Appendix A
SEELS SAMPLING, DATA COLLECTION, AND ANALYSIS PROCEDURES: WAVE 1

This appendix describes several aspects of the SEELS methodology relevant to the Wave 1 parent interview/survey, including:

- Sampling local education agencies (LEAs), schools, and students
- Parent interview and survey procedures and response rates
- Weighting of the parent interview/survey data
- Estimating and using standard errors
- Calculating statistical significance
- Measurement issues.

SEELS Sample Overview

The SEELS sample was constructed in two stages. A sample of 1,124 LEAs was selected randomly from the universe of approximately 14,000 LEAs that serve students receiving special education in at least one grade from first to seventh grade.[1] These districts and 77 state-supported special schools that serve primarily students with hearing and vision impairments and multiple disabilities were invited to participate in the study. A total of 245 LEAs and 32 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 receiving special education from each LEA[2] and special school was stratified by disability category. Students then were randomly selected from each disability category. Sampling fractions were calculated that would produce enough students in each category so that, in the final study year, we can generalize to most categories individually with an acceptable level of precision, accounting for attrition and for response rates to both the parent interview and the direct assessment. A total of 11,512 students were selected and eligible to participate in the SEELS parent interview/survey sample.

Details of the LEA and student samples are provided below.

The SEELS LEA Sample

Defining the Universe of LEAs

The SEELS 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, such as correctional facilities;

[1] The 1999 Quality Education Data, Inc. (QED) database was used to construct the sampling frame.

[2] 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.

LEAs from U.S. territories; and LEAs with 10 or fewer students in the SEELS age range, which would be unlikely to have students with disabilities.

The public school universe data file maintained by Quality Education Data (QED, 1998) 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 (1997). Correcting for errors and duplications resulted in a master list of 13,426 LEAs that were expected to have at least one student receiving special education in the appropriate age range. These comprised the SEELS LEA sampling frame.

Stratification

The SEELS LEA sample was stratified to increase the precision of estimates by eliminating between-strata variance, 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 SEELS 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, as well as 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 include Northeast, Southeast, Midwest, and West).

LEA size (student enrollment). LEAs vary considerably by size, the most useful available measure of which is pupil 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 enrollment greater than 17,411 in grades 1 through 7)
- **Large** (estimated enrollment from 4,707 to 17,411 in grades 1 through 7)
- **Medium** (estimated enrollment from 1,548 to 4,706 in grades 1 through 7)
- **Small** (estimated enrollment between 10 and 1,547 in grades 1 through 7).

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) is a well-accepted measure. The distribution of Orshansky index scores was organized into four categories of LEA/community wealth, each containing approximately 25% of the student population in grades 2 through 7:

- High (0% to 12% Orshansky)
- Medium (13% to 34% Orshansky)

- Low (35% to 45% Orshansky)
- Very low (over 45% 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, 297 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 1,124 LEAs was invited to participate, from which 297 participating LEAs might be recruited. A total of 245 LEAs actually provided students for the sample. Although the sample of LEAs was somewhat smaller than anticipated, analyses of the characteristics of the LEA sample, in weighted and unweighted form, on the sampling variables of region, LEA size, and LEA wealth confirmed that that the weighted LEA sample closely resembled the LEA universe with respect to those variables, thus yielding an initial sample of LEAs that was representative of the nation.

In addition to ensuring that the LEA sample matched the universe of LEAs on variables used in the sampling, it was important to ascertain whether this stratified random sampling approach resulted in skewed distributions on relevant variables not included in the stratification scheme. Two variables from the QED database were chosen to compare the "fit" between the first-stage sample and the population: the LEA's metropolitan status and its proportion of minority students. Analyses revealed that the fit between the weighted LEA sample and the LEA universe was quite good.

The SEELS Student Sample

Determining the size of the SEELS student sample took into account the duration of the study, desired levels of precision, and assumptions regarding attrition and response rates. We calculated that approximately three students would need to be sampled for each one student who would have both a parent/guardian interview and a direct assessment in Wave 3 of SEELS data collection.

The SEELS sample design emphasizes the need to generate fairly precise estimates of proportions and ratios for students receiving special education as a whole and for each of the 12 special education disability categories. A level of precision for standard errors of 3.6% was considered sufficient for study purposes. Thus, by sampling 1,150 students per disability category (except for TBI and deaf-blind) in year 1, we estimated there would be 388 students per category with both a parent interview and a direct assessment in year 5. Assuming a 50% sampling efficiency (which will tend to be exceeded for almost all disability categories), the 388 students would achieve a standard error of estimate of 3.6%. In addition, all students with traumatic brain injury or with deaf-blindness in participating LEAs and special schools were selected

SRI contacted LEAs and special schools to obtain their agreement to participate in the study and request rosters of students receiving special education who were between the ages of 6 and

12 on September 1, 1999 and in at least first 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 to SRI).

After estimating the number of students receiving special education in the SEELS age range, the appropriate fraction of students in each category was selected randomly from each LEA. In addition, from the state-supported special schools, 100% of students with deaf-blindness, 50% of students with visual impairments, and 15% of those with hearing impairments were sampled. In cases in which more than one child in a family was included on a roster, only one child was eligible to be selected. LEAs and special schools were notified of the students selected and contact information for their parents/guardians was requested.

Parent Interview/Questionnaire

The data source for the findings reported here was parents/guardians of SEELS sample members, who were interviewed by telephone or through a questionnaire sent through the mail. The SEELS conceptual framework holds that a child'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 student outcomes. Parents/guardians are the most knowledgeable about these aspects of students' lives.

Matches of names, addresses, and telephone numbers of SEELS parents with existing national locator databases were conducted to maximize the completeness and accuracy of contact information and subsequent response rates. Letters were sent to parents to notify them that their child had been selected for SEELS and that we would be attempting to contact them by telephone. A toll-free telephone number was included in the letter for parents to call in to be interviewed if they could not be reached by telephone or to make an appointment for the interview at a convenient time. If the computer match of contact information, letters mailed to parents, and attempted telephone interviews revealed that neither a working telephone number or accurate address was available for a student, that student was considered ineligible for the study and removed from the sample. Students who had no adult in the household who spoke either English or Spanish were ineligible for the study.

Computer-assisted telephone interviewing (CATI) was used for parent interviews, which were conducted between from mid-July through early December 2000. Interviews were conducted in both English and Spanish.

All parents with an accurate address who could not be reached by telephone were mailed a self-administered questionnaire in a period that extended from December 2000 through March 2001. The questionnaire contained a subset of key items from the telephone interview. Exhibit A-1 reports the responses to the telephone and mail surveys.

---

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

Exhibit A-1
RESPONSE RATES FOR PARENT/GUARDIAN TELEPHONE INTERVIEW AND MAIL SURVEY

	Number	Percentage
Total eligible sample	11,512	100.00
Respondents		
Completed telephone interview	8,624	74.9
Partial telephone interview completed	132	1.2
Complete mail questionnaire	1,068	9.3
Total respondents	9,824	85.3
Nonrespondents		
Refused	455	4.0
Language barrier	156	1.4
No response	1,077	9.4

Overall, 93% of respondents reported that they were parents of sample members (biological, adoptive, or step), and almost 1% were foster parents. Four percent were relatives other than parents, 1% were nonrelative legal guardians, and fewer than 1% reported other relationships to sample members.

Direct Assessment

Several of the dependent variables that are the subject of this report come from the SEELS direct assessment. Study designers felt that for students at this age level, some outcomes could only be assessed through a face to face assessment. The assessment was designed to measure a range of topics from academics to self concept and provide a mechanism to include the student "voice" in study data. The resulting standard assessment battery draws on the following published instruments to achieve these goals:

- Rapid letter naming and segmenting from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgeson, & Rashotte, 1999).
- Oral reading fluency from the Standard Reading Passages (Marston & Deno, 1986).
- Letter word identification (research edition) from the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001).
- Passage comprehension (research edition) from the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001).
- Mathematics calculation (research edition) from the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001).
- Math problem solving (research edition) from the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001).
- Student self concept scale (Gresham & Elliott, 1991).
- Student attitude measure (Wick, 1991).
- Loneliness scale (Asher, 1986).

Students whose educational programs depart from that of the general population and who are judged by their teachers to be ineligible for the standard assessment were eligible for a teacher

completed alternate assessment that draws on the following published instruments to achieve these goals:

- Scale of independent behavior-revised (SIBR; Bruininks, Woodcock, Weatherman & Hill, 1996).
- AAMR Adaptive Behavior Scales-School (ABS-S:2) (Lambert, Nihira & Leland, 1990).

The assessment data presented in this report come from the standard assessment. Eligibility for the assessment process included a complete parent interview or family questionnaire, parental consent, and availability of assessors in the area. Local assessors were hired by the study to conduct assessments. These assessors were predominantly school psychologists with backgrounds in assessment as well as some special education teachers. Assessors were responsible for completing between 9 and 30 assessments each. These assessments were conducted from March 2001 through August 2001.

Several steps were followed in order to complete assessments. (1) A screening questionnaire was conducted with teachers knowledgeable about student abilities to determine eligibility for standard vs. alternate assessment, specific subtests, and necessary accommodations. Students received the standard assessment as long as they were able to complete the 1st item on WJ3 letter word identification test. Accommodations during the assessment were intended to reflect the same ones used during instruction. (2) Arrange a suitable time and place to conduct the assessment. Most SEELS assessments were conducted in students' school sites, but some were conducted in family homes. (3) Assessments were conducted as arranged and data were sent to SRI. 4,912 completed standard or alternate assessments were returned for 7,806 eligible sample members (63% response).

School Data Collection

Additional data sources for the analyses reported here were primary language arts teachers of SEELS sample members and teachers most knowledgeable of students' overall programs, who were surveyed by mail. The SEELS conceptual framework holds that language arts instruction is central to the educational experiences of students with disabilities and that classroom context, curriculum, instruction, accommodations, and assessment are crucial to student outcomes and are most amenable to intervention. Language arts teachers are the most knowledgeable about these aspects of students' language arts programs. Further, student experiences span the school day and that content classes, related services, IEP goals, participation in district/state assessments all describe student experiences and relate to student progress. These data are best provided by teachers who are most knowledgeable about the student's program.

The first step in the school data collection process was to identify the current school attended by the sampled students during the 2000-2001 school year. School attendance data had been collected during the parent interview during the summer and fall of 2000. Parent responses relating to schools were coded (e.g., address, phone) using the Quality Education Data (QED) database. For identified schools not in the QED or for students for whom there was no complete parent interview, school district records collected for sampling were used. School attendance data was sent to schools for verification using the School Enrollment Form (SER). In addition to

verification of attendance, the SER form requested that schools provide the name of the teacher who provided primary language arts instruction for the sampled student (for the teacher survey), as well as the name of the teacher who was most knowledgeable about the student's overall school program (for the school program survey).

In March 2001, packets were sent to each school (n=3,827), which included a teacher survey for each sample member, a school program survey for each sample member, and a single school characteristics survey for the school. A second packet was sent in April 2001. Additional mailings were conducted to individual teachers in May 2001 and September 2001. By December 2001, completed teacher surveys were returned for 6,250 out of 10,410 eligible sample members (60% response), and completed school program surveys were returned for 6,213 out of 10,410 eligible sample members (59% response).

Combining Data from Multiple Data Sources

The multivariate analyses reported in Chapters 3 through 6 combine data from multiple sources (e.g., a dependent variable taken from the parent interview and independent variables from the school program survey). Although any single data source has a reasonably high response rate, a smaller number of students 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 students that are omitted from an analysis as the sample declines. SEELS approaches to these two issues are described in this section.

Maintaining the Analytic Sample Size

Two approaches are used in SEELS 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 SEELS analyses can be measured using data from more than one source. For example, parents were asked to describe students' overall grades, and school staff were asked to report students' grades in specific general education and special education language arts 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. Thus, the grades variable includes students who have either a parent interview, a teacher survey, which results in a much larger number of students included in analyses of grades than would result from including those with a single data source. The other variable constructed from a combination of parent and school data is the measure of whether students have been declassified from special education. In that case, preference was given to school-provided information, with parents' reports used if the school program survey item was missing.

Other examples of composite variables that use data from more than one instrument involve classroom characteristics and practices. Measures involving receipt of particular interventions or

services (i.e., tutoring and modifications or accommodations to instruction or testing, presentation or communication, or those related to social adjustment) gave preference to data provided about such programs or services that were indicated on students' Individualized Education Plans (IEPs). If the school program survey was missing for a given student, but he or she had a teacher survey, information about accommodations or services provided in the class reported in that survey was used.

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 that were mentioned previous.

Thus, it can be beneficial to impute values on key variables for students who otherwise would be excluded from analyses because of missing data. Imputation procedures involve assigning a value for a student with missing data that is the best prediction for that student given what else is known about him or her. Although there are a variety of procedures for imputation, SEELS has employed a straightforward assignment of mean values that are calculated for a subset of students who resemble the students 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 students who share his or her disability category and whose head of household has the same level of education. These criteria for subsetting students for purposes of imputation were selected because they relate 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. Exhibit A-2 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.

Exhibit A-2
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 students with same disability category and number of domains with functional limitation	97 to 225 (6.1% to 7.7%)
Functional cognitive skills scale	Mean value of students with same disability category and number of domains with functional limitation	1 to 19 (.1% to .5%)
Household income	Mean value of students with same disability category, head of household education, and race/ethnicity	83 to 241 (6.5% to 7.0%)
Family involvement at home	Mean value of students with same disability category and head of household education	0 to 5 (<.1%)
Family involvement at school	Mean value of students with same disability category and head of household education	30 to 122 (.8% to 4.1%)
School mobility—number of school changes other than grade-level progression	Mean value of students with same disability category, student age, and household income	3 to 7 (.1% to .4%)
Absences excluding suspensions and expulsions (used as an independent variable only)	Mean value of students with same disability category	288 to 809 (20.2 to 28.0%)
Percent of classes in general education	Mean value of students with same disability category	8 to 139 (.4% to 8.8%)
Number of minutes per week in language arts instruction	Mean value of students with the same disability category	20 to 376 (1.1% to 10.4%)

Understanding the characteristics of students included in analyses. As mentioned above, combining data from multiple sources in a given analysis necessarily limits the students included in it to those who have both data sources. It is important to understand the extent to which the included subset of students is similar to or differs from the full sample in order to know whether the results of the analysis generalize to all students or only to those represented in the subset. To address this question, SEELS compared means for all dependent and independent variables used in each multivariate model reported in this document with those of the full sample of students for whom there are data. The number of cases included in each model and the results of the analyses of means and standard errors are reported in Exhibit A-3. There are a number differences in the subsamples of data used for various models from their means for the entire sample. However, for the most part, these differences are small and, thus, unlikely to affect the results of the multivariate analyses. Below we describe these differences for dependent as well as independent variables.

