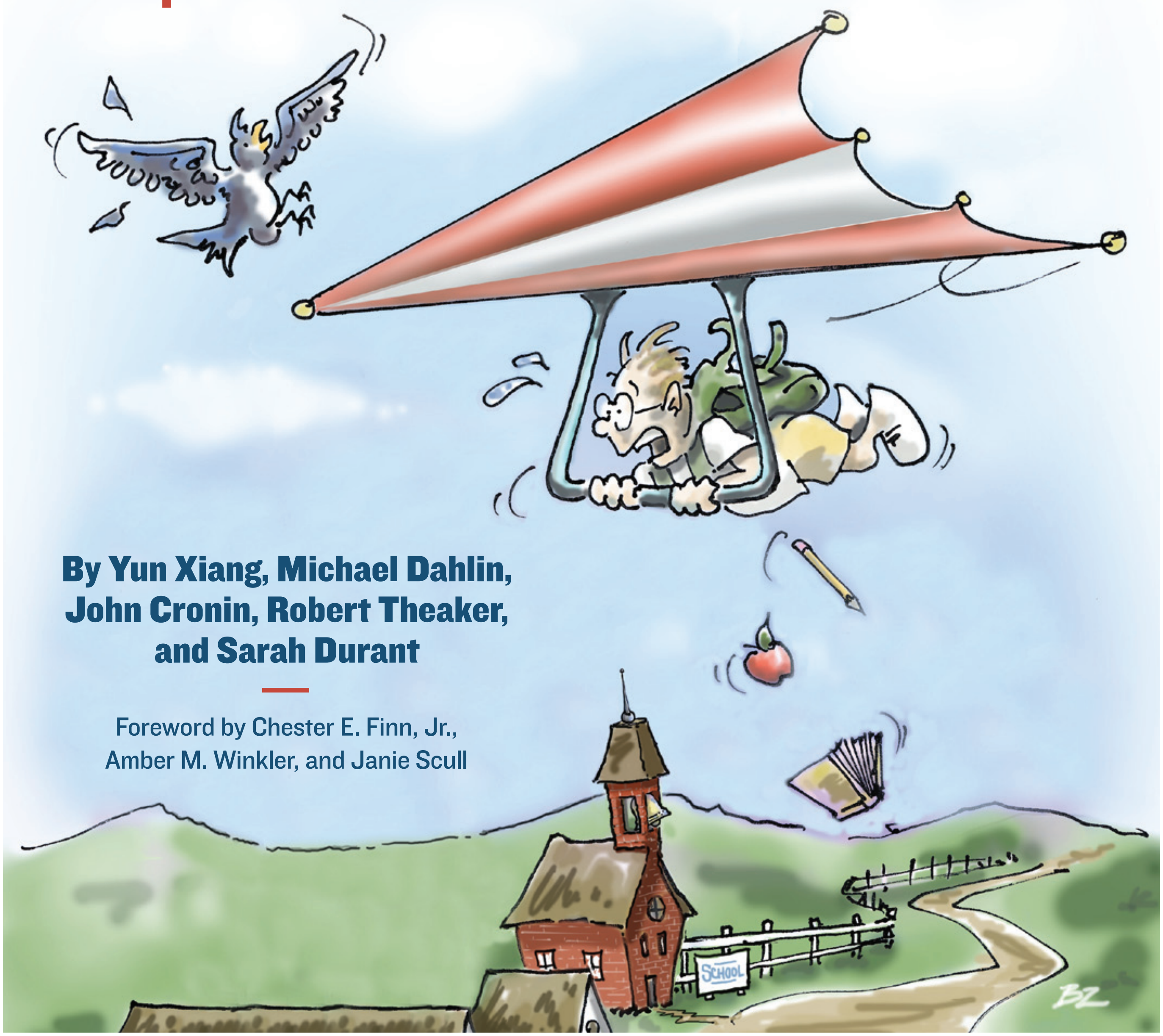
SEPTEMBER 2011

DO HIGH FLYERS MAINTAIN THEIR ALTITUDE?

Performance Trends of Top Students



By Yun Xiang, Michael Dahlin, John Cronin, Robert Theaker, and Sarah Durant

Foreword by Chester E. Finn, Jr., Amber M. Winkler, and Janie Scull

THOMAS B. FORDHAM INSTITUTE
Advancing Educational Excellence

NWEA
Northwest Evaluation Association

CONTENTS

The Thomas B. Fordham Institute is the nation's leader in advancing educational excellence for every child through quality research, analysis, and commentary, as well as on-the-ground action and advocacy in Ohio. It is affiliated with the Thomas B. Fordham Foundation, and this publication is a joint project of the Foundation and the Institute. For further information, please visit our website at www.edexcellence.net or write to the Institute at 1016 16th St. NW, 8th Floor, Washington, D.C. 20036. The Institute is neither connected with nor sponsored by Fordham University.

FOREWORD

By Chester E. Finn, Jr., Amber M. Winkler, and Janie Scull

This groundbreaking study is the first ever to examine the achievement of high-performing students over time *at the individual level*. It poses—and seeks to answer—this straightforward question: Do students who outscore their peers on standardized achievement tests remain at the top of the pack year after year? Put differently, how many "high flyers" maintain their "altitude" over time? How many fall back toward Earth as they make their way through school, losing the academic edge they once enjoyed?

The reason this question is important should be obvious: If America is to remain internationally competitive with other advanced nations, we need to maximize the potential of our top students. Yet many analysts worry that various policies and programs, including the federal No Child Left Behind Act (NCLB), tend to "level" student achievement by focusing on the lowest-achieving students and ignoring—or, worse, driving resources away from—our strongest students.

Thanks to previous Fordham research by Tom Loveless, we know that from 2000 to 2007 achievement for the highest-performing students (as measured by the National Assessment of Educational Progress) stagnated, while the lowest-performing students made significant gains (Duffett, Farkas, and Loveless, 2008). Other studies have corroborated this finding. But these earlier analyses examined different cohorts of students—eighth graders in 2000 versus eighth graders in 2007, for example—not individual students over time.

In the present study, analysts from the Northwest Evaluation Association™ (NWEA) examine achievement trends for students who scored extremely well on the NWEA assessment, known as the Measures of Academic Progress™ (MAP). If the schools these students attend are adequately challenging them to continue learning at high levels—and providing them the instruction they need to do so—one would logically expect most of them to maintain their high standing over time. On the other hand, if these youngsters are left to fend for themselves while attention and resources are showered on their lower-achieving peers, one might expect them to drop closer to average.

