Article

Exceptional Children
2015, Vol. 81(3) 370–387

DOI: 10.1177/0014402914563695
ec.sagepub.com
SAGE

Once, Sometimes, or Always in Special Education: Mathematics Growth and Achievement Gaps

Ann C. Schulte[1] and Joseph J. Stevens[2]

Abstract

This study used a statewide longitudinal sample to examine mathematics achievement gaps and growth in students with and without disabilities and to examine the impact of different methods of determining disability group membership on achievement gaps and growth. When disability status was determined on the basis of special education placement each year, the achievement gap was larger across grades than when the subgroup of students with disabilities (SWDs) was defined more broadly, including students who had exited special education or who were in special education anytime between Grades 3 and 7. Regardless of the identification criteria, the SWD subgroup showed lower average achievement and slower growth than students without disabilities. The results suggest that the present way of identifying the SWD subgroup in reporting achievement outcomes may be biased and that even students who exit special education continue to be at risk for lower mathematics achievement.

The academic achievement of students with disabilities (SWDs) has been a long-standing concern (McDonnell, McLaughlin, & Morison, 1997; Wagner, Newman, Cameto, & Levine, 2006). For example, on the 2013 National Assessment of Educational Progress, students with disabilities scored more than one standard deviation below their peers without disabilities in Grades 4, 8, and 12 in reading and almost a standard deviation below these same peers in mathematics in Grade 4, with a larger gap of over one standard deviation in Grades 8 and 12 (U.S. Department of Education, Institute of Education Sciences, 2014).

No Child Left Behind and Disaggregated Achievement Reports

In response to concerns about this persistent low achievement, the No Child Left Behind Act of 2001 (NCLB; 2006) requires states, districts, and schools to report annual achievement test score results for students with disabilities separately as well as to include these students' scores in aggregate reporting. Disaggregated reporting also is required along three additional dimensions where specific subgroups have been historically at risk for low achievement: ethnic-racial identity, economic disadvantage, and low English proficiency. NCLB also set annual achievement targets, termed *adequate yearly progress* (AYP). AYP goals require students in all subgroups to reach 100% proficiency in reading and mathematics by 2013–2014 and mandated a "conjunctive approach" (Koretz & Hamilton, 2006) to determining success in meeting these

[1]Arizona State University
[2]University of Oregon

Corresponding Author:
Ann Schulte, T. Denny Sanford School of Social and Family Dynamics, Arizona State University, P.O. Box 873701, Tempe, AZ 85287-3701, USA.
E-mail: ann.schulte@asu.edu

targets, meaning failure of one subgroup to meet AYP in a year would result in overall failure for the state, district, or school.

Of the subgroups targeted for disaggregation, the SWD subgroup has proved the most challenging in terms of meeting AYP targets. For example, Eckes and Swando (2009), in a study of three states, found that schools most often failed to make AYP because of low achievement for the SWD subgroup. A number of concerns have been raised about NCLB policies relative to this subgroup (e.g., Eckes & Swando, 2009; Wei, Blackorby, & Schiller, 2011; Wei, Lenz, & Blackorby, 2013). Among these is the concern that SWDs are the only subgroup where the basis for membership can include cognitive limitations that make it difficult for some students in this subgroup to reach grade-level proficiency. Another is that SWDs, on average, start out with lower test scores than students in other at-risk subgroups, making it less likely that the subgroup will be able to reach grade-level proficiency standards in the time frame originally required by NCLB.

In addition to concerns about whether the ambitious targets set by NCLB for SWDs can actually be achieved, Ysseldyke and Bielinski (2002) raised concerns that identifying the SWD subgroup on the basis of annual participation in special education masks true gains in achievement for this subgroup. They argued that this approach results in a downward bias for SWD subgroup results, because each year students who are academically successful "exit from" special education, whereas students who are experiencing academic difficulty in general education "enter" special education. Using large-scale assessment data from one state, Ysseldyke and Bielinski reported special education turnover rates of approximately 20% per year and found that achievement gaps between SWDs and non-SWDs were smaller when a stable subgroup of SWDs, defined on the basis of special education membership at one point in time, rather than on a year-to-year basis, was examined.

Another complication with the cross-sectional or year-to-year approach for characterizing the achievement gap between SWDs and non-SWDs is that specific exceptionality groups enter and exit special education at different grades. As such, the mix of mild, moderate, and severe disabilities may change from grade to grade. For example, there is a much higher prevalence of students with speech or language impairments, one of the disability groups whose achievement is most similar to non-SWDs (Puranik, Petscher, Al Otaiba, Catts, & Lonigan, 2008; Wei et al., 2013), in early elementary school than in middle or high school (U.S. Department of Education, Office of Special Education Programs, 2013). Other disabilities (e.g., specific learning disability) are more likely to be diagnosed in later elementary school, in part because of the definitional criteria and in part because these children's difficulties may become apparent only as the expectations for content mastery rise (Francis, Shaywitz, Steubing, Shaywitz, & Fletcher, 1996). Although this complication does not affect year-to-year comparisons for the same grade (e.g., third grade in one year to third grade in the next year), changes in the composition of the SWD subgroup across grades have the potential to distort portrayals of changes in the size of the achievement gap across grades for SWDs.

In recent years, some states have modified their basis for defining the SWD subgroup by including students who have exited special education along with those currently served in special education in the SWD subgroup. For example, starting in 2009, North Carolina allowed schools to include students who had exited special education up to 2 years earlier in their SWD subgroup reporting (North Carolina Department of Public Instruction, 2011). This modification addresses some of the concerns about the bias introduced by reporting outcomes only for SWDs who are receiving services in the current year but does not fully address the impact that changes in SWD subgroup composition across grades may have on the achievement gap.

Growth Models and Their Implications for the SWD Subgroup

As concerns have been raised about the NCLB requirements and procedures for the SWD subgroup, others have criticized NCLB's "status" model (Castellano & Ho, 2013). With this model, schools are held to a uniform set of grade-level proficiency targets for all students, regardless of their achievement level when they enter the grade. Among the concerns cited are the status model's fairness and its validity in representing schools' true influence on students (e.g., Marion et al., 2002; Stevens, 2005). These concerns have led to interest in reshaping states' accountability systems to focus on student growth rather than attainment of a uniform set of grade-level proficiency targets (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, 2011). In response, researchers have begun to examine achievement growth longitudinally, seeking to understand the "natural developmental progress" that typifies children's grade-to-grade progress through school (e.g., Bloom, Hill, Black, & Lipsey, 2008 p. 298; Lee, 2010). An understanding of typical development in an academic area, and the extent to which individual or subgroup deviations from the typical trajectory are observed, is important in setting ambitious, but "scientifically based," growth targets for school accountability systems. Knowledge of growth patterns also can be helpful in judging whether specific curricula or interventions have a meaningful impact on achievement by situating effect sizes within the context of typical achievement growth for that age, grade, and subgroup (Bloom et al., 2008).

