unit change in the independent variable: Synonyms and antonyms	Applied problem solving	Social studies content knowledge
Accommodations			
Breaks or multiple sessions	**-4.16*** **	**-3.62** **	**-4.97*** **
ASL or ASL interpreter	**-6.93*** **	**-5.34*** **	**-4.83** **
Braille or large print materials	**+4.65***	+1.56	+0.36
Special furniture or lighting	**+4.61** **	-0.29	+2.76
Calculator		**-2.65** **	

Multivariate analyses require that for categorical variables, such as disability category, each category be compared with another specified category. Learning disability was chosen as the category against which to compare the relationships for other disability categories because it is the largest category and, therefore, most closely resembles the characteristics of youth with disabilities as a whole. Similarly, White youth are the group against which to compare results for other racial/ethnic groups because it is the largest group.

*p < .05, **p < .01, ***p < .001. Note that approximately 100 relationships are included in this table; about 5 would be expected to occur by chance.

Table reads: The applied problem solving standard score of youth with autism is 10.1 points lower than the score of youth with learning disabilities, other factors being equal. The social science score of youth from low-income households is 3.7 points lower than the scores of youth from moderate-income households, independent of other factors.


DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA