Appendix A
SEELS SAMPLING, DATA COLLECTION, AND ANALYSIS PROCEDURES: WAVES 1 & 2

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, school questionnaire, and assessment procedures and response rates
- Weighting of the SEELS 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.

Because of the need to include a large urban LEA whose rosters were received to late participate in CATI process, 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

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

December 2000 through March 2001. The questionnaire contained a subset of key items from the telephone interview.

This process was repeated for Wave 2 from April through July, 2002. The paper questionnaire was not required for Wave 2. In terms of response rates, of 11,512 eligible respondents in Wave 1, 9,824 interviews/questionnaires were completed (85% response). In Wave 2, 7,126 interviews were completed from an eligible pool of 9,475 (75% response).

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, 1997).
- 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 for Wave 1, and again from March to August, 2002 for Wave 2.

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.

In Wave 1, 4,912 completed standard or alternate assessments were returned for 7,806 eligible sample members (63% response). In Wave 2, 5,963 completed standard or alternate assessments were returned for 8,095 eligible sample members (74% 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.

For Wave 2, this process was repeated from March through August of 2002. There were several changes made in Wave 2. Teachers were provided a $5 incentive for returning the teacher or school program questionnaires. In addition, paid school coordinators were hired at the school sites to facilitate the data collection.

In Wave 1, 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). In Wave 2, completed teacher surveys were returned for 5,733 out of 9,775 eligible sample members (59% response), and completed school program surveys were returned for 5,789 out of 9,775 eligible sample members (59% response).

Weighting SEELS 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-1
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:

$$\frac{(P_1P_2)^2}{SE_1^2 + SE_2^2} > 1.96^2$$

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.

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.

Exhibit A-2

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").

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.