That said, it would be naïve to expect *all* high-achieving pupils to stay that way forever—just as it's naïve to expect 100 percent of students to reach "proficient." Some students, for example, might face life crises that would depress their

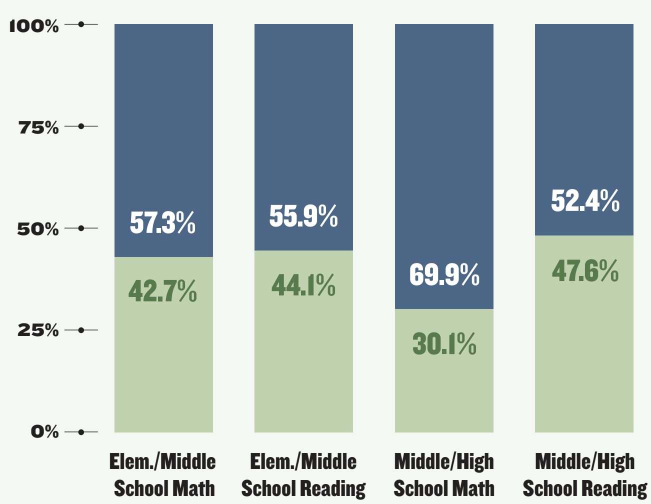


FIGURE ES-1

Outcomes of Initial High Flyers

Note: For each cohort and subject, the figure shows what percentage of the initial high achievers was still high-achieving in the final year of the study (Steady High Flyers), and what percentage was not (Descenders). Students were tracked from grades three to eight in elementary/middle school and grades six to ten in middle/high school. For example, 57.3 percent of the initial high math achievers in the elementary/middle school cohort remained high-achieving from third to eighth grade (i.e., they were Steady High Flyers), while 42.7 percent were no longer high-achieving by eighth grade (i.e., they were Descenders).

achievement, such as a divorce in the family or the loss of a loved one. Moreover, high achievement ought not be a zero-sum game, whereby some students are kicked out of their lofty rank to make room for others. (In this study, we define high achievers as those students who score at the 90th percentile or above according to external norms, but allow for as many students within the subset being tracked to enter those ranks as qualify to do so.) In the end, great schools should be able to take middling students and, over the course of several years, help them to near the head of the class.

How about those who were "high flyers" to start with? Some will surely lose altitude. But if a great many of them do so, this should set off alarm bells. It would be especially concerning if many high flyers within particular groups—say, girls, or minority children, or those attending high-poverty schools—faltered over time.

What did we learn?

Finding #1. A majority of high flyers maintained their status over time, but substantial numbers "lost altitude." Nearly three in five students identified as high-achieving in the initial year of the study remained high-achieving in the final year (Figure ES-1). Specifically, 57.3 percent of high-achieving third-grade math students remained that way by eighth grade, while 55.9 percent did so in reading (we call them "Steady High Flyers"). A full 69.9 percent of high-achieving sixth-grade math students remained high-achieving by tenth grade; 52.4 percent did so in reading. Of course, that also means that roughly 30 to 50 percent of students in the initial high-achieving group lost their top-tier academic status over time (we call them "Descenders").

That doesn't mean the pool of high-achieving students shrank. On the contrary, it grew, thanks to a greater number of students—we call them the "Late Bloomers"—who entered the high-achieving ranks over time. For instance, the percentage of high flyers in math at the elementary/middle school level grew from 12.4 percent of all third graders to 14.1 percent in eighth grade.[1]

Finding #2. Most Descenders don't fall far. Even those students who "lost altitude" (dropped below our 90th percentile cutoff) didn't glide too far from the high-achieving ranks. Most stayed at the 70th percentile or higher. And those who soared into the high-achieving ranks by eighth grade did not have far to fly. In math, for instance, they performed on average at the 74th percentile in third grade.

Interestingly, the proportions of minority students within the high-achieving group proved relatively stable (albeit smaller than one would hope) over time. In terms of gender, however, boys were overrepresented among high achievers in math and girls in reading, but high-achieving boys were more likely to lose ground in *both* subjects.

Finding #3. High flyers grew academically at similar rates to low and middle achievers in math, but grew at slightly slower rates than low and middle achievers in reading. We compared the growth rates of high-achieving students in reading and math in relation to middle achievers (those performing between the 45th and 54th percentiles, inclusive) and low achievers (those below the 10th percentile). In math, the gaps between high, middle, and low achievers changed very little over time. In reading, however, low- and middle-achieving students demonstrated faster rates of improvement than high achievers.

Finally, NWEA conducted a separate analysis that used a different definition of high achievement. Instead of using an external standard, they defined high-achieving students as those whose math or reading scores placed them within the top 10 percent *of their individual grades and schools*. In those analyses, NWEA researchers were primarily interested in examining the impact that school-level factors might have on student growth. They tracked both an elementary school cohort from third grade to fifth grade, and a middle school cohort from sixth grade to eighth grade.

1 The external benchmark for high-achieving status was set at the 90th percentile, but since the study sample and normed sample are independent, more than 10 percent of students could qualify in the former as high-achieving at either the beginning or ending point.

This analysis was exploratory due to a non-representative sample (plus analysts were not able to fully account for student mobility). Still, they uncovered a surprising and encouraging trend: **High-achieving students attending high-poverty schools made about the same amount of academic growth over time as their high-achieving peers in low-poverty schools.** For example, high flyers at low-poverty schools performed on average at the 97th percentile in third-grade math, while high flyers at high-poverty schools scored at the 83rd percentile—a difference representing over a year's worth of growth. By fifth grade, however, they scored at the 97th and 82nd percentiles, respectively. While high achievers in high-poverty schools grew slightly less than those in low-poverty schools, the difference was very small. It appears that the relationship between a school's poverty rate and the growth of its highest-achieving students is weak. In other words, attending a *low-poverty* school adds little to the average high achiever's prospects for growth.

Implications

What to make of all this? We see four takeaways.

First, many students maintained their high-flying status and many lost it. Hence you can choose to view the results as a glass half-empty or half-full. Some may say there's no good reason that a child who initially performed at the head of the class shouldn't continue doing so. And they'd be right. There are real consequences for graduates who descend from the 90th to 70th percentile in terms of merit-based aid and choice of college. It is up to the parents, schools, teachers, and so on, they'd say, to ensure that a child with that much demonstrated potential maintains buoyancy.

On the other hand, we ended up with more high achievers overall than we started with. The proportion of "Late Bloomers" surpassed the "Descenders" at both grade-level bands and in both reading and math. Surely that's good news.

Second, and more distressing, the progress of the high achievers didn't keep up with that of their lower-achieving peers, at least in reading. (In math, the performance gaps between high, middle, and low achievers changed very little with time.) In fact, high achievers grew about half as fast from third grade to eighth grade as low-achieving elementary/middle school students, reducing the gap between the two groups by over a third. One could celebrate such gap-closing, but one could also be dismayed by the "leveling" at work. We can hypothesize that many factors contributed to these results—perhaps NCLB's focus on low-performing schools or Reading First's focus on struggling readers. We simply don't know. But we are concerned nonetheless.

Third, poverty amongst one's schoolmates may not be the thief of high performance that we once thought. Exploratory findings here cast doubt on the notion that wealthy suburban schools produce greater academic gains for students than their poorer counterparts. These findings carry echoes of the original 1966 Coleman report. Perhaps growth over time for the highest-achieving students has little to do with the schools they attend and much to do with what's happening for them personally and at home. Perhaps.

Finally, while the progress that many students made in this study is praiseworthy, it's not staggering. Yes, let's laud our "Late Bloomers." But keep in mind that most of these kids were already above average at our starting point, with nearly all performing between the 50th and 89th percentiles in third grade. (By definition, they could not perform in the 90th percentile or above). Even our "Descenders" aren't altogether pitiable; while they no longer performed at or above the 90th percentile as they did in third grade, the vast majority remained above the 70th percentile in eighth grade.

What we're not seeing is students scrape and claw their way into the high-achieving ranks from the 20th, 30th, even 40th or 50th percentiles. Instead, students come in and out of the top decile but basically stay within the top third of students. These are our high-achieving "bubble kids," standing between academic fit and stretch; between sufficient and life-changing opportunities; between adequate and stellar futures. No, these aren't the kids that education-reform outfits fuss about. They aren't catalysts for campaigns to expand school choice, or initiate weighted student funding, or end last-in-first-out policies. They don't tug at the heartstrings like the needy children in our most wretched school systems. (Some of them reside there, but most don't.) But they deserve attention, too: Eight, ten, twelve, seventeen years old, with little more than a coin toss determining whether they wind up their school careers simply "above average" or among the country's top achievers and brightest hopes for the future.

What will we do to bolster their odds?

Acknowledgments

Generous support for this project was provided by the Kern Family Foundation, as well as by our sister organization, the Thomas B. Fordham Foundation. Many thanks to the Northwest Evaluation Association—specifically Yun Xiang, Michael Dahlin, John Cronin, Robert Theaker, and Sarah Durant—for the countless hours spent on both this "short" report and the more comprehensive report we'll release later this year. We appreciate their willingness to accommodate our suggestions for study design and presentation, as well as their patience with nonstop edits. Thanks also to NWEA's Rebecca Moore for her help with editing and proofreading and to Clay Johnson for his research assistance.

We're also grateful to John Papay, assistant professor of Education at Brown University, who served as external reviewer for the project. Dr. Papay provided immeasurably useful input on the study design and on consecutive drafts of the report. A big thank you also goes to the whole Fordham team for assistance on this project, especially Michael Petrilli and Chester E. Finn, Jr., for their project guidance and feedback; to Janie Scull for skillful editing and production management; and to Candice Santomauro, Joe Portnoy, and Daniela Fairchild for dissemination. Shannon Last served as copy editor, Alton Creative as layout designer, and David Flanagan as cover illustrator.

INTRODUCTION

If America is to remain internationally competitive with other advanced nations, we must maximize the academic potential of our top students. Over the last decade, however, federal and state education accountability systems—particularly in the wake of the No Child Left Behind Act (NCLB) of 2001—have placed primary emphasis on moving low-performing students toward proficiency. The sanctions stemming from these systems have cast greater attention on schools that fail to attain proficiency for most students—a necessary and noble endeavor. But they have also fueled concerns that the academic needs of high-performing learners, who in many states are largely unaffected by accountability systems, have been neglected.

To date, few research studies have examined the progress of individual high achievers over time in relation to other students. (See *Rundown on the Research on High Flyers* on page 6.) A few studies—including Duffett, Farkas, and Loveless (2008), commissioned by the Fordham Institute—have followed the top tenth of students in an effort to determine how these students rise—or fall—on absolute scales of academic performance. But we know of no research to date that has followed the progress of *individual* high achievers over time. This analysis helps to fill that gap. Using data from an extensive student-level database maintained by the Northwest Evaluation Association™ (NWEA) and its Measures of Academic Progress™ (MAP) assessments, we compared the performance and growth of high achievers to that of their peers over multiple years, examining two groups of students: an elementary/middle school cohort, followed from third through eighth grades; and a middle/high school cohort, followed from sixth through tenth grades. We sought answers to three key questions:

- **Do high achievers maintain their altitude?** In other words, are the nation's star third graders the same students that graduate eighth grade at the top of the pack? Or do up-and-coming peers surpass them? To find out, we compared student achievement at the initial and final years of the analysis—third and eighth grades for elementary/middle school students, and sixth and tenth grades for middle/high school students.
- **For those students who "lose altitude" over time, how far do they fall? And for those who climb into the top tier, how did they perform academically in earlier grades?** We tracked the achievement of these volatile high flyers to determine whether they experienced large swings in performance or remained relatively solid students throughout their school careers.
- **How much do high achievers grow academically over time?** While high achievers, by definition, perform better than 90 percent of their peers, do they get further ahead each year? Or do low- and middle-achieving students gain ground relative to them? We examined the performance gaps between these three groups of students and whether those gaps grew or narrowed over time.

The study also briefly investigated *which* students—by race, gender, and school environment—remained high achievers throughout their careers, and whether certain types of high achievers (or high achievers in certain types of schools) displayed different rates of academic growth over time.[2]

Methods in Brief

Data were drawn from NWEA's Growth Research Database, a longitudinal repository containing MAP assessment results. MAP tests are a series of computer-based adaptive assessments offered in mathematics, reading, language usage, and science that are typically administered to students in grades two through ten. The full repository includes data from 4,800 school systems and approximately five million students.

2 This report is the short version of a more comprehensive report that will be released later this year. Expect then to hear more about how school-level factors impacted high-achieving students, among other lines of inquiry.

RUNDOWN ON THE RESEARCH ON HIGH FLYERS

The body of research regarding the academic performance of high achievers is relatively limited, although the Thomas B. Fordham Institute has published several reports in this area. The most relevant of these is *High-Achieving Students in the Era of NCLB* (Duffett, Farkas, and Loveless, 2008). Using data from the National Assessment of Educational Progress (NAEP), Loveless concluded that low achievers made big strides in performance from 2000 to 2007, but that the progress of high achievers remained consistently meager over time. Unfortunately, NAEP data cannot be used to trace the performance of individual students over time, so the Loveless analysis relied on cross-sectional comparisons.

Several studies on this general topic have focused on the effect that proficiency-centered accountability systems (NCLB in particular) may have on the growth of high-performing learners. Neal and Schanzenback (2010) found that Chicago's shift to a high-stakes test led to achievement gains among students at the threshold of proficiency. In a study of Texas data, Reback (2008) found that low-achieving students performed better than expected when their scores were important to a school's rating, while the performance of high-achieving students did not change. These findings have been reinforced by some other research (Dee & Jacob, 2011; Duffett, Farkas, and Loveless, 2008), but not all findings in this area have been consistent. A rigorous study by Ballou and Springer (2008) examined test scores over three years in one western state and found gains across the achievement spectrum. A 2011 update of this study (Kober, McMurrer, and Silva, 2011) again reported no gains posted by low performers at the expense of high performers, but did find that the former showed *larger* gains than the latter. Other evidence suggests that the achievement of high-performing students has not suffered under NCLB (Cronin, Kingsbury, McCall, & Bowe, 2005; Chudowsky, Chudowsky, & Kober, 2009).

The existing body of research has several limitations. First, prior studies are generally limited to short time frames or a few grade levels. Second, school poverty and other factors related to school context are not frequently considered. Third, most prior research fails to acknowledge that the distribution of high-achieving students is uneven: If one defines the threshold of high achievement as students performing at or above the 90th percentile, middle- and high-income students are certain to be overrepresented relative to low-income students, and a low-income student at the top of his class—but at the 85th percentile overall—would be overlooked. This study was designed to address all of these gaps. We examine multiple grade levels over several years, consider the impact of school-based factors, and adopt a school-based definition of "high achiever" in a separate, preliminary analysis (see *A Closer Look at High Flyers in High-Poverty Schools* on page 15).

In this study, high achievers—dubbed "high flyers" in these pages—were defined as those students who scored at or above the 90th normed percentile on their MAP math and reading assessments, according to the NWEA 2008 RIT Scale Norms (NWEA, 2008).[3] We tracked two groups of high achievers over time: an elementary/middle school cohort and a middle/high school cohort. The elementary/middle school cohort comprised 81,767 students in math and 93,182 students in reading, from more than 1,500 schools in thirty states. Of those students tracked in math, 10,116 (12.4 percent) qualified as high flyers in third grade, while 10,925 third-grade students (11.7 percent) were high flyers in reading.[4] We followed this cohort from 2004-05 through 2009-10, as those students progressed from third grade to eighth grade.

3 The RIT scale (Rasch unit) is an IRT-based equal-interval scale used to measure student achievement, somewhat akin to using feet and inches on a yardstick to measure height. The scale can be used to chart a student's academic growth from year to year. We recognize that some problems may be introduced when identifying high performers using a cut score defined from a norm. For example, a norm does not necessarily provide a fixed standard; that is, the 90th percentile in third-grade math within a test's 2008 norms may not be the same as the 90th percentile within the same test's 2011 norms. Norming groups may improve or slip in their performance over time.

4 The MAP scores of students in our sample were evaluated against NWEA's 2008 norm population. The norm and study populations are therefore distinct; thus, the percentage of students performing at or above the 90th normed percentile could in practice be more or less than 10 percent of the study population. The data show that this was indeed the case here, as the proportion of high flyers in the initial sample (elementary/middle school) was greater than 10 percent (Table 2, page 9).

The middle/high school cohort comprised 43,423 students in math and 48,220 students in reading, from more than 800 schools in twenty-eight states. Among the math students, 2,912 (6.7 percent) were high flyers in sixth grade, while 4,394 (9.1 percent) of sixth-grade reading students were high flyers. We followed these students from 2005-06 through 2009-10, as they progressed from sixth grade to tenth grade.[5] Table 1 shows the two cohorts of students followed over time.

TABLE 1

Cohorts 1 and 2 by Grade and Year

Cohort	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Cohort 1 (Elementary/Middle School)	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cohort 2 (Middle/High School)	N/A	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10

The two cohorts consisted of students who had MAP scores in both the initial and final years of the study. We did not require that students have MAP scores for each intervening year, since very few students met that criterion. While the choice of third and sixth grades is in one sense arbitrary, both were selected because they represent a form of entry point: third grade because it is the first NCLB-tested grade and is the beginning of the intermediate grades in many schools, and sixth grade because it is the typical entry point for middle school.[6]

5 We did not track students through higher grades due to the fact that smaller numbers of students participate in testing at those grades.

6 See full methodology and limitations in Appendix I. Additional data tables are available online at the Thomas B. Fordham Institute's website at http://www.edexcellence.net/publications-issues/publications/high-flyers.html and at the Kingsbury Center Data Gallery at http://kingsburycenter.org/gallery/high-achievers.

FINDINGS: DO HIGH FLYERS MAINTAIN THEIR ALTITUDE?

What are the odds that a star third grader will still rank at the top of the pack by eighth grade? Or that a bright, budding sixth grader will remain a model student in high school? To find out, this analysis traced high-achieving students across multiple years to determine how many of them remained high-achieving over time; how many lost their high-achieving status; and how many gained that distinction. Students were sorted into the following categories:

- **Steady High Flyers:** Students who were high-achieving in both the initial and final years of the study (i.e., third and eighth grades for elementary/middle school students, sixth and tenth grades for middle/high school students)
- **Descenders:** Students who were high-achieving in the initial, but not the final, year
- **Late Bloomers:** Students who were high-achieving in the final, but not the initial, year
- **Never High Flyers:** Students who were not high achievers in the initial or the final year

FINDING #1 A majority of high flyers maintained their status over time, but substantial numbers "lost altitude."

As shown in Figure 1, a majority of high achievers remained that way over time, earning them the designation "Steady High Flyers." **Nearly three in five students identified as high-achieving in the initial year of the study remained high-achieving in the final year**. That is, 57.3 percent of high-achieving third-grade math students remained that way by eighth grade, while 55.9 percent did so in reading. A full 69.9 percent of high-achieving sixth-grade math students remained high-achieving by tenth grade; 52.4 percent did so in reading. The converse of these students, of course, are the 30 to 50 percent of initially high-achieving students that proved unstable and lost that status over time—earning them the designation of "Descenders."[7]

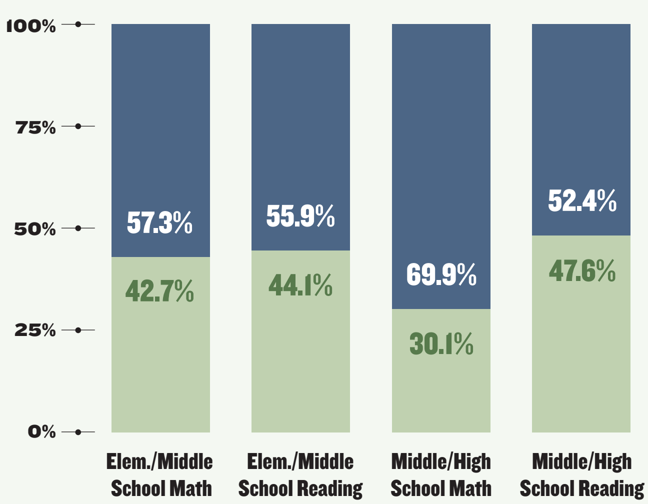


FIGURE 1

Outcomes of Initial High Flyers

Steady High Flyers | Descenders

Note: For each cohort and subject, the figure shows what percentage of the initial high achievers was still high-achieving in the final year of the study (Steady High Flyers), and what percentage was not (Descenders). Students were tracked from grades three to eight in elementary/middle school and grades six to ten in middle/high school. For example, 57.3 percent of the initial high math achievers in the elementary/middle school cohort remained high-achieving from third to eighth grade (i.e., they were Steady High Flyers), while 42.7 percent were no longer high-achieving by eighth grade (i.e., they were Descenders).

7 The rate of attrition is somewhat related to measurement error; for more information, see Appendix I.

Though substantial proportions of the high achievers lost that status over time, that isn't to say that the pool of high-achieving students shrank; on the contrary, it grew (Table 2), thanks to students ascending into the high-achieving ranks.[8] The percentage of high flyers in math at the elementary/middle level, for instance, grew from 12.4 percent of all students in third grade to 14.1 percent in eighth grade.

TABLE 2

High Achievers in Initial and Final Years

	Total Number of Students in Cohort	Number of High Flyers in Initial Year	Percentage of High Flyers in Initial Year	Number of High Flyers in Final Year	Percentage of High Flyers in Final Year	Change in High-Flyer Percentage
ELEMENTARY/MIDDLE SCHOOL COHORT						
Math	81,767	10,116	12.4%	11,544	14.1%	+1.7%
Reading	93,182	10,925	11.7%	12,429	13.3%	+1.6%
MIDDLE/HIGH SCHOOL COHORT						
Math	43,423	2,912	6.7%	4,779	11.0%	+4.3%
Reading	48,220	4,394	9.1%	4,677	9.7%	+0.6%

These increases were fueled by greater numbers of Late Bloomers entering the high-achieving ranks (Table 3). Within the full elementary/middle school cohort, 5.3 percent of students in math were Descenders, while 7.0 percent proved to be Late Bloomers. In reading, 5.2 percent of those students were Descenders, while 6.8 percent proved to be Late Bloomers. In the full middle/high school cohort, 2.0 percent of students in math were Descenders, compared with 6.3 percent who were Late Bloomers. In reading, 4.3 percent of students were Descenders, while 4.9 percent of students were Late Bloomers.

TABLE 3

Migration of High Achievers

	Total Number of Students in Cohort	Number of Descenders	Percentage of Descenders	Number of Late Bloomers	Percentage of Late Bloomers
ELEMENTARY/MIDDLE SCHOOL COHORT					
Math	81,767	4,317	5.3%	5,745	7.0%
Reading	93,182	4,817	5.2%	6,321	6.8%
MIDDLE/HIGH SCHOOL COHORT					
Math	43,423	878	2.0%	2,745	6.3%
Reading	48,220	2,090	4.3%	2,373	4.9%

8 Given that "high-achieving" status is defined as those students performing at or above the 90th normed percentile, one might assume that the Descenders' loss is the Late Bloomers' gain; that is, that the Late Bloomers simply assume the other group's place in the academic pecking order. Yet, there is no such thing as a "zero-sum game" here since the norm population is independent of the study population. See footnote 4 and/or Appendix I for additional discussion.

FINDING #2 The majority of students who attained high-flyer status at one point in time did not stray far from it.

While the Descenders fell below the 90th percentile by eighth or tenth grades, most did not fall far below. Take, for instance, those students who were high-performing in third-grade math but not in eighth-grade math. On average, those students still performed at the 77th percentile by eighth grade (Figure 2). Put another way, those students dropped from the top 10 percent of their grade to the top 30 percent. Late Bloomers also did not typically have far to climb to become high math achievers by eighth grade—on average, those students performed at the 74th percentile in third-grade math. (Results were similar for elementary/middle school reading and middle/high school math and reading.[9]) So while the

WHICH STUDENTS WERE MOST LIKELY TO REMAIN HIGH FLYERS?

Nearly half of high flyers lost their altitude over time, and many students who were not originally high flyers eventually earned that designation. This volatility in the high-achieving group invites the question: Which students fell, and which students rose? Are they distinguishable by race, gender, or school-level poverty? Findings are summarized below. Data can be found in Tables A-4 and A-5 on pages 19-20.

Minority status: While minority students were underrepresented among high achievers at both the elementary/middle and middle/high school levels, the proportions of minority students within the high-achieving groups proved relatively stable and, in most cases, increased slightly over time.[1] Elementary/middle school math was the only subject in which minority representation *didn't* increase: Minorities represented 8.2 percent of high flyers in both third and eighth grades in that subject. In reading, however, minorities grew from 9.0 percent of third-grade high flyers to 9.4 percent in eighth grade. In middle/high school, minority students grew from 7.3 percent of high flyers in sixth-grade math to 7.8 percent in tenth grade, and from 6.7 percent in reading to 7.3 percent.