Dependent Variables

The means values of the dependent variables in the subsamples used in most of the models do not differ from the means for the entire sample. Exceptions are that the percentage of students who belong to groups differs from the entire sample by 1 percentage point, mean grades differ by .1 point on a 9-point scale, and the gap between mathematics standardized tests and grade level differs by .12 of a grade level. In addition, the mean scores on the classroom behavior scores in general and special education differ by .2 points and .35 points, respectively, on a 5-point scale.

Independent Variables

The means of the following independent variables differ in the subsamples for one or more of the models from the entire sample.

Disability Characteristics

- The percentage of students with most disabilities in the subsamples used for the models does not differ from the total sample by more than 8 percentage points.

- The subsamples used for the models include from 6 to 13 percentage points more students with ADD/ADHD than the entire sample.

- Students included in the models of absenteeism, disciplinary action, belonging to groups, and classroom behavior in general education classes were an average of 1 to 4 months younger than students entire sample when their disability was discovered than students in the entire sample. Students included the models of motivation, passage comprehension, calculation, grades, retention in grade, locus of control, and classroom behavior in general education classes were an average of 1.8 to 5.0 months older when their disability was discovered than students entire sample.

- The mean number of domains in which students have problems differs in the subsamples for some of the models from the entire sample, but never by more than .1 problems, except for the model of classroom behavior in special education classes. In that subsample, the mean number of health problems exceeds that of the entire sample by .3 problems.

- The mean general health score of students in the subsample for the model of behavior in general education classes is .2 points higher than that of the entire sample on a 5-point scale. In contrast, the mean general health score of students in the subsample for the model of behavior in special education classes is .08 points lower than that of the entire sample.

Functioning

- Compared to the entire sample, students included in the analyses of motivation, passage comprehension, calculation, grades, retention, reading discrepancy,

mathematics discrepancy, and behavior in general education classrooms have higher average self-care skill scale scores functional mental skills scale scores, and social skills scale scores than the entire sample. Students included in the model of locus of control also have higher average functional mental skills than the entire sample. In contrast, students included in the models of days absent, disciplinary action, belonging to groups, and behavior in special education classrooms have lower average functional cognitive scale scores than the entire sample. Except for the models of classroom behavior, mean scores on the self-care skills scale do not differ from the entire sample by more than .2 on a 7-point scale, mean scores on the functional cognitive skills scale do not differ by more than .3 on a 13-point scale, and mean scores on the social skills scale do not differ by more than .3 on a 19-point scale. Compared with the entire sample of students, the mean scores on the self care ability scale, functional cognitive scale, and social skills scale for the subsample for the model of behavior in general education classes is are .1 points higher, 1.2 points higher, and .6 points higher, respectively, and the mean scores for the subample for the model of behavior in special education classes are .3, 1.1, and .3 points lower.

- Students' mean value on the persistence scale in the subsets for the models does not differ from the entire sample by more than .04 of a point on a 13-point scale.

Demographics

- Students included in the models of grades, retention in grade, locus of control, and behavior in special education classrooms were an average of approximately 3.1 months, 1.6 months, 1.7 months, and 1.6 months older, respectively, than the entire sample of students.. Students included in the model of classroom behavior in general education classrooms were an average of 3.7 months younger than the entire sample of students.

- In most of the models, the subsamples included a smaller proportion of minority students than the total sample; however, the greatest differences between the total sample and a subsample for any model were 6 percentage points for African American students, except for the model of behavior in general education classrooms, and 4 percentage points for Hispanic students. The subsample for the model of behavior in general education classrooms included .12 percentage points fewer African American students than the entire sample. Differences between the percentage of students who spoke a language other than English in the home in the samples for the models and the entire sample did not exceed 6 percentage points.

Household Characteristics

- The mean family income of students in subsamples for the models is higher than the mean for students in the entire sample, but never by more than .7 on 16-point scale, except for subsample for the model of behavior in general education classrooms, in which case it is 1.8 points higher than for the entire sample.

- The level of involvement of student's families at home is higher in subsamples for all models except absenteeism and behavior in special education classrooms than in the entire sample, but not by more than .4 on a 8-point scale. The mean level of involvement for the subsample for behavior in special education classrooms is 1 point lower than for the entire sample.

- The level of involvement of student's families in their schools is higher in all models except retention in grades and behavior in special education classrooms than in the entire sample, but never by more than .5 on a 13-point scale, except for behavior in general education classrooms. For that subsample, the difference is 1.0 points higher than for the entire sample. The mean for the subsample for the model of retention in grade does not differ from the mean for the entire sample, and the mean for the subsample for the model of behavior in special education classrooms is .2 points lower than for the entire sample.

- Expectations for postsecondary achievement in the subsets for the models do not differ from the entire sample by more than .1 point on a 4 scale, except for the models of classroom behavior. Compared with the mean for the entire sample, the mean for the subsample for the model of behavior in general education classrooms is higher by .3 points, and the mean for the subsample for the model of behavior in special education classrooms is lower by .2 points.

School Programs and Experiences

Maximum differences between the subsamples for the models, except models of classroom behavior, and the entire sample are as follows:

- Mean percentage of classes in general education: 5 percentage points.

- Mean class size: .6 of a student.

- Mean degree of modifications to the language arts curriculum: .4 on a 6-point scale.

- Mean number of modifications to tests: .3 modifications.

- Mean number of communication aides: .2 communication aides.

- Mean number of days absent: .2.

- Mean frequency with which students receive which whole-class instruction, small group instruction, and individual instruction from teachers: .2, .1, and .1, respectively on a 4-point scale.

- Mean level of participation in general instructional activities, literature reading activities, and skill building reading activities: 1.3 on a 29-point scale, .7 points on a 13-point scale, and .4 points on a 12-point scale, respectively.

- Mean grades: .06 on a 9-point scale.

- Mean number of times a students changed schools: .1 school changes.

- Percentage of students retained in grade: 2 percentage points.

- Percentage of students who receive tutoring from an adult: <.1 percentage point.

Compared with the means for the entire sample, differences between the subsamples for the models of classroom behavior and the entire sample are as follows:

- The mean percentage of classes in general education is 26.5 percentage points higher for the subsample for the model of behavior in general education classrooms and 23.6 percentage points lower for the subsample for the model of behavior in special education classes. Such differences are expected, given that one subsample purposely includes only students in general education classes and the other purposely includes only students in special education classes.

- The mean size of the student's language arts class is smaller by 7.5 for the subsample for the model of behavior in general education classrooms and smaller by 6.3 students for the model of behavior in special education classrooms.

- The mean number of modifications to tests is .73 smaller for the subsample for the model of behavior in general education and .73 larger for the subsample for the model of behavior in special education.

- The mean number days absent is .23 smaller for the subsample for the model of behavior in general education and .08 larger for the subsample for the model of behavior in special education.

- On a 4-point scale, the mean frequencies with which students receive which whole-class instruction, small group instruction, individual instruction from a teacher, and individual instruction from another adult are .4, .1, .1, and .1 points higher, respectively, for the subsample for the model of behavior in general education and .2, .2, .1, and .1 points lower, respectively, for the subsample for the model of behavior in special education.

- The mean levels of participation in literature reading activities and skill building reading activities are 1.1 points higher and .22 points higher, respectively, for the subsamples for the model of behavior in general education classrooms. The mean levels of participation in general instructional activities and literature reading activities are 20.2 points lower and .6 points lower, respectively, lower for the subsample for behavior in special education classrooms. The mean level of participation in skill building activities does not differ for the subsample for behavior in special education classrooms.

- Mean grades are .1 higher for the subsample for the model of behavior in general education classrooms and .1 lower for the model of behavior in special education classrooms.

- The mean number of times a student changed schools is .2 changes lower for the subsample for the model of behavior in general education classrooms and .1 changes higher or the model of behavior in special education classrooms.

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS

		Multivariate Analysis of:													
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Sample Size	**10,739**	**2,190**	**1,303**	**1,554**	**1,466**	**2,038**	**1,587**	**3,662**	**3,662**	**3,566**	**1,664**	**2,228**	**2,211**	**1,109**	**1,553**
DEPENDENT VARIABLES															
Average:															
Days absent per month	1.320	1.265	-	-	-	-	-	-	-	-	-	-	-	-	-
	(0.037)	(0.053)	-	-	-	-	-	-	-	-	-	-	-	-	-
Motivation for schooling	9.997	-	10.003	-	-	-	-	-	-	-	-	-	-	-	-
	(0.011)	-	(0.017)	-	-	-	-	-	-	-	-	-	-	-	-
Standard score on passage comprehension	80.117	-	-	79.331	-	-	-	-	-	-	-	-	-	-	-
	(0.367)	-	-	(0.571)	-	-	-	-	-	-	-	-	-	-	-
Standard score on calculation	89.079	-	-	-	88.401	-	-	-	-	-	-	-	-	-	-
	(0.335)	-	-	-	(0.514)	-	-	-	-	-	-	-	-	-	-
Overall grades across all subjects (9-point scale)	6.429	-	-	-	-	6.525	-	-	-	-	-	-	-	-	-
	(0.020)	-	-	-	-	(0.037)	-	-	-	-	-	-	-	-	-

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

		Multivariate Model:													
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Locus of control	10.005	-	-	-	-	-	-	-	-	-	10.044	-	-	-	-
	(0.012)	-	-	-	-	-	-	-	-	-	(0.017)	-	-	-	-
Discrepancy between grade level and reading level on standardized tests (in years)	-1.483	-	-	-	-	-	-	-	-	-	-	-1.569	-	-	-
	(0.026)	-	-	-	-	-	-	-	-	-	-	(0.041)	-	-	-
Classroom behavior scale score in general education classrooms (scale: 2 to 6)	12.446	-	-	-	-	-	-	-	-	-	-	-	-	12.251	-
	(0.047)	-	-	-	-	-	-	-	-	-	-	-	-	(0.072)	-
Classroom behavior scale score in special education classrooms (scale: 2 to 6)	11.550	-	-	-	-	-	-	-	-	-	-	-	-	-	11.904
	(0.045)	-	-	-	-	-	-	-	-	-	-	-	-	-	(0.063)
Discrepancy between grade level and mathematics level on standardized tests (in years)	-1.359	-	-	-	-	-	-	-	-	-	-	-	-1.476	-	-
	(0.024)	-	-	-	-	-	-	-	-	-	-	-	(0.037)	-	-
Percentage:														-	-
With disciplinary actions in the past year	0.153	-	-	-	-	-	-	0.112	-	-	-	-	-	-	-
	(0.003)	-	-	-	-	-	-	(0.005)	-	-	-	-	-	-	-
Retained at grade level in the past 3 years	0.118	-	-	-	-	-	0.137	-	-	-	-	-	-	-	-
	(0.005)	-	-	-	-	-	(0.009)	-	-	-	-	-	-	-	-
Belong to a group	0.606	-	-	-	-	-	-	-	0.633	-	-	-	-	-	-
	(0.005)	-	-	-	-	-	-	-	(0.008)	-	-	-	-	-	-