A substantial body of longitudinal studies of achievement growth in reading and mathematics is beginning to accumulate (e.g., Lee, 2010; Shin, Davison, Long, Chan, & Heistad, 2013). These studies generally depict growth in both subject areas as rapid in the early grades and diminishing as grade level increases (Bloom et al., 2008; Lee, 2010; Shin et al., 2013). However, accurate depictions of achievement growth for the SWD subgroup present a number of challenges. First, the corpus of longitudinal studies of SWD achievement growth across grades is quite limited. With a few exceptions (e.g., Shin et al., 2013; Wei et al., 2011, 2013), most studies of growth in the SWD population have examined students in one or two disability classifications only, primarily students with speech or language impairments or learning disabilities (e.g., Francis et al., 1996; Judge & Watson, 2011; Morgan, Farkas, & Wu, 2011). Second, the existing studies that examine achievement growth in SWDs have varied on multiple dimensions, including the age or grade range studied, whether a typically achieving comparison group was included, and whether demographic characteristics were included as predictors along with disability status. These studies generally depict SWDs with lower achievement than students without disabilities initially, with the achievement gap remaining stable or growing larger over time (Judge & Watson, 2011; Wei et al., 2013). In addition, SWDs follow the same "growth-to-plateau" pattern over grades observed in students without disabilities (Francis et al., 1996; Morgan et al., 2011; Wei et al., 2013). However, the paucity of studies and lack of consistency in methodology limit the confidence with which generalizations about achievement growth for SWDs can be made.

Shifting classifications complicate the interpretation of longitudinal studies of achievement growth for students with disabilities....

Last, just as the unstable nature of disability classification affects portrayals of SWD achievement within NCLB (Ysseldyke & Bielinski, 2002), shifting classifications complicate the interpretation of longitudinal studies of achievement growth for SWDs as well. Given that the distribution of different disability categories varies depending on the age when disability membership is determined, portrayals of SWD achievement growth and achievement gaps across grades may depend on the age or grade when group membership is determined. In addition, the fact that a student may enter or exit special education any time during the course of a longitudinal study raises the question of how to treat students who transition into or out of special education following their initial assignment into the focal group (SWD) or the comparison group (non-SWD).

Judge and Watson (2011) explicitly examined whether time of SWD identification affected depictions of mathematics achievement levels and growth in a longitudinal study of students with learning disabilities (LD) from kindergarten to fifth grade. They divided students with LD into three groups depending on when they were identified, comparing early-emerging (kindergarten and first grade), emerging (second and third grade), and late-emerging (fourth and fifth grade) LD groups. They found that all LD groups had lower mathematics achievement in kindergarten than the non-LD comparison group and that the mathematics achievement gap increased over time for each of the LD groups due to slower growth relative to the growth in the non-LD comparison group. Differences in growth among the three LD groups were minimal.

In contrast, findings from Puranik et al. (2008) suggest that groups of students who differ in the timing of their exit from special education may differ in achievement. They compared within-year growth in oral reading fluency in Grades 1 through 3 for students who were identified with speech or language impairments in first grade and then exited from special education and students with speech or language impairments who remained in special education in Grades 2 or 3. Students with language impairments who remained in special education had markedly lower growth in oral reading fluency than students with language impairments who had exited special education. This distinction was less important for the students with speech impairments, who showed oral reading fluency growth and levels similar to the nondisabled comparison group, regardless of whether they had exited from special education or continued to receive services in second and third grade.

In sum, SWDs are one of the targeted subgroups at risk for low achievement per NCLB. However, unlike relatively stable characteristics of students, such as gender or race or ethnicity, a student's disability status can change on an annual basis. Special education entrances and exits appear to be related to students' achievement, which means estimates of the achievement gap and descriptions of achievement growth for this population may not accurately represent the entire population of SWDs but only the subset who happen to be present in special education at the point in time when group membership was determined.

Research Questions

The purpose of the present study was to examine mathematics growth using three different longitudinal methods for identifying the SWD subgroup and to contrast achievement outcomes from these methods with the results obtained when SWD membership is determined on an annual basis, as specified in NCLB. The mathematics achievement of a statewide third-grade cohort of students was followed for 5 years (2001–2005). The three longitudinal methods were identification as an SWD on the first wave of measurement (i.e., third grade), identification as an SWD any time during Grades 3 to 7, and continuous identification as an SWD across the 5 years. These three longitudinal methods were contrasted with the cross-sectional approach used in NCLB. For each of the methods of identifying the SWD group, the total group of students remained the same. Differences in growth and achievement gap estimates were only a result of changes in which children were assigned to the SWD and non-SWD groups.

The specific research questions examined were as follows:

1. Do cross-sectional and longitudinal depictions of the mathematics achievement gap across grades differ?
2. Does the characterization of mathematics achievement growth change when the method for identifying the SWD subgroup varies?

Method

Study Database and Analytic Sample

The study data set represented all 2000–2001 students who were in third grade for the first time that year (i.e., not repeating third grade) and who were present in the state's large-scale achievement database (*N* = 103,123). To create an analytic sample that was appropriate for our research questions, we systematically excluded a number of individuals. Among those excluded were students who did not follow the typical grade-level sequence from Grades 3 to 7, largely due to grade retention (*n* = 8,315; 8.1%), students who never participated in the large-scale mathematics test in any of the Grades 3 to 7 (*n* = 1,729; 1.7%), and students with missing values on predictor variables. Eighteen students were missing an ethnicity code, 11 were missing gender codes, and 1,206 (1.2%) had no code for parental education. We also excluded 30 students coded as "other" ethnicity in 2001 because the category was dropped in subsequent years. Students in exceptionality categories with a sample size of 100 or less in third grade (multiple disability, orthopedically impaired, traumatic brain injury, and visually impaired; *n* = 321, <1%) and students with missing exceptionality codes for third grade (*n* = 314, <1%) also were excluded. After any student who met one or more of these exclusionary criteria was eliminated, the analytic sample comprised 92,045 students (89.3% of the original sample).

Students who were present in the database in 2000–2001 were matched to all succeeding years through Grade 7 (2004–2005). New students entering the system in future years were not added to the cohort (e.g., new Grade 4 students in 2001–2002, new Grade 5 students in 2002–2003). We tracked the cohort only through seventh grade because the state administered a new edition of the mathematics test in 2005–2006, and we wanted to avoid confounding mathematics growth with any changes in performance due to changes in test edition. We included any student who had at least one score on the mathematics general assessment during the study period. Of the 92,045 students in the analytic sample, 80.9% had mathematics scores in all 5 years, another 5.1% had scores in 4 years, 4.3% had scores in 3 years, 3.8% had scores in 2 years, and 6.0 % had only one mathematics score during the 5-year study period. Student test data could be missing because a student moved out of state or to a nonparticipating private school or because the student took an alternate assessment in that year.

Given the large sample sizes, changes of less than 1% in demographic characteristics were statistically significant. Therefore, we used Cohen's h, a measure of effect size for differences in proportions to examine the extent to which the two samples differed. Values of h ranged from .01 to .07; all would be considered small using Cohen's interpretive guidelines for this statistic (Cohen, 1988, pp. 184-185).

After creating the analytic sample, we divided the sample into students with and without disabilities (SWD and non-SWD) using four different decision rules for determining membership in the SWD group: (a) participation in special education determined for each year separately, as in NCLB (special education in current year); (b) participation in special education in third grade (special education at Wave 1); (c) participation in special education for at least one grade during the third- to seventh-grade span (ever in special education); and (d) participation in special education in all grades from 3 to 7 (always in special education). The non-SWD group for each method of determining SWD status consisted of any student who was not classified as an SWD based on the SWD criterion being applied. Table 1 depicts the demographic characteristics for the SWD and non-SWD groups with the three longitudinal methods of identifying SWD. The fourth identification method, special education in current year, resulted in a changing group of SWD in each grade. This group's demographic characteristics were identical to the demographic characteristics for the group in special education at Wave 1 in Table 1 (because in third grade, both groups were defined by special education participation in that grade). In subsequent grades, the demographic characteristics for the SWD and non-SWD subgroups in special education in the current year generally varied by only a small amount from the third-grade figures. (More detailed information on the grade-to-grade demographic changes when the SWD subgroup is defined annually is available from the authors at www.NCAASE.com.)

Table 1. Characteristics of Students With (SWD) and Without Disabilities (Non-SWD) With the Three Longitudinal Methods of Identifying SWDs.

	Basis for identifying SWD group					
	Special education at Wave 1		Ever in special education		Always in special education	
Characteristic	Non-SWD	SWD	Non-SWD	SWD	Non-SWD	SWD
n	81,179 (88.2)	10,866 (11.8)	77,079 (83.7)	14,966 (16.3)	86,497 (94.0)	5,548 (6.0)
Female	42,819 (52.7)	3,545 (32.6)	41,333 (53.6)	5,031 (33.6)	44,671 (51.6)	1,693 (30.5)
Ethnicity						
American Indian	1,166 (1.4)	187 (1.7)	1,077 (1.4)	276 (1.8)	1,258 (1.5)	95 (1.7)
Asian	1,720 (2.1)	71 (0.7)	1,668 (2.2)	123 (0.8)	1,765 (2.0)	26 (0.5)
Black	22,640 (27.9)	3,456 (31.8)	21,252 (27.6)	4,844 (32.4)	24,000 (27.7)	2,096 (37.8)
Hispanic	4,233 (5.2)	322 (3.0)	4,020 (5.2)	535 (3.6)	4,413 (5.1)	142 (2.6)
Multiracial	1,458 (1.8)	181 (1.7)	1,386 (1.8)	253 (1.7)	1,545 (1.8)	94 (1.7)
White	49,962 (61.5)	6,649 (61.2)	47,676 (61.9)	8,935 (59.7)	53,516 (61.9)	3,095 (55.8)
Limited English	2,555 (3.1)	169 (1.6)	2,368 (3.1)	356 (2.4)	2,649 (3.1)	75 (1.4)
Free/reduced lunch	31,744 (39.1)	5,522 (50.8)	29,729 (38.6)	7,537 (50.4)	34,137 (39.5)	3,129 (56.4)
Parental education level						
M (range 1–6)	3.1	2.6	3.1	2.6	3.1	2.4
SD	1.5	1.4	1.5	1.4	1.5	1.3
Cases with complete achievement data	67,273 (82.9)	7,172 (66.0)	63,966 (82.9)	10,586 (70.7)	70,605 (81.6)	3,840 (69.2)

Note. Percentages shown in parentheses.

Measures

Mathematics Achievement. For all analyses reported, the outcome measure was the standardized, second-edition North Carolina End of Grade (EOG) Test score in mathematics. The North Carolina EOG Mathematics Tests are multiple-choice tests designed to measure competency goals and objectives that are organized into four strands: (a) number sense, numeration, and numerical operations; (b) spatial sense, measurement, and geometry; (c) patterns, relationships, and functions; and (d) data, probability, and statistics (North Carolina Department of Public Instruction, 2006). There were a total of 80 items across these four areas at each grade, 30% calculator inactive, 70% calculator active, on the second edition of the test.

Scores from the North Carolina mathematics tests are reported as developmental scale scores and proficiency levels. The developmental scale is constructed using a three-parameter logistic item response theory model (Thissen & Orlando, 2001). To allow comparisons across years, the state has performed vertical linking using a common-items design in which adjacent grades are forward linked to create a common developmental scale that spans all grades.

Reliability and validity information for the EOG Mathematics Tests (second edition) is summarized in a technical report (North Carolina Department of Public Instruction, 2006). Internal consistency reliability estimates (coefficient alpha) for the EOG Mathematics Tests ranged from .94 to .96 for Grades 3 through 8. These estimates were the same when examined separately by gender, ethnicity, disability status, or limited English proficiency status. Evidence of both content and criterion-related validity includes analyses of how well items are aligned with the four content strands and correlations of student scores with teacher judgments of expected mathematics test scores and grades.

Student Background Variables. All student background variables were obtained from the test files used to construct the longitudinal database. Demographic information was coded onto answer sheets at the school level. Parental educational level was rated on an ordinal six-point scale where 1 was *did not finish high school* and 6 was *graduate school degree*.

Analytic Methods

The four methods of identifying the SWD subgroup were contrasted using two different approaches to describing and estimating mathematics growth. First, we used descriptive statistics to provide an overall picture of growth. We calculated the mean mathematics score at each grade for students with and without disabilities using each of the four methods of identifying special education students. As part of this descriptive analysis, we also examined the number of students participating in the general assessment in mathematics at each grade.

Second, we modeled student mathematics growth using multilevel, longitudinal analyses (Raudenbush & Bryk, 2002). We modeled mathematics growth only with the identification methods that produced a consistent group of SWDs across grades (omitting the method for current year in special education) and conducted three separate sets of multilevel analyses, with each set examining growth in special education students using one of the three remaining ways of identifying students as SWD (special education at Wave 1, ever in special education, and always in special education). Time was centered at the first occasion (Grade 3). Because the multilevel modeling approach to longitudinal analysis does not require data to be balanced within subjects (Raudenbush & Bryk, 2002), we included student data from any available time point even when data from other time points were missing.

For each of the three sets of multilevel analyses, we first applied an unconditional growth model followed by conditional models that added student disability status and student demographic variables as predictors. All multilevel analyses were conducted using HLM 7.0 (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011), full maximum likelihood estimation, and specification of model parameters as random effects. Both quadratic and cubic models showed good fits with the data. However, based on analyses reported elsewhere (Stevens, Schulte, Elliott, Nese, & Tindal, 2014), a quadratic model was determined to be the most parsimonious description of the data. The initial Level 1 model can be described as follows:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(\text{time}) + \pi_{2i}(\text{time squared}) + e_{ti}, \quad (1)$$

where Y is the mathematics scale score, π_{0i} is the initial status or intercept, π_{1i} is the initial linear rate of change, π_{2i} is the curvature in each student's growth trajectory, and e_{ti} is the residual for each student.

At Level 2, the Level 1 parameters (π_{0i}, π_{1i}, and π_{2i}) were modeled using student-level predictors. All dichotomous predictors were uncentered and the parental education variable that had multiple categories was grand mean centered. The Level 2 equations for the mathematics initial status and growth rate parameters were as follows:

$$\text{Initial Status}, \pi_{0i} = \beta_{00} + \Sigma\beta_{0k}(\text{Predictor}_i) + r_{0i}; \quad (2)$$

$$\text{Linear Change}, \pi_{1i} = \beta_{10} + \Sigma\beta_{1k}(\text{Predictor}_i) + r_{1i}; \text{ and} \quad (3)$$

$$\text{Curvature}, \pi_{2i} = \beta_{20} + \Sigma\beta_{2k}(\text{Predictor}_i) + r_{2i}, \quad (4)$$

where each Level 1 parameter was predicted by student disability status and demographic characteristics.

Results

Descriptive Statistics

Average Observed Growth for Students With and Without Disabilities. Table 2 provides the means and standard deviations of mathematics scale scores by grade for all students in the

Table 2. Mathematics Scale Score Means and Standard Deviations by Grade When Students With Disabilities (SWD) Identification Is Based on Special Education Status Determined Annually.

	Grade				
Basis for disability group inclusion	3	4	5	6	7
All students					
n analytic sample	92,045	86,654	82,994	80,741	78,754
n tested	90,017	84,529	82,153	80,296	78,135
% tested (of analytic sample)	97.80	97.55	98.99	99.45	99.21
M	251.44	257.39	263.22	267.15	270.15
SD	7.49	8.21	8.68	9.18	10.88
Special education in current year					
Non-SWD					
N analytic sample	81,179	75,928	72,785	71,317	69,976
N tested	80,045	75,536	72,738	71,265	69,875
% tested	99.05	99.48	99.94	99.93	99.86
M	251.99	258.04	264.13	268.26	271.33
SD	7.31	8.05	8.27	8.62	10.45
SWD					
n analytic sample	10,866	10,726	10,209	9,424	8,778
n tested	9,612	8,993	9,415	9,031	8,260
% tested	88.46	83.84	92.22	95.83	94.10
M	246.84	251.93	256.15	258.45	260.13
SD	7.46	7.47	8.58	8.79	9.13
Achievement gap effect size	−0.69	−0.74	−0.92	−1.07	−1.03

analytic sample and for the SWD and non-SWD groups when SWD status was defined annually as in NCLB. Table 3 provides the means and standard deviations of mathematic scale scores by grade for the SWD and non-SWD groups for each of the three methods of identifying SWDs that were stable across time. A complicating factor in characterizing growth in achievement across grades for SWDs is that an individual student's participation in the general or alternate assessment can vary from year to year (Chudowsky, Chudowsky, & Kober, 2009). For this reason, Tables 2 and 3 report the number of students present in each grade but also report the percentage of students (of those present that year) for whom scores on the general mathematics assessment were available. The percentage of students who were present who had test scores was large and quite stable across grades for the non-SWD group (98% to 99%) but more variable in the SWD subgroup, regardless of which identification method was used to identify who was in the SWD subgroup. The lowest percentage tested figures were for the SWD group always in special education (see Table 3). Using this SWD identification method, we found the percentage of students participating in the general assessment for mathematics ranged from 76% to 94% per year.

As is evident in Tables 2 and 3, the means for the non-SWD groups tracked each other closely across grades, regardless of how the SWD group was composed. In addition, the means for the SWD groups in special education at Wave 1 and ever in special education were quite similar across grades and higher than the other two methods for classifying SWD status. In contrast, after the 1st year (where the groups for current year in special education and special education at Wave 1 were the same children), the cross-sectional approach to determining SWD membership consistently resulted in lower mean mathematics achievement than the SWD groups in special education at Wave 1 and ever in

Table 3. Mathematics Scale Score Means, Standard Deviations, and Achievement Gap Effect Sizes Using Three Longitudinal Methods of Identifying Students With Disabilities (SWD).

	Grade				
Basis for disability group inclusion	3	4	5	6	7
Special education at Wave 1					
Non-SWD					
n analytic sample	81,179	76,491	73,263	71,266	69,586
n tested	80,045	75,948	73,155	71,170	69,453
% tested	99.05	99.29	99.85	99.87	99.81
M	251.99	257.95	263.98	268.07	271.13
SD	7.31	8.07	8.33	8.72	10.55
SWD					
n analytic sample	10,866	10,214	9,735	9,488	9,274
n tested	9,612	8,588	9,000	9,126	8,783
% tested	88.46	84.08	92.45	96.18	94.71
M	246.84	252.51	257.07	260.05	262.43
SD	7.46	7.84	9.02	9.57	10.33
Achievement gap effect size	−0.69	−0.66	−0.80	−0.87	−0.80
Ever in special education					
Non-SWD					
n analytic sample	77,079	72,406	69,321	67,414	65,834
n tested	76,381	72,038	69,282	67,358	65,747
% tested	99.09	99.49	99.94	99.92	99.87
M	252.29	258.26	264.32	268.46	271.57
SD	7.18	7.96	8.18	8.54	10.41
SWD					
n analytic sample	14,966	14,299	13,677	13,340	13,026
n tested	13,636	12,948	12,873	12,938	12,489
% tested	91.11	87.40	94.12	96.99	95.88
M	246.65	252.40	257.30	260.37	262.70
SD	7.42	7.83	8.90	9.39	10.22
Achievement gap effect size	−0.75	−0.71	−0.81	−0.88	−0.82
Always in special education					
Non-SWD					
n analytic sample	86,497	81,157	77,450	75,206	73,312
n tested	85,508	80,338	77,264	75,069	73,136
% tested	98.86	98.99	99.76	99.82	99.76
M	251.78	257.75	263.77	267.85	270.90
SD	7.36	8.10	8.41	8.80	10.60
SWD					
n analytic sample	5,548	5,548	5,548	5,548	5,548
n tested	4,509	4,198	4,891	5,227	5,100
% tested	81.27	75.67	88.16	94.21	91.93
M	245.01	250.55	254.53	257.18	259.37
SD	7.06	7.22	8.38	8.65	9.01
Achievement gap effect size	−0.90	−0.88	−1.06	−1.16	−1.06

special education. Figure 1 depicts the mean achievement across grades when SWD and non-SWD membership is determined at Wave 1 versus reconstituted annually.