Gender: Girls were underrepresented among high achievers in math and were slightly overrepresented among high achievers in reading; still, their proportions in both subjects grew over time.[2] In elementary/middle school math, girls rose from 41.9 to 44.0 percent of all high flyers from third grade to eighth grade, and in reading from 51.7 to 53.0 percent of high flyers. In the middle/high school cohort, the proportion of female high flyers grew from 39.0 to 41.7 percent in math, and from 49.8 to 52.6 percent in reading. Though girls remained underrepresented in math, the increasing proportions of girls in both subjects rendered the relative decline of boys among the top-performing portion of American students increasingly apparent.

School poverty: Students in high-poverty schools were predictably underrepresented among high flyers, but unlike minority and female students, their proportions declined over time.[3] In third-grade math, 19.4 percent of high achievers attended high-poverty schools; that fell to 16.1 percent by eighth grade. In elementary/middle school reading, the proportion fell slightly from 13.5 to 13.4 percent. In the middle/high school cohort, students in high poverty schools accounted for 18.1 percent of high achievers in sixth-grade math; they totaled 15.3 percent by tenth grade. In reading, they declined from 16.6 to 14.7 percent from sixth grade to tenth grade.

1 Minority students were defined as children from traditionally disadvantaged ethnic groups and included African American, Hispanic, and Native American students. Non-minority students included Anglo and Asian students. Of the total study sample, approximately 23 percent of students were minority, while 77 percent were non-minority.

2 The total study sample consisted of relatively equal proportions of girls (49.6 percent) and boys (50.4 percent).

3 Low poverty was defined as schools in which less than 50 percent of students received free or reduced-price lunch, while high poverty refers to a school in which more than 50 percent did so. In the study sample, 31 percent of students attended high-poverty schools, and 69 percent attended low-poverty schools.

9 For figures depicting results of these additional cohorts and subjects, see the Thomas B. Fordham Institute website at http://www.edexcellence.net/publications-issues/publications/high-flyers.html and the Kingsbury Center Data Gallery at http://kingsburycenter.org/gallery/high-achievers.

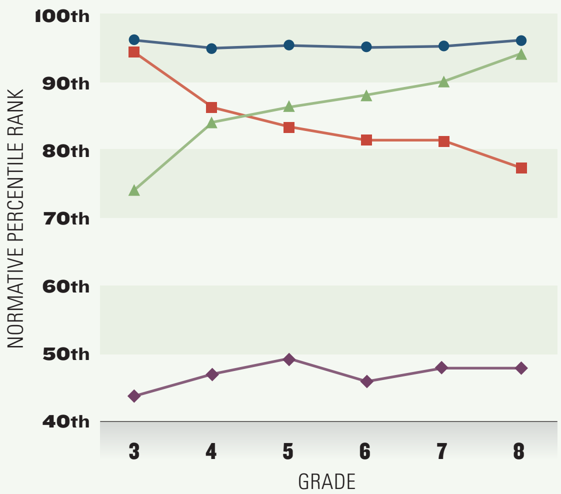


FIGURE 2

Migration Patterns by High-Flyer Status (Elementary/Middle School Math)

Steady High Flyers | Descenders | Late Bloomers | Never High Flyers

Note: The figure shows the mean achievement percentiles in math for four groups of students at each grade. For example, between third and eighth grades, Steady High Flyers consistently ranked near the 95th percentile in math; Descenders dropped from around the 94th to the 77th percentile; Late Bloomers rose from the 74th to the 94th percentile; and Never High Flyers wavered between the 43rd and the 49th percentile.

pool of high achievers did experience turnover, **migration in and out of high-achieving status was concentrated among students performing above the 70th percentile.**

Descenders showed gradual movement away from the 90th percentile over time while Late Bloomers showed similarly gradual progress toward this benchmark—unsurprising findings, considering how these groups were defined. As Figure 2 shows, the biggest movements occurred between third and fourth grades and between seventh and eighth grades. While explaining these developments is beyond the scope of this study, a portion of the large drop between third and fourth grades is likely attributable to some measurement error (see Appendix I for further discussion).

The achievement of Descenders and Late Bloomers is explored more thoroughly in Figures 3 and 4. Figure 3 illustrates the full range of achievement of the Descenders in eighth-grade math. While these students no longer performed at or above the 90th percentile, as they did in third grade, the vast majority still performed near it. Only a small percentage of these students performed below the 50th percentile—meaning that the vast majority of initial high achievers remained above average throughout their school years.

Similarly, Figure 4 illustrates the full range of achievement of Late Bloomers in third-grade math. How did these students, who were high-achieving by eighth grade, perform in their earlier years? The vast majority of them were above-

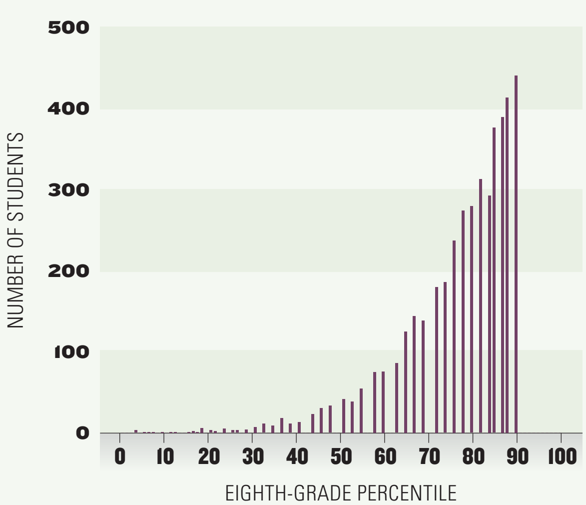


FIGURE 3

Eighth-Grade Performance of Descenders (Elementary/Middle Math)

Note: This figure illustrates the performance range of Descenders—those students who performed at or above the 90th percentile in third-grade math, but below the 90th percentile in eighth-grade math—in eighth grade by displaying the number of Descenders performing at each eighth-grade percentile. By eighth grade, few of those students performed below the 50th percentile, while many continued to perform near the 90th percentile.