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

		Multivariate Model:													
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
See friends at least weekly	0.551	-	-	-	-	-	-	-	-	0.556	-	-	-	-	-
	(0.005)	-	-	-	-	-	-	-	-	(0.008)	-	-	-	-	-
EXPLANATORY VARIABLES															
Individual Characteristics															
Percentage with:															
Speech impairment	0.094	0.036	0.044	0.041	0.042	0.043	0.037	0.039	0.039	0.040	0.041	0.043	0.043	0.068	0.018
	(0.003)	(0.004)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)	(0.004)	(0.008)	(0.003)
Mental retardation	0.097	0.100	0.091	0.093	0.087	0.087	0.093	0.093	0.093	0.095	0.090	0.091	0.090	0.039	0.134
	(0.003)	(0.006)	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)	(0.006)	(0.006)	(0.006)	(0.009)
Emotional disturbance	0.099	0.076	0.087	0.083	0.086	0.091	0.085	0.078	0.078	0.079	0.091	0.084	0.085	0.071	0.084
	(0.003)	(0.006)	(0.008)	(0.007)	(0.007)	(0.006)	(0.007)	(0.004)	(0.004)	(0.005)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)
Hearing impairment	0.109	0.120	0.143	0.147	0.148	0.138	0.131	0.121	0.121	0.107	0.140	0.133	0.134	0.115	0.135
	(0.003)	(0.007)	(0.010)	(0.009)	(0.009)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)	(0.009)	(0.007)	(0.007)	(0.010)	(0.009)
Visual impairment	0.085	0.079	0.060	0.064	0.065	0.079	0.069	0.082	0.082	0.079	0.061	0.079	0.077	0.119	0.045
	(0.003)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)	(0.010)	(0.005)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Orthopedic impairment	0.103	0.093	0.102	0.099	0.100	0.100	0.084	0.100	0.100	0.103	0.109	0.101	0.099	0.128	0.073
	(0.003)	(0.006)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.008)	(0.006)	(0.006)	(0.010)	(0.007)
Other health impairment	0.082	0.109	0.141	0.133	0.136	0.128	0.141	0.116	0.116	0.119	0.149	0.124	0.122	0.145	0.091
	(0.003)	(0.007)	(0.010)	(0.009)	(0.009)	(0.007)	(0.009)	(0.005)	(0.005)	(0.005)	(0.009)	(0.007)	(0.007)	(0.011)	(0.007)
Autism	0.098	0.160	0.114	0.131	0.126	0.119	0.115	0.149	0.149	0.152	0.110	0.138	0.137	0.131	0.162
	(0.003)	(0.008)	(0.009)	(0.009)	(0.009)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)	(0.008)	(0.007)	(0.007)	(0.010)	(0.009)
Traumatic brain injury	0.038	0.037	0.033	0.035	0.035	0.037	0.033	0.035	0.035	0.036	0.034	0.038	0.037	0.032	0.039
	(0.002)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Multiple disabilities or Deaf/blindness	0.079	0.084	0.052	0.052	0.048	0.053	0.068	0.083	0.083	0.084	0.052	0.055	0.057	0.034	0.118
	(0.003)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.008)
ADHD/HD	0.323	0.384	0.450	0.439	0.449	0.418	0.407	0.392	0.392	0.393	0.448	0.409	0.411	0.379	0.402
	(0.005)	(0.010)	(0.014)	(0.013)	(0.013)	(0.011)	(0.012)	(0.008)	(0.008)	(0.008)	(0.012)	(0.010)	(0.010)	(0.015)	(0.012)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Average:															
Age when child started having this difficulty/condition	2.886	2.715	3.134	3.035	3.100	3.089	3.298	2.753	2.753	2.802	3.081	2.911	2.930	3.057	2.578
	(0.030)	(0.057)	(0.076)	(0.069)	(0.071)	(0.061)	(0.073)	(0.044)	(0.044)	(0.045)	(0.067)	(0.057)	(0.057)	(0.082)	(0.067)
Number of problems reported for seeing, speaking, conversing, communicating, appendage use, and/or health	1.770	1.860	1.729	1.768	1.723	1.704	1.689	1.858	1.858	1.852	1.730	1.747	1.743	1.536	2.085
	(0.011)	(0.027)	(0.034)	(0.031)	(0.032)	(0.027)	(0.032)	(0.021)	(0.021)	(0.021)	(0.030)	(0.026)	(0.026)	(0.034)	(0.033)
General health score (scale: 1 to 5)	3.962	3.997	-	-	-	-	-	-	-	-	-	-	-	4.167	3.880
	(0.012)	(0.023)	-	-	-	-	-	-	-	-	-	-	-	(0.030)	(0.028)
Self care ability scale score (scale: 2 to 8)	7.043	7.006	7.266	7.228	7.269	7.239	7.243	7.039	7.039	7.031	7.259	7.207	7.214	7.185	6.789
	(0.013)	(0.030)	(0.030)	(0.029)	(0.029)	(0.025)	(0.030)	(0.023)	(0.023)	(0.023)	(0.027)	(0.025)	(0.025)	(0.034)	(0.037)
Functional cognitive skills scale score (scale: 4 to 16)	11.165	10.946	11.724	11.553	11.731	11.816	12.058	11.032	11.032	11.037	11.797	11.486	11.490	12.333	10.081
	(0.031)	(0.075)	(0.080)	(0.076)	(0.075)	(0.068)	(0.080)	(0.058)	(0.058)	(0.059)	(0.073)	(0.067)	(0.067)	(0.086)	(0.090)
Social skills scale score: 9 To 27)	19.732	19.860	19.998	19.930	19.987	20.052	20.011	19.809	19.809	19.802	-	20.036	20.048	20.344	19.429
	(0.036)	(0.071)	(0.089)	(0.082)	(0.084)	(0.070)	(0.080)	(0.055)	(0.055)	(0.055)	-	(0.068)	(0.068)	(0.096)	(0.084)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motiva-tion for Schooling	Passage Comprehension	Calcu-lation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Persistence scale score (scale: 6 to 18)	2.131	2.161	2.160	2.156	2.164	2.174	2.162	2.137	2.137	2.132	2.162	2.170	2.170	2.170	2.119
	(0.007)	(0.015)	(0.019)	(0.017)	(0.018)	(0.015)	(0.017)	(0.011)	(0.011)	(0.012)	(0.017)	(0.014)	(0.015)	(0.020)	(0.018)
Age as of 12/01/2000	10.311	10.241	10.376	10.347	10.363	10.568	11.655	10.296	10.296	10.278	10.452	10.278	10.275	10.005	10.443
	(0.017)	(0.039)	(0.049)	(0.045)	(0.047)	(0.039)	(0.028)	(0.030)	(0.030)	(0.030)	(0.043)	(0.038)	(0.038)	(0.053)	(0.047)
Percentage															
Male	0.661	0.672	0.664	0.668	0.667	0.664	0.674	0.664	0.664	0.665	0.664	0.664	0.666	0.654	-
	(0.004)	(0.010)	(0.013)	(0.012)	(0.012)	(0.010)	(0.012)	(0.008)	(0.008)	(0.008)	(0.012)	(0.010)	(0.010)	(0.014)	-
African American	0.217	0.173	0.157	0.157	0.149	0.169	0.175	0.165	0.165	0.165	0.153	0.158	0.157	0.094	0.223
	(0.004)	(0.008)	(0.010)	(0.009)	(0.009)	(0.008)	(0.010)	(0.006)	(0.006)	(0.006)	(0.009)	(0.008)	(0.008)	(0.009)	(0.011)
Hispanic	0.129	0.126	0.092	0.095	0.098	0.128	0.123	0.122	0.122	0.121	0.088	0.126	0.127	0.097	0.149
	(0.003)	(0.007)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.009)	(0.009)
Other race/ethnicity	0.035	0.025	0.024	0.024	0.025	0.021	0.025	0.025	0.025	0.024	0.024	0.026	0.025	0.031	0.020
	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)	(0.005)	(0.004)
Language other than English regularly spoken in home	0.179	0.174	0.142	0.149	0.149	0.176	0.169	0.169	0.169	0.161	0.138	0.173	0.174	0.115	0.219
	(0.004)	(0.008)	(0.010)	(0.009)	(0.009)	(0.008)	(0.009)	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.010)	(0.010)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Household characteristics															
Average:															
Family income (scale: 1 to 16)	7.809	8.380	8.429	8.446	8.517	8.281	8.467	8.505	8.505	8.561	8.541	8.413	8.438	9.565	7.504
	(0.042)	(0.102)	(0.130)	(0.120)	(0.124)	(0.105)	(0.120)	(0.079)	(0.079)	(0.080)	(0.115)	(0.101)	(0.101)	(0.142)	(0.118)
Family involvement at home (scale 1 to 8)	7.640	7.715	8.009	7.997	8.039	8.020	7.771	7.731	7.731	7.727	7.996	7.973	7.979	7.234	6.617
	(0.015)	(0.038)	(0.037)	(0.035)	(0.035)	(0.031)	(0.039)	(0.029)	(0.029)	(0.029)	(0.033)	(0.031)	(0.031)	(0.033)	(0.041)
Family involvement at school (scale: 0 to 12)	4.073	4.297	4.550	4.556	4.588	4.443	4.200	4.432	4.432	4.482	4.595	4.495	4.504	5.077	3.874
	(0.024)	(0.062)	(0.082)	(0.075)	(0.077)	(0.065)	(0.073)	(0.049)	(0.049)	(0.050)	(0.073)	(0.062)	(0.063)	(0.090)	(0.071)
Expectations for postsecondary education score (scale: 1 to 4)	2.905	2.903	3.005	2.981	3.000	3.002	2.901	2.907	2.907	2.908	3.006	2.995	2.996	3.196	2.701
	(0.009)	(0.018)	(0.021)	(0.020)	(0.020)	(0.017)	(0.021)	(0.014)	(0.014)	(0.014)	(0.019)	(0.017)	(0.017)	(0.021)	(0.022)
School program and experiences															
Average:															
Percentage of classes in general education	56.753	53.243	60.945	58.640	59.991	58.098	53.937	55.726	55.726	56.943	61.712	58.526	58.476	83.287	33.180
	(0.258)	(0.721)	(0.873)	(0.816)	(0.831)	(0.719)	(0.822)	(0.551)	(0.551)	(0.548)	(0.764)	(0.686)	(0.687)	(0.520)	(0.667)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motiva-tion for Schoolin g	Passage Compreh ension	Calcu-lation	Grades	Retained in Grade	Disciplina ry Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepan cy	Mathematic s Discrepanc y	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Size of language arts class	15.661	15.016	16.045	15.557	15.801	15.875	15.661	-	-	-	-	-	-	23.154	9.350
	(0.113)	(0.176)	(0.228)	(0.207)	(0.214)	(0.189)	(0.222)	-	-	-	-	-	-	(0.167)	(0.110)
Number of social adjustment supports	0.441	0.443	0.437	0.442	0.443	0.427	-	0.432	0.432	0.421	0.429	0.435	0.439	0.344	0.502
	(0.010)	(0.016)	(0.021)	(0.019)	(0.020)	(0.017)	-	(0.012)	(0.012)	(0.012)	(0.019)	(0.016)	(0.016)	(0.021)	(0.020)
Scale of change in language arts curriculum (scale: 2-8)	4.219	4.199	3.816	3.897	3.852	3.890	4.064	-	-	-	-	3.929	3.928	3.087	4.972
	(0.023)	(0.036)	(0.039)	(0.037)	(0.038)	(0.033)	(0.040)	-	-	-	-	(0.032)	(0.033)	(0.036)	(0.039)
Number of modifications to tests, assignments, grades, etc	4.059	-	4.187	4.231	4.215	4.269	4.321	-	-	-	4.200	4.260	4.270	3.331	4.782
	(0.035)	-	(0.074)	(0.068)	(0.070)	(0.059)	(0.068)	-	-	-	(0.066)	(0.057)	(0.057)	(0.077)	(0.067)
Number of presentation/communication aides	0.793	-	0.622	0.650	0.625	0.670	0.727	-	-	-	0.651	0.691	0.688	-	-
	(0.015)	-	(0.030)	(0.028)	(0.028)	(0.024)	(0.027)	-	-	-	(0.027)	(0.023)	(0.023)	-	-
Days absent from class	1.329	-	1.213	1.174	1.175	1.224	1.326	1.295	1.295	1.303	1.218	1.250	1.249	1.097	1.412
	(0.016)	-	(0.058)	(0.050)	(0.051)	(0.046)	(0.054)	(0.037)	(0.037)	(0.037)	(0.049)	(0.047)	(0.047)	(0.051)	(0.064)
Highest level of teacher's education	2.288	-	2.308	2.311	2.319	2.264	2.294	-	-	-	-	2.297	2.297	-	-
	(0.012)	-	(0.026)	(0.024)	(0.025)	(0.021)	(0.023)	-	-	-	-	(0.020)	(0.020)	-	-