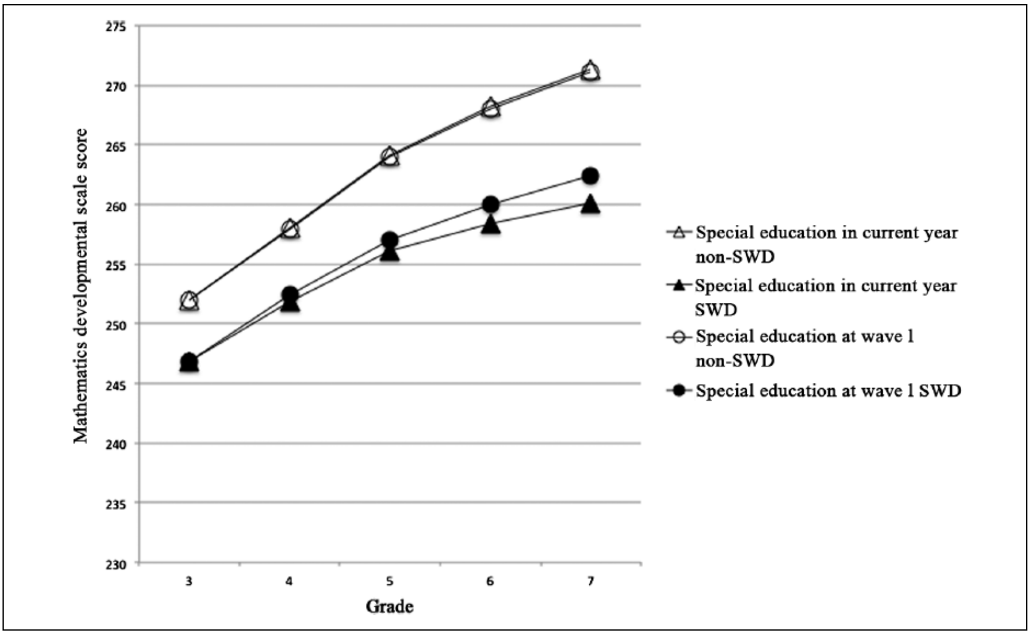

Figure 1. Observed mathematics means for students with disabilities (SWD) and without disabilities (non-SWD) when disability status is based on special education membership in third grade versus determined annually.

Achievement Gap Effect Sizes. Tables 2 and 3 also report achievement gap effect sizes, which describe mean differences between SWD and non-SWD groups at each grade divided by the standard deviation for all students at that grade. The achievement gap effect sizes based on cross-sectional, year-by-year identification of special education status depict the achievement gap widening across grades, increasing from −0.69 in third grade to more than a standard deviation by Grades 6 and 7 (see Figure 2). A somewhat different picture emerges by examining the three SWD groups where a stable cohort of students is identified over grades. The achievement gaps for the SWD groups in special education at Wave 1 and ever in special education tend to be smaller than the achievement gaps when SWD membership is determined annually and, after Grade 4, result in achievement gaps that are similar to each other and relatively stable over time. At all grades, the achievement gap is largest for the SWD group always in special education, and the gap widens to more than a standard deviation in Grades 5 to 7.

Multilevel Growth Models

The first step in the multilevel growth modeling analyses was the calculation of an unconditional random-effects growth model. Across all students, the estimated mean mathematics scale score in third grade was 251.08. The average initial linear change was a statistically significant growth rate of 6.92 points per grade ($z = 520.02$, $SE = 0.01$, $p < .001$). The curvature in the growth function was also significantly different from zero ($z = -163.89$, $SE = 0.003$, $p < .001$) and was estimated as an average deceleration in growth of −0.55 scale score points. Inspection of the variance components for each parameter (intercept, linear change, curvature) showed that there was significant variation in parameters across students for all three random-effect model parameters ($p < .001$).

We then applied a conditional model that added predictors for SWD status and student demographic and background characteristics. We repeated this analysis three times, each time changing only the criterion for identifying

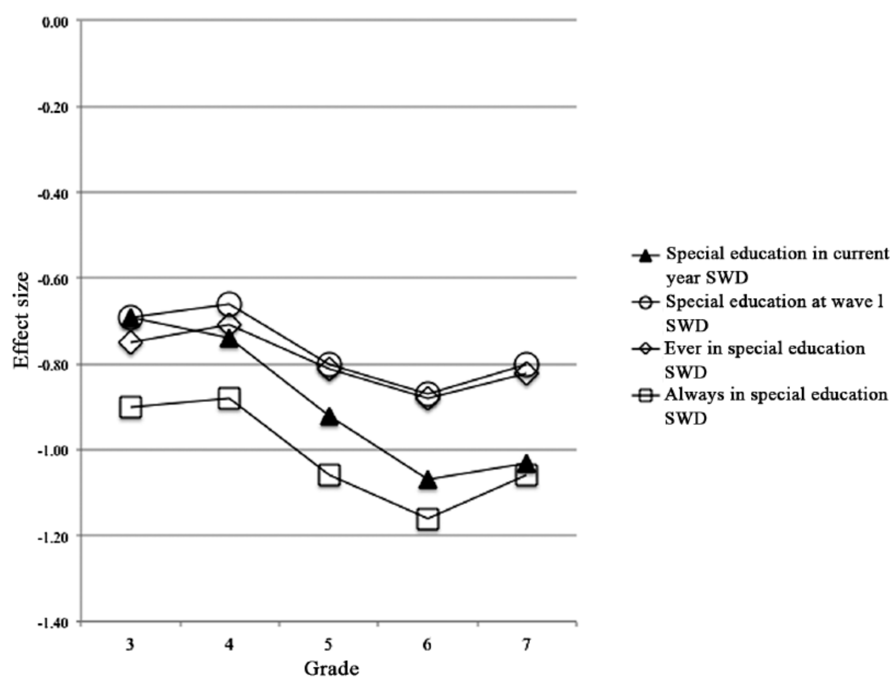

Figure 2. Achievement gap effect sizes for four different methods of identifying students with disabilities (SWD).

SWD (special education at Wave 1, ever in special education, or always in special education). The results for all three conditional analyses are presented in Table 4. The estimated intercept in each of the three models represents the average mathematics achievement in Grade 3 for White male students who were non-SWD (by the SWD identification criteria used in the particular model), not limited English proficient or receiving a free or reduced lunch, and whose parent education level fell at the grand mean. The differences in the intercept across the three analyses are the result of changes in the reference group because the composition of the non-SWD group changes somewhat with each SWD identification method.