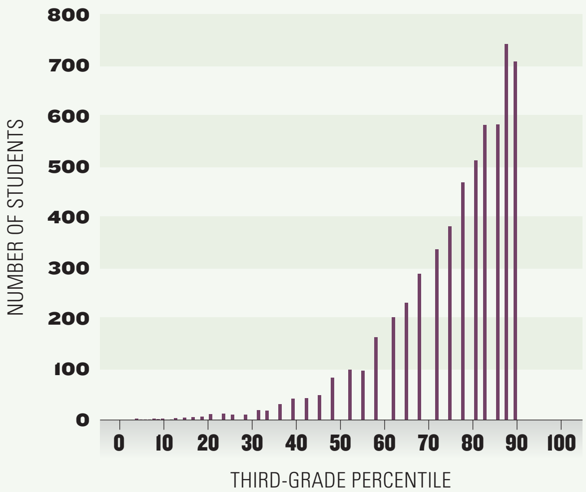


FIGURE 4

Third-Grade Performance of Late Bloomers (Elementary/Middle Math)

Note: This figure illustrates the performance range of Late Bloomers—those students who performed below the 90th percentile in third-grade math, but at or above the 90th percentile in eighth-grade math—in third grade by displaying the number of Late Bloomers performing at each third-grade percentile. In third grade, few of those students performed below the 50th percentile, while many performed near the 90th percentile.

average third graders, with overwhelming numbers performing between the 50th and 89th percentiles in third grade (by definition, they could not perform in the 90th percentile or above).

FINDING #3 High flyers grew academically at similar rates to low and middle achievers in math, but grew at slightly slower rates than low and middle achievers in reading.

As already noted, individual high flyers follow different trajectories throughout their academic careers: Some rise, some descend, and some maintain their altitude throughout their schooling. But every subject and grade has its high flyers; as a group, how much do they improve academically over time? Do they further outpace their low- and middle-achieving peers, or do those groups gain on the high achievers? To find out, we compared the academic growth rates of high-achieving students in reading and math in relation to middle achievers (those performing between the 45th and 54th percentiles, inclusive) and low achievers (those below the 10th percentile).[10]

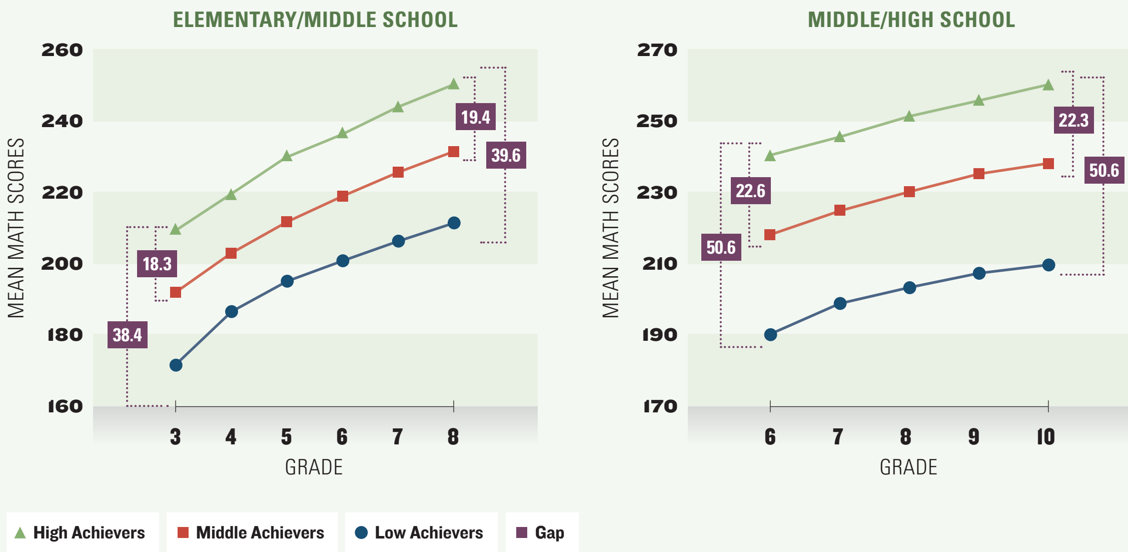
The performance gaps between high, middle, and low achievers were, as one would expect, quite large. In math, changes in those gaps over time were minimal. Elementary/middle school high achievers slightly increased their performance advantage over the other two groups between third grade and eighth grade, but those differences only amounted to an additional 25 percent of a year's growth for a typical high achiever (Figure 5). Even in eighth grade, the mean math scores of the low-performing group did not match the high achievers' third-grade marks, and middle-achieving eighth graders only ever matched the high achievers' fifth-grade marks. The pattern was similar for the middle/high school group: Gaps in mathematics performance between high, middle, and low achievers remained about the same over the four years (though the gaps between high and low performers were larger in magnitude at the middle/high school level than at the elementary/middle school level).

In reading, however, low- and middle-achieving students demonstrated faster rates of improvement than high achievers (Figure 6). The resulting narrowing of these performance gaps can be attributed to sluggish growth of those students at the

10 Growth here refers to the rate at which students increased their mean scores. To be included in the study, a student must have had test results for both the initial and final grades of the cohort. Thus the difference in average scores at these two points represents the actual growth of the group between these grades. Because members of the cohorts were not required to have a test result in each grade, the averages at the other grades do not necessarily reflect the actual mean growth of the group.

FIGURE 5

Academic Growth of High, Middle, and Low Achievers (Math)