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONTINUED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Frequency of whole-class instruction for student (scale: 1 to 4)	3.256	-	-	3.394	3.413	3.416	3.415	-	-	-	-	3.373	3.369	3.656	3.064
	(0.014)	-	-	(0.024)	(0.024)	(0.020)	(0.023)	-	-	-	-	(0.020)	(0.020)	(0.018)	(0.029)
Frequency of small group instruction for student (scale: 1 to 4)	3.355	-	-	3.450	3.447	3.419	3.358	-	-	-	-	3.451	3.451	3.249	3.560
	(0.011)	-	-	(0.018)	(0.019)	(0.016)	(0.019)	-	-	-	-	(0.015)	(0.015)	(0.021)	(0.019)
Frequency of individual instruction from teacher for student (scale: 1 to 4)	3.324	-	-	3.293	3.274	3.275	3.263	-	-	-	-	3.306	3.306	3.177	3.444
	(0.010)	-	-	(0.018)	(0.018)	(0.016)	(0.019)	-	-	-	-	(0.015)	(0.015)	(0.022)	(0.018)
Frequency of individual instruction from another for student (scale: 1 to 4)	3.324													2.726	2.722
	(0.010)													(0.032)	(0.031)
Participation in general instructional activities (scale 0-28)	20.858	21.126	22.113	21.858	22.000	22.047	21.638	-	-	-	-	21.860	21.873	-	0.674
	(0.062)	(0.096)	(0.088)	(0.086)	(0.086)	(0.079)	(0.104)	-	-	-	-	(0.075)	(0.075)	-	(0.012)
Participation in literature reading activities overall (scale 0-12)	8.776	-	9.492	9.334	9.448	9.410	9.188	-	-	-	-	9.331	9.342	9.851	8.185
	(0.037)	-	(0.055)	(0.054)	(0.052)	(0.048)	(0.062)	-	-	-	-	(0.047)	(0.047)	(0.055)	(0.075)

Exhibit A-3
UNWEIGHTED SAMPLE SIZES OF FULL SAMPLE AND EACH MULTIVARIATE ANALYSIS AND MEANS AND STANDARD ERRORS OF VARIABLES INCLUDED IN THE FULL SAMPLE AND IN EACH MULTIVARIATE ANALYSIS (CONCLUDED)

	Multivariate Model:														
	Entire Sample	Days Absent	Motivation for Schooling	Passage Comprehension	Calculation	Grades	Retained in Grade	Disciplinary Action	Belongs to Group	Sees Friends	Locus of Control	Reading Discrepancy	Mathematics Discrepancy	Classroom Behavior Scale (General Education)	Classroom Behavior Scale (Special Education)
Participation in skill building reading activities (scale 0-12)	9.384	-	9.779	9.725	9.766	9.739	9.284	-	-	-	-	9.775	9.787	9.606	9.352
	(0.031)	-	(0.049)	(0.046)	(0.047)	(0.041)	(0.052)	-	-	-	-	(0.039)	(0.039)	(0.055)	(0.064)
Typical grades received for coursework (scale: 1 to 9)	2.692	-	-	-	-	-	-	2.634	2.634	-	-	-	-	2.764	2.605
	(0.010)	-	-	-	-	-	-	(0.017)	(0.017)	-	-	-	-	(0.031)	(0.028)
Number of school changes	0.904	0.918	0.854	0.869	0.875	0.913	1.026	0.893	0.893	0.885	0.865	0.883	0.888	0.718	1.040
	(0.010)	(0.026)	(0.033)	(0.030)	(0.031)	(0.027)	(0.032)	(0.020)	(0.020)	(0.020)	(0.029)	(0.025)	(0.026)	(0.033)	(0.032)
Percentage:															
Ever retained in grade	0.238	0.243	0.262	0.267	0.268	0.260	-	0.248	0.248	0.246	0.255	0.263	0.261	0.241	0.258
	(0.004)	(0.009)	(0.012)	(0.011)	(0.012)	(0.010)	-	(0.007)	(0.007)	(0.007)	(0.011)	(0.009)	(0.009)	(0.013)	(0.011)
Who receive tutoring from an adult	0.568	-	0.523	0.542	0.533	0.551	0.551	-	-	-	-	0.560	0.560	-	-
	(0.006)	-	(0.014)	(0.013)	(0.013)	(0.011)	(0.012)	-	-	-	-	(0.011)	(0.011)	-	-

- =Variable not in model.

Weighting Wave 1 Data

The percentages and means reported in the data tables are estimates of the true values for the population of students with disabilities in the SEELS age range. The estimates are calculated from responses of parents of SEELS sample members. The response for each sample member is weighted to represent the number of students 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.

Exhibit A-4 illustrates the concept of sample weighting and its effect on percentages or means that are calculated for students 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 value regarding whether that student participated in organized group activities outside of school (1 for yes, 0 for no). Six students participated in such activities, which would result in an unweighted value of 60% participating. However, this would not accurately represent the national population of students with disabilities because many more students 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 students in the population that are from each disability category (actual SEELS 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 87%. The percentages in all SEELS 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 Exhibit A-4).

Exhibit A-4
EXAMPLE OF WEIGHTED PERCENTAGE CALCULATION

Disability Category	A Number in Sample	B Participated in Group Activities	C Weight for Category	D Weighted Value for Category
Learning disability	1	1	4.3	4.3
Speech/language impairment	1	1	3.0	3.0
Mental retardation	1	1	1.0	1.0
Emotional disturbance	1	0	.8	0
Hearing impairment	1	1	.1	.1
Visual impairment	1	1	.1	.1
Orthopedic impairment	1	0	.1	0
Other health impairment	1	1	.4	.4
Autism	1	0	.1	0
Multiple disabilities	1	0	.1	0
TOTAL	10	6	10	8.9
	Unweighted sample percentage = 60% (Column B total divided by Column A total)		Weighted population estimate = 89% (Column D total divided by Column C total)	

The students in LEAs and state schools with parent interview/survey data 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.
- For each of the 64 LEA cells, the number of students in each 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), then we would estimate there to be 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.
- The initial student sampling weights were adjusted by disability category so that the sum of the weights (that is, the initial student sampling weights multiplied by the number of

students with completed interviews) was equal to the number of students in the geographical and wealth cells of each size strata. 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 5 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 (discussed below).