The difference in intercept or initial status in Grade 3 between the non-SWD reference group (i.e., the group coded zero on all vectors) and SWD group was significant regardless of how the SWD group was identified ($p < .001$). The largest difference (−6.21) occurred when the SWD group was identified on the basis of continuous placement in special education (always in special education), meaning that after controlling for student demographic characteristics, being in the SWD group always in special education was associated with a 6.21-point lower score in mathematics in third grade. The intercept coefficients for special education status using the identification criterion of being in special education at Wave 1 or having ever been in special education were similar (−5.01 vs. −5.30). Given that the SWD group in special education at Wave 1 was composed exclusively of students who were in special education in third grade, and the SWD group ever in special education included students who had yet to be placed in special education, the similar intercept coefficients for the two groups suggest that the students identified after Grade 3 (who are the additional students ever in special education) were evidencing lower achievement prior to their placement in special education.

Table 4. HLM Models of Mathematics Growth Trajectories for Three Methods of Identifying Students With Disabilities, Grades 3 to 7.

	Basis for disability group inclusion								
	Special education at Wave 1			Ever in special education, Grades 3 to 7			Always in special education, Grades 3 to 7		
Predictor	Intercept	Linear	Quadratic	Intercept	Linear	Quadratic	Intercept	Linear	Quadratic
Grand mean	253.86 (.04)	6.97 (0.02)	−.52 (.01)	254.18 (0.04)	6.96 (0.02)	−.52 (.01)	253.55 (0.04)	6.96 (0.02)	−.53 (.01)
Special education	−5.01 (0.07)	−0.64 (0.05)	.01† (.01)	−5.30 (0.06)	−0.39 (0.04)	−.04 (.01)	−6.21 (0.09)	−0.92 (0.06)	.02† (.02)
Limited English	−2.90 (0.16)	−0.11† (0.11)	.03† (.03)	−2.72 (0.16)	−0.09 (0.11)	.03† (.03)	−2.85 (0.16)	−0.11† (0.11)	.03† (.03)
Parental education	1.53 (0.02)	0.04 (0.01)	.04 (.00)	1.50 (0.02)	0.04 (0.01)	.03 (.00)	1.55 (0.02)	0.04 (0.01)	.04 (.00)
Gender	−0.42 (0.04)	0.02† (0.03)	.04 (.01)	−0.58 (0.04)	0.03† (0.03)	.03 (.01)	−0.30 (0.04)	0.02† (0.03)	.04 (.01)
Free/reduced lunch	−1.43 (0.05)	−0.19 (0.03)	−.01† (.01)	−1.39 (0.05)	−0.20 (0.03)	−.01† (.01)	−1.46 (0.05)	−0.19 (0.03)	−.02† (.01)
Asian	0.47 (0.17)	1.27 (0.10)	−.07 (.03)	0.33† (0.17)	1.28 (0.10)	−.07 (.03)	0.61 (0.17)	1.28 (0.10)	−.07 (.03)
Black	−4.42 (0.05)	0.21 (0.03)	−.11 (.01)	−4.40 (0.05)	0.22 (0.03)	−.11 (.01)	−4.31 (0.05)	0.22 (0.03)	−.11 (.01)
Hispanic	−0.89 (0.13)	0.92 (0.08)	−.15 (.02)	−1.02 (0.12)	0.93 (0.08)	−.15 (.02)	−0.79 (0.13)	0.93 (0.08)	−.15 (.02)
American Indian	−1.99 (0.18)	−1.47 (0.11)	.29 (.03)	−1.90 (0.17)	−1.46 (0.12)	.29 (.03)	−1.97 (0.18)	−1.47 (0.12)	.29 (.03)

Note. HLM = hierarchical linear modeling. Standard errors shown in parentheses.
†Not statistically significant, $p > .05$.

In terms of demographic characteristics and their relationship to students' initial status, across all three models, students who were female, were limited English proficient, qualified for a free or reduced-price lunch, or were Black, Hispanic, or American Indian showed significantly lower levels of achievement ($p < .001$) than students in the reference group. Students who were Asian or whose parents had educational levels higher than average showed higher initial achievement in all three models, with one exception: When using the criterion of having ever been in special education for identifying SWDs, the initial achievement for Asian students did not differ significantly from the reference group.

In terms of slope, the reference group showed a significant rate of linear change of just less than seven points per year, regardless of how the SWD group was identified. In each analysis, SWDs showed lower growth in contrast to the reference group, with slope coefficients differing by more than a factor of two depending on the SWD identification criteria (ever in special education, −0.39; special education at Wave 1, −0.64; always in special education, −0.92). Linear growth also varied significantly by student demographic characteristics. Students who were Asian, Black, or Hispanic showed significantly higher initial growth compared to the reference group, whereas students who were receiving free or reduced lunch or who were American Indian showed lower growth on average. Coefficients for demographic characteristics were quite similar regardless of the criterion used to identify the SWD group.

For the quadratic term, the reference group in each analysis had a statistically significant deceleration in growth rate of about −0.5 scale score points per year. The SWD group deceleration did not differ from that of the reference group, except when the criterion of having ever been in special education was used. Then, the SWD group showed a small, but significant, moderation in deceleration. Moderation in the deceleration rate was also found across the three analyses for students whose parents had higher levels of education, for females, and for students who were American Indian compared to students in the reference group. Significantly more deceleration was observed for students who were Asian, Black, or Hispanic.

There was a substantial subset of students who were identified as general education students in some grades and special education students in other grades....

Discussion

The present study used a statewide longitudinal sample to examine mathematics gaps and growth in students with and without disabilities. Based on concerns that the present way of identifying the SWD subgroup in NCLB may not accurately characterize achievement gaps, and that single versus multiple time point definitions of disability status may yield different descriptions of growth for disability groups, we contrasted depictions of the achievement gap and growth for SWDs with four identification methods.

NCLB reporting typically identifies students as belonging to the SWD subgroup on a year-by-year basis. Across the span of Grades 3 to 7, this method identified approximately 12% of students each year as belonging to the SWD subgroup in this sample. This figure is slightly lower than national estimates of the percentage of students served in special education as a percentage of total school enrollment (approximately 13%; U.S. Department of Education, Office of Special Education Programs, 2012) but is comparable considering that students consistently taking the alternate assessment were not included in the sample.

When examined longitudinally, just over 16% of students in the present sample were identified as special education students at some time between third and seventh grade. Although there was a stable core of students who remained in special education across all grades, composing about 6% of all students in the present study, there was a substantial subset of students who were identified as general education students in some grades and special education students in other grades.

Major Findings

The first research question posed was whether cross-sectional or longitudinal depictions of the achievement gap across grades for the SWD subgroup differed. Regardless of whether the identification method used to compose the SWD subgroup was cross-sectional or longitudinal, a substantial gap between the achievement of SWDs and non-SWDs was observed at each grade level. However, the gap widened more across grades when SWD subgroup membership was allowed to vary by year as in NCLB. These findings reinforce concern over the under-achievement of the SWD subgroup (Wagner et al., 2006) but also suggest that alternate ways of portraying the achievement gap are needed to obtain an accurate picture of how students who are placed in special education fare as they progress through school, as suggested by Ysseldyke and Bielinski (2002). The

Cross-sectional portrayals of the achievement gap frequently reported in the research and policy literature may underestimate academic growth for SWDs.