- 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 1999-2000 school year (OSEP, 2001).

As mentioned earlier, the imposition of constraints on the adjusted weights increased sampling efficiency at the cost of introducing a small amount of bias. The largest increases in sampling efficiency and the largest biases occurred for the categories of autism and visual impairment; the smallest increase in efficiency and biases occurred for specific learning disabilities. The principal bias for autism was the reduction in the proportion of students from the Northeast (from 22% to 18%), from the West/Southwest (from 34% to 30%) and from small LEAs (from 16% to 13%). The principal bias for visual impairment is in small LEAs (from 12% to 4%), in very wealthy LEAs (from 20% to 17%). For the category of learning disability, all biases introduced by the imposition of constraints on the student weights are negligible. Considering the increase in sampling efficiency for autism (from 23% to 53%) and visual impairment (from 18% to 53%), we consider these biases to be acceptable.

The reason for the reduction in the proportion of students represented in the cells mentioned above is that there were relatively few students with interview/survey data in those cells. For example, in small LEAs, there were only six students with visual impairments with data, requiring that they represent an estimated 1,771 students with visual impairments from small LEAs. The weighting program determined that the average weight required (i.e., 295) violated the constraints, and therefore reduced these weights to a more reasonable value (i.e., 84.4).

Estimating Standard Errors

The SEELS sample is both stratified and clustered, so that calculating standard errors by formula is not straightforward. Standard errors for means and proportions can also be estimated using pseudo-replication, a procedure that is widely used by the U.S. Census Bureau and other federal agencies involved in fielding complex surveys. To that end, we developed a set of weights for each of 50 half-replicate subsamples. Each half-replicate involved randomly selecting half of the total set of LEAs that provided contact information and then weighting that half to represent the entire universe. Randomization was accomplished within each of the 64 sampling cells. 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) calculate the squares of the deviations of the half-sample estimate from the full sample estimate; and 3) adding the squared deviations and divide by (n-1) where n is the number of half-replicates.

Although the procedure of pseudo-replication 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 SEELS, and it is computationally expensive. In the past, we have found that 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*.

SEELS respondents are not independent of each other because they are clustered in LEAs and the intra-cluster correlation is not zero. However, the intra-cluster correlation traditionally has been quite small, so that the formula for the effective sample size shown above has worked well. To be conservative, however, we multiplied the initial estimate by a "safety factor" that assures that we will not underestimate the standard error of estimate.

To determine the adequacy of fit of the variance estimate based on the effective sample size and to estimate the required safety factor, we selected 24 questions with 95 categorical and 2 continuous responses. We calculated standard errors of estimates 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 pseudo-replicate standard error estimate for 92% of the categorical responses and 89% of the mean responses. Because the pseudo-replicate estimates of standard error are themselves estimates of the true standard error, and are therefore subject to sampling variability, we considered this to be an adequate margin of safety. All standard errors in Wave 1 are 3% or less, except for categories of deaf-blindness and traumatic brain injury, where sample sizes are very small.

Calculating Significance Levels

Readers may want to compare percentages or means for different subgroups to determine, for example, whether the difference in the percentage of students in poverty between students with learning disabilities and those with mental retardation is greater than would be expected to occur by chance. To calculate whether the difference between percentages is statistically significant with 95% confidence (often denoted as $p<.05$), 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 product of this calculation is 6.63 to 10.79, the significance level is .01, products of 10.8 or greater are significant at the .001 level.

Multivariate Analysis Methods

Multivariate techniques are used in this report to assess the independent relationships between outcome measures and characteristics of individual students, their households, and their school program and experiences.

Multiple linear regression analysis is used to examine the variation in ordinal dependent variables (i.e., days absent, classroom engagement behavior scale scores, grades, discrepancies in reading and math levels, and household responsibilities scale scores). 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 X's are the values of the independent variables. When the dependent variables are dichotomous (i.e., whether student belong to groups, see friends at least weekly, have been subject to disciplinary actions, have been involved with the criminal justice system, or hold a job), logistic regression is used [e.g., log(probability of criminal justice system involvement/no involvement) $= a + b_1X_1 + b_2X_2 + ... + b_nX_n$]. Both types of regression allow the modeling of the simultaneous influence of predictor variables on the dependent variable and provide estimates of model fit. For ease of interpretation, coefficients of logistic regression analyses are transformed into differences in the probabilities of the dependent variable occurring given a specified increment of difference in the independent variables.

SEELS multivariate analyses and correlations are unweighted. In general, results are reported for analyses that include the full set of individual, household, and school factors simultaneously. The one exception is that analyses of the relationships of individual social adjustment interventions or supports to related outcomes reported in Chapter 5, Exhibit 5-10 (i.e., receipt of mental health, social work, or behavior intervention services; participation in an anger management program; or having a behavior management plan) considered each of those interventions separately (along with all other individual, household, or school factors) because of higher intercorrelations among them. Coefficients for the individual, household, and other school factors in those analyses are those resulting from analyses that exclude the individual interventions.

In reporting the explained variation for multivariate analyses, an r^2 is used for a linear regression, which describes the percent of the variance in a continuous variable explained by the model. Although an r^2 can be calculated for dichotomous variables used in logistic regression, it

is much less useful than for continuous variables owing to the near constancy of variance over wide ranges of underlying probabilities of success. Many alternative pseudo-r^2 statistics have been proposed to measure "goodness of fit" of logistic regression models, but most of these are quite complex and difficult to interpret.

This report uses a statistic that we have called "predictive improvement". This statistic is scaled from 0 to 1, like r^2 is easier to interpret than pseudo-r^2 statistics, and heuristically represents the proportion of the maximum possible improvement in predictive ability associated with the explanatory (independent) variables in a logistic regression.[4] Referred to as "predictive improvement" (PI), the statistic is calculated in the following way:

$$PI = 1 - (e_0 + e_1)$$

Where

e_0 is the model's "rate of error" in predicting observations that actually have a value of 1 on the dependent variable This is obtained by taking the mean of the values predicted by the model for those observations.

and

e_1 is1 model's rate of error in predicting observations that actually have a value of 0 on the dependent variable. This is obtained by taking 1 minus the mean of the values predicted by the model for those observations.

This simple statistic represents the percentage of improvement in predictive power that a specific logistic model gives over a logistic model that includes only a constant term.[5] For a model that

[4] The PI statistic was developed by Harold Javitz.

[5] More specifically, consider an experiment in which two logistic models are used to predict the probability of a positive outcome. One of these models includes only a constant, and the other includes a variety of explanatory (independent) variables. After fitting the model, the data set is divided into two groups—individuals with a positive outcome and individuals with a negative outcome. A large number of individuals (say 1,000) are selected from the first group randomly and with replacement. The same number of individuals are selected from the second group randomly and with replacement. Using the logistic model that includes only a constant term, the experimenter estimates the probability of a positive outcome for each of these 2,000 selections. (When the model only includes a constant term, this probability will always equal the proportion of positive outcomes in the original dataset). Once this probability is estimated for an individual, the experimenter flips a coin with that same probability for heads. If the coin comes up heads and the individual actually had a positive outcome, or if the coin comes up tails and the individual actually had a negative outcome, then the experimenter scores a success; otherwise the experimenter scores a failure. Using the logistic model with only a constant term, the overall proportion of successes for these 2,000 randomly selected individuals will be approximately 50%. The experimenter now repeats this process using the logistic model with one or more explanatory variables. (In this case, the estimated probability of success will vary from person to person, and therefore the coin that the experimenter flips will have probability of a heads that also varies from person to person). The overall proportion of successes for the same 2,000 randomly selected individuals will typically be greater than 50% (depending on the extent to which the explanatory variables improve predictive accuracy). Suppose that the overall proportion of successes is 74%. Then the use of the explanatory variables has increased the proportion of correct guesses from 50% to 74%. This is an improvement of 24%. Since the maximum improvement is 50% (i.e., improving predictive accuracy from 50% to 100%), the percent

predicts no better than chance, PI has a value of 0. As a model's predictive power improves, the value of PI increases, so that if a model were able to predict every observation perfectly, PI would have a value of 1.