Inferences about how children grow academically and how this growth is influenced by intrinsic and extrinsic factors, such as disabilities or school environments, are most likely to be accurate when they are made on the basis of examining change over time within individuals. Modeling the growth of a target group of students where the distinguishing characteristic, such as disability status, can change over time requires that the researcher or policy maker decide how to operationalize the target group. The second research question posed was whether the characterization of achievement growth changed when the method of identifying the SWD group was altered. In the multilevel modeling portion of this study, regardless of the identification criterion used to form the SWD group, growth was curvilinear across grades, and the intercept and linear slope coefficients for disability status were negative, meaning the SWD subgroup showed lower initial achievement and slower growth. The decelerating curvilinear growth pattern has been observed in most growth studies of mathematics achievement, for both students with and without disabilities (e.g., Lee, 2010; Wei et al., 2013). Although the three different SWD identification methods examined here all resulted in negative intercept and linear slope coefficients, the extent of the deficits varied depending on the SWD identification criterion used. As in the Puranik et al. (2008) study, students identified based on a continued need for special education services (always in special education) were the farthest behind and experienced the least growth across grades. Students identified on the basis of third-grade status (special education in Wave 1) or any time in Grades 3 to 7 (ever in special education) had higher and similar third-grade achievement and growth across grades.

These results add to the body of studies reporting slower mathematics achievement growth for students with disabilities (Judge & Watson, 2011; Morgan et al., 2011) and extend them by including both a broad range of exceptionalities and a comparison group of students without disabilities and by clarifying the ways in which depictions of growth may vary based on how consistently students are served in special education over time. By including both disability status and demographic characteristics in the growth models, this study also allowed the impact of disability status on achievement level and growth to be estimated while controlling for differences in demographic characteristics between students with and without disabilities. The study's findings relative to demographic characteristics were similar to those reported in previous research. For example, being female, Black, Hispanic, or eligible for free or reduced-price lunch all were associated with lower initial levels of mathematics achievement, as in previous studies (Shin et al., 2013; Wei et al., 2011). However, unlike some studies (e.g., Morgan et al., 2011; Wei et al., 2011), we observed more rapid mathematics growth for students who were Hispanic or Black. Most previous growth studies have used nationally normed mathematics tests; in contrast, the mathematic test employed in the present study was designed to align with the mathematics curriculum students received. It may be that closer curriculum–test alignment made the test more sensitive to differences in mathematics growth, or perhaps the large sample size in the present study provided more power to detect small differences in mathematics growth across grades.

Policy and Research Implications

The primary purpose of this study was to examine whether the way in which a disability subgroup is composed can affect portrayals of achievement gaps and growth. In both cases, the answer to the question is yes. Most notable in terms of the achievement gap was the difference in the size of the achievement gap when the SWD subgroup was identified with a method that tracked students from special education into general education (special education in Wave 1 and ever in special education), rather than including only students currently participating in special education. This finding suggests that alternate methods of identifying the SWD subgroup for NCLB or other public reporting of disaggregated test results should be considered that include tracking students after they leave special education. Adding to the argument for including students who have exited special education in the SWD subgroup is the fact that these students appear to continue to be at risk for low mathematics achievement and growth, as indicated by the achievement gaps and slower growth observed even when the SWD subgroup included students who had exited special education.

It is likely, however, that neither of the identification methods examined here that tracked SWDs as they entered general education would be acceptable alternatives for use in public reporting of subgroup outcomes as required by NCLB or other accountability programs that may be implemented in the future. The criterion of being in special education in Wave 1 would fail to include the substantial number of students who were identified for special education after Grade 3, and the criterion of ever having been in special education can be applied only retrospectively, once the full group of students who were identified across grades is known. However, this study does support the utility of exploring reporting alternatives that include outcomes for a more broadly defined SWD subgroup. For example, each year, SWD subgroup reporting might include all students identified for special education in the current and previous years or include students for a set time period after they exit special education, as is already done in some states.

[A] system that tracks students' growth longitudinally would provide...more information about whether individual students are responding to instruction....

It is important to note that regardless of how the SWD group was constituted, substantial achievement gaps were observed between SWDs and non-SWDs at each grade, as was slower growth for the SWD subgroup in all of the growth analyses. Given that closing the achievement gap would require more rapid growth in mathematics achievement for the SWD subgroup, these findings suggest that substantial changes may be needed in the type or intensity of mathematics instruction for SWD if the achievement gap is to be narrowed or closed. What types of instructional changes will produce accelerated mathematics growth for SWDs, particularly given the transition to the rigorous Common Core State Standards (National Governor's Association Center for Best Practices & Council of Chief State School Officers, 2010) is not known, although previous research suggests that early intervention, explicit instruction, and development of foundational skills are likely to be fruitful strategies for enhancing mathematics outcomes for SWDs in both general and special education settings (Gersten et al., 2009; Powell, Fuchs, & Fuchs, 2013).

The finding of slower growth in the SWD subgroup suggests that it is likely the SWD subgroup would continue to underperform in mathematics relative to non-SWDs even if growth-based accountability models replaced NCLB's (2002) status model. However, a system that tracks students' growth longitudinally would provide school staff and policy makers with more information about whether individual students are responding to instruction or falling further behind their nondisabled peers as they progress across the grades.

Although the number of longitudinal research studies examining the mathematics and reading achievement growth of SWDs is increasing (e.g., Shin et al., 2013; Wei et al., 2011, 2013), these studies have varied in how the disabilities subgroup was composed. By holding the group of students constant, but varying the criteria for determining which students are considered to exhibit a disability, this study suggests that some of the variation found in growth study results for SWDs may be a result of the age at which students are identified or the stringency of the inclusion criteria. More research is needed that explicitly examines how achievement growth estimates for groups whose composition may change by age or grade are affected by the identification methods used in forming the focal and comparison groups. In addition, the choice of the age or grade at which group membership is determined should be an explicit aspect of longitudinal study design.

Limitations

As with any study, there are a number of limitations that must be taken into account when the study results are considered. First, a number of students were excluded from the analytical sample, including students who participated in alternate assessments across all grades, were retained, or were missing demographic information. The deletions on the basis of grade retention and lack of participation in the general assessment are likely to have resulted in a sample that was achieving at a higher level than a complete sample of SWDs. Second, this study made use of large-scale assessment data from only one state rather than a nationally representative sample. Identification criteria for exceptionality categories and SWD prevalence vary by state, which makes the extent to which these results generalize to outcomes for the SWD subgroup with other state assessment systems unknown. Third, this study examined only mathematics achievement, and reading achievement results may differ. Finally, differences in growth model results were not tested directly due to the overlap in the students identified with the three longitudinal methods. Although the sample size offers some assurance that the differences among SWD subgroup results found in the multilevel growth modeling analyses were robust, future studies should be designed to allow direct tests of these comparisons and replication with data from other states.

In sum, this study is a first step in examining the complexity of addressing questions about the growth of SWDs. Continued efforts to move beyond the "fuzzy data" (Chudowsky et al., 2009, p. 1) presently available about outcomes for SWDs to accurate depictions of their achievement and growth are essential to close the achievement gap. Such efforts will also assure that schools have accurate information about when their efforts to serve this population are successful or are falling short of the goal of having all students reach their full potential.

Authors' Note

Shortly before this issue of Exceptional Children went to press, the United States Department of Education announced the creation of Results-Driven Accountability (RDA). Under the RDA, the Department will also include educational results and outcomes for students with disabilities in making each state's annual determination under the Individuals with Disabilities Education Act (IDEA).

References

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*, 289–328. doi:10.1080/19345740802400072

Castellano, K. E., & Ho, A.D. (2013). *A practitioner's guide to growth models*. Washington, DC: Council of Chief State School Officers. http://scholar.harvard.edu/files/andrewho/filesa_pracitioners_guide_to_growth_models.pdf

Chudowsky, N., Chudowsky, V., & Kober, N. (2009). *State test score trends through 2007–2008, Part 4: Has progress been made in raising achievement for students with disabilities?* Washington, DC: Center on Education Policy.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Eckes, S., & Swando, J. (2009). Special education subgroups under NCLB: Issues to consider. *Teachers College Record*, *111*, 2479–2504.

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, *88*, 3–17. doi:10.1037/0022-0663.88.1.3

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*, 1202–1242. doi:10.3102/0034654309334431

Judge, S., & Watson, S. M. R. (2011). Longitudinal outcomes for mathematics achievement for students with learning disabilities. *Journal of Educational Research*, *104*, 147–157. doi:10.1080/00220671003636729

Koretz, D. M., & Hamilton, L. S. (2006). Testing for accountability in K–12. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 531–578). Westport, CT: American Council on Education and Praeger.

Lee, J. (2010). Tripartite growth trajectories of reading and math achievement: Tracking national academic progress at primary, middle, and high school levels. *American Educational Research Journal*, *47*, 800–832. doi:10.3102/0002831210365009

Marion, S., White, C., Carlson, D., Erpenbach, W. J., Rabinowitz, S., & Sheinker, J. (2002). *Valid and reliable decisions in determining adequate yearly progress*. Washington, DC: Council of Chief State School Officers.

McDonnell, L. M., McLaughlin, M. J., & Morison, P. (Eds.). (1997). *Educating one and all: Students with disabilities and standards-based reform*. Washington, DC: National Academy Press.

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children's growth trajectories in reading and mathematics: who falls increasingly behind? *Journal of Learning Disabilities*, *44*, 472–488. doi:10.1177/0022219411414010

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards mathematics*. Washington, DC: Author.

No Child Left Behind (NCLB) Act of 2001, 20 U.S.C. §§ 6301 *et seq*. (2006 & Supp.V.2011).

North Carolina Department of Public Instruction. (2006). *The North Carolina Mathematics Tests: Technical report*. Raleigh, NC: Author. Retrieved from http://www.ncpublicschools.org/docs/accountability/testing/mathtechmanual.pdf

North Carolina Department of Public Instruction. (2011). *Determining adequate yearly progress (AYP) 2010–11*. Raleigh, NC: Author. Retrieved from http://www.ncpublicschools.org/docs/accountability/reporting/abc/2010-11/aypstatus.pdf

Powell, S. R., Fuchs, L. S., & Fuchs, D. (2013). Reaching the mountaintop: Addressing the Common Core Standards in mathematics for students with mathematics difficulties. *Learning Disabilities Research & Practice*, *28*, 38–48. doi:10.1111/ldrp.12001

Puranik, C. S., Petscher, Y., Al Otaiba, S., Catts, H. W., & Lonigan, C. J. (2008). Development of oral reading fluency in children with speech or language impairments: A growth curve analysis. *Journal of Learning Disabilities*, *41*, 545–560. doi:10.1177/0022219408317858

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: SAGE.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Shin, T., Davison, M. L., Long, J. D., Chan, C., & Heistad, D. (2013). Exploring gains in reading and mathematics achievement among regular and exceptional students using growth curve modeling. *Learning and Individual Differences*, *23*, 92–100. doi:10.1016/j.lindif.2012.10.002

Stevens, J. (2005). The study of school effectiveness as a problem in research design. In R. Lissitz (Ed.), *Value-added models in education: Theory and applications* (pp. 166–208). Maple Grove, MN: JAM Press.

Stevens, J. J., Schulte, A. C., Elliott, S. N., Nese, J. F. T., & Tindal, G. (2014). Mathematics achievement growth of students with and without disabilities on a statewide achievement test. Journal of School Psychology. http://dx.doi.org/10.1016/j.jsp.2014.11.001

Thissen, D., & Orlando, M. (2001). Item response theory for items scored in two categories. In D. Thissen & H. Wainer (Eds.), *Test scoring* (pp. 73–140). Mahwah, NJ: Lawrence Erlbaum.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2013). *National Assessment of Education Progress data explorer*. Rockville, MD: Author. Retrieved from http://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2011). *Final report on the evaluation of the growth model pilot project*. Washington, DC: Author. Retrieved from http://www2.ed.gov/rschstat/eval/disadv/growth-model-pilot/index.html

U.S. Department of Education, Office of Special Education Programs. (2012). *31st annual report to Congress on the implementation of the Individuals With Disabilities Education Act, 2009*. Washington, DC: Author.

U.S. Department of Education, Office of Special Education Programs. (2013). *Data accountability center, IDEA data explorer*. Rockville, MD: DAC/Westat. Retrieved from https://www.ideadata.org/DACAnalyticTool/Intro_2.asp

Wagner, M., Newman, R.S., Cameto, R., & Levine, P. (2006). *The academic achievement and functional performance of youth with disabilities: A report from the National Longitudinal Transition Study–2 (NLTS2)*. Menlo Park, CA: SRI International.

Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, *78*, 89–106.

Wei, X., Lenz, P. B., & Blackorby, J. (2013). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioecononmic status differences from ages 7 to 17. *Remedial and Special Education*. Advance online publication. doi: 10.1177/0741932512448253

Ysseldyke, J., & Bielinski, J. (2002). Effect of different methods of reporting and reclassification on trends in test scores for students with disabilities. *Exceptional Children*, *68*, 189–200.

Authors' Note

This work was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R32C110004 awarded to the University of Oregon. The opinions expressed are those of the authors and do not necessarily represent views of the institute or the U.S. Department of Education.

Manuscript received June 23, 2013; accepted March 7, 2014.