Measurement Issues

The chapters in this report include information on specific variables included in analyses. However, several general points about SEELS measures that are used repeatedly in analyses should be clear to readers as they consider the findings reported here.

Categorizing students by primary disability. Information about the nature of students' disabilities came from rosters of all students in the SEELS age range receiving special education in the 1999-2000 school year under the auspices of participating LEAs and state-supported special schools. In data tables included 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. Definitions of disability categories and criteria and methods for assigning students to them vary from state and to state and even between districts within states. Because we have relied on category assignments made by schools and districts, SEELS 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 disability by their school or district. Hence, descriptive data are nationally generalizable to students in the SEELS age range who were classified as having a particular disability in the 1999-2000 school year.

Measuring course grades. Teacher grades are a key dependent variable for the academic performance outcome domain discussed in Chapter 4 and is an independent variable used in analyses of some other outcomes. As a dependent variable, grade information is taken from the parent interview. Respondents were asked to report students' overall grades on a 9-point scale (e.g., mostly As, mostly As and Bs, mostly Bs, etc.). For students with no parent interview, teachers of general or special education classes were asked to report students' grades in their classes on the same 9-point scale. Data were used for the setting in which students take the most classes. Only students who receive this kind of letter grade are included in the analysis of this outcome measure.

Parents and teachers also were given an option of reporting qualitative indicators of student performance (e.g., excellent, good, fair, poor, or passing/not passing) if students do not receive traditional letter grades. When grades are used as an independent variable, it was considered important to include all students, including both those who receive letter grades and those who receive grades that are measured on a qualitative scale. Thus, the letter grade metric and various qualitative metrics needed to be combined. To do so, a 4-category variable was created. Letter grades from the 9-point scale were collapsed as indicated in the first column of Exhibit A-9. The corresponding qualitative grades appear in the second column.

---

improvement is 24% x 2 = 48%. It can be shown mathematically that this is the same value as would be obtained by using the formula for PI given above.

Exhibit A-9

CORRESPONDENCE OF LETTER AND QUALITATIVE GRADES IN CONSTRUCTING A COMPOSITE GRADE VARIABLE

Letter Grades	Qualitative Grades
Mostly As/Mostly As and Bs	Excellent
Mostly Bs/Mostly Bs and Cs	Good
Mostly Cs/Mostly Cs and Ds	Fair
Mostly Ds/Mostly Ds and Fs/Mostly Fs	Poor/Unsatisfactory/ Failing

Note that grades reported as "needs improvement", "satisfactory," or "passing" were not included in the analyses because their correspondence to a letter grade category was not clear.

Measuring motivation for schooling. This outcome is presented as a measure of engagement in Chapter 3. The student interview portion of the direct assessment includes a series of seven semantic differential items from the Motivation for Schooling subscale from the School Attitude Measure (Wick, 1991). The SAM includes different sets of items for students in the age groups 6 and 7 years, 8 and 9 years, 10 and 11 years, 12 and 13 years and 14 years or older. The response categories for the 6- and 7-year-old group were dichotomous, with 0=no and 1=yes. For the remaining age groups, the response categories were as follows: 1=never agree; 2=sometimes agree; 3=usually agree; and 4=always agree. To create a common motivation for schooling variable across the age groups, dichotomous responses for the 6- and 7-year-olds were recoded into the following categories so that 0 (no)=1 (never agree) and 1 (yes)=4 (always agree). The scale includes the following items common across age groups:

- I am happiest when I am at school
- School is the best place for me to learn
- Mondays are great because I get to come back to school
- School will help me have a better life
- Going to school is not boring for me
- I am excited about school and look forward to it
- I am looking forward to several more years of school

A scale was created by summing values on these items, which ranges from 7 (all responses "never agree") to 28 (all responses "always agree").

Measuring mobility for students with visual impairments. This outcome is presented as part of the discussion of independence in Chapter 6. The student's school program survey included series of 10 items to be completed by respondents for all students with a visual impairment as either their primary or a secondary disability. With advice from experts in the mobility of those with visual impairments, items were selected from the teacher checklist for orientation and mobility used at the Texas School for the Blind and Visually Impaired. Respondents indicated whether students could do the following "very well," "pretty well," or "not very well":

- Travel using sighted guide to familiar locations

- Travel indoors using rotely learned routes
- Travel to other areas using rotely learned routes
- Create new routes between familiar places indoors
- Execute route within building w/verbal directions
- Execute route in another building w/directions
- Locate unfamiliar place by numbering systems
- Orient oneself to unfamiliar room
- Solicit help to orient oneself to a building
- Solicit help to orient oneself to the school campus or a workplace.

A scale was created by summing values on these items, which ranges from 10 (all tasks done "not at all well") to 30 (all tasks done "very well").

Measuring locus of control. This outcome is also presented in Chapter 6. The student interview portion of the direct assessment included a series of four semantic differential items from the Locus of Control subscale from the School Attitude Measure (Wick, 1990). As noted regarding the measure of motivation for schooling, the SAM includes different sets of items for students in different age groups. The dichotomous response categories for the 6- and 7-year-olds were recoded so that 0 (no)=1 (never agree) and 1 (yes)=4 (always agree). The scale included the following items common across age groups:

- Most things I do at school turn out wrong
- A student like me will not get good grades
- I have no control over the grades I get
- I don't know how to do better in school

A scale was created by summing values on these items, which ranges from 4 (all responses "never agree") to 16 (all responses "always agree").

Comparisons with the general population of students. Many of the analyses reported here do not have precise statistical comparisons with the general population of students. Instead, we usually have drawn comparisons using published data. For many of these comparisons, differences in samples (e.g., ages of students) or measurement (e.g., question wording on surveys) reduce the direct comparability of SEELS and general population data. Where these limitations affect the comparisons, they are pointed out in the text and the implications for the comparisons are noted. Comparisons using data from the National Household Education Survey (NHES) are more precise because an analysis file was created from the publicly available data to match the age of SEELS students.

APPENDIX A REFERENCES

Asher, (1984). Loneliness in children. *Child Development, 55*(4), 1456-1464.

Bruininks, R. H., Woodcock, R. W., Weatherman, R. F., & Hill, B. K. (1996). *Scales of independent behavior-revised*. Chicago: Riverside Publishing.

Gresham, F. M., & Elliot, S. N. (1990b). *Student self concept scale.* Circle Pines, MN: American Guidance Service.

Lambert, N., Nihira, K., & Leland H. (1993). *AAMR adaptive behavior scales-school* (ABS-S:2). Austin, TX: Pro-Ed.

Wagner, R., Torgeson, J., & Rashotte, C. (1999). *Comprehensive test of phonological processing.* Austin, TX: Pro-Ed.

Marston & Deno, (1986). *Standard Reading Passages.* Minneapolis, MN: Children's Educational Publishing.

Wick, J. (1990). *School attitude measure.* Iowa City, IA: American College Testing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

Office of Special Education Programs (OSEP). (2001). *21st annual report to congress on the implementation of the Individuals with Disabilities Education Act.* Washington, DC: Author.