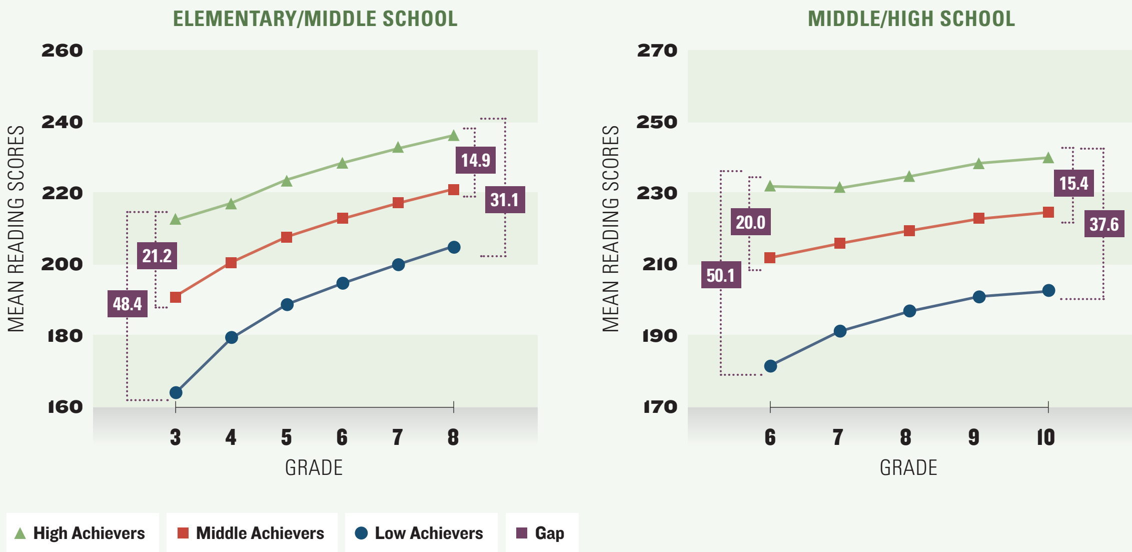
Note: These figures illustrate the growth in math achievement made by high, middle, and low achievers by plotting each group's mean scores between third grade and eighth grade (for elementary/middle school students) and between sixth grade and tenth grade (for middle/high school students). Performance is measured by NWEA's MAP assessments; scores can range from about one hundred to about 350. For example, low achievers in elementary/middle school improved their mean score from 172.0 to 211.2 in that time, while high achievers in elementary/middle school improved their mean score from 210.4 to 250.8, slightly increasing the performance gap between the two groups by 1.2 points.

top. From third grade to eighth grade, low-achieving elementary/middle school students grew nearly twice as fast on the assessment as high achievers, reducing the performance gap between the two groups by over a third. Middle achievers reduced their performance gap with high achievers by approximately 30 percent. High achievers still outperformed middle and low achievers by large gaps—once again low achievers never surpassed the third-grade mean of high achievers, and middle achievers never surpassed the high achievers' fifth-grade mean—but high achievers did not soar quite as high above their peers in eighth grade as they did in third grade.[11] Patterns were again similar in the middle/high school group, though the reduction in gaps was not as dramatic. Both low and middle achievers reduced their performance gaps with high achievers by about 25 percent.

11 The sluggish growth in reading as students advance in grade sometimes raises questions about possible ceiling effects on the test. This is commonly characterized as a lack of "room to grow." The assumption is that students are testing at or near the highest possible score on the test. The MAP test is adaptive, however, meaning high- and low-performing students receive more items targeted to their current achievement levels than they would receive on fixed-form assessments. Thus, there is less likelihood of ceiling effects. As evidence of this, standard errors on the reading test at the eighth-grade 90th percentile are not significantly different from those found in the middle of the distribution (NWEA, 2008), which typically means that students performing at the cut point are not challenging the ceiling of the test. Ironically, what appears to be sluggish reading growth may actually be tied to how reading development manifests itself among high achievers. At some point, reading development becomes subject-dependent, and tests of general reading may not adequately measure it. For example, a general test of reading ability typically will not include highly specialized science reading passages (e.g., an excerpt from a scholarly paper on genetic engineering), because students would require prior knowledge to understand such text. But it is precisely this type of specialized reading that many high achievers confront in high school.

FIGURE 6

Academic Growth of High, Middle, and Low Achievers (Reading)



Note: These figures illustrate the growth in reading achievement made by high, middle, and low achievers by plotting each group's mean scores between third grade and eighth grade (for elementary/middle school students) and between sixth grade and tenth grade (for middle/high school students). Performance is measured by NWEA's MAP assessments; scores can range from about one hundred to about 350. For example, low achievers in elementary/middle school improved their mean score from 164.3 to 205.1 in that time, while high achievers in elementary/middle school improved their mean score from 212.7 to 236.2, reducing the performance gap between the two groups by 17.3 points.

A CLOSER LOOK AT HIGH FLYERS IN HIGH-POVERTY SCHOOLS

In the current report, we defined high-achieving math and reading students as those with scores at or above the 90th percentile on NWEA's MAP assessments. This definition, however, excluded many students attending high-poverty schools; even when those students were high-performing relative to their peers, many did not perform at or near the externally normed 90th percentile. In an additional line of inquiry (to be described and discussed more fully in a forthcoming report), we examined a different group of students using a new definition of what it meant to be a high achiever. For those analyses, we defined high-achieving students as those whose math or reading scores placed them within the top 10 percent *of their individual grades and schools*. Using that school-based definition, we examined the relationship between school poverty and high achievers' academic performance and growth. We tracked an elementary school cohort from third grade to fifth grade, and a middle school cohort from sixth grade to eighth grade.

From the start, it was clear that this school-based definition of "high achiever" captured a different group of students: Many students in high-poverty schools who ranked at the top of their own classes did not rank at or above the larger 90th percentile based on overall NWEA norms. In other words, higher poverty rates generally predicted lower overall academic performance. In math, only 76.1 percent of third graders who were high-achieving within their schools achieved at or above the external 90th percentile—and this dropped to just 69.3 percent by fifth grade. In reading, 80.7 percent of high-achieving third graders performed at or above the external 90th percentile, and this declined to just 63.8 percent by fifth grade. Middle school students fared similarly, with 87.2 percent of high-achieving sixth graders surpassing the 90th percentile in math and only 69.3 percent doing so in eighth grade; in reading, the proportion fell from 83.9 to 61.4 percent between sixth and eighth grades. (Data not shown in tables.)

In terms of growth, however, we did uncover a surprising and encouraging trend: School poverty was not a strong predictor of student progress. High flyers at low-poverty schools performed on average at the 97th percentile in third grade math, while high flyers at high-poverty schools scored at the 83rd percentile—a difference representing over a year's worth of growth. By fifth grade, however, they scored at the 97th and 82nd percentiles, respectively. While high achievers in high-poverty schools grew slightly less than those in low-poverty schools, the difference was marginal. The same pattern held for middle school math. For both elementary and middle school reading, the gaps between high-achieving students in high- and low-poverty schools slightly diminished over time, but again, only marginally.

These findings suggest that the relationship between a school's poverty rate and extent of growth among its high-achieving students is very weak. In fact, both high- and low-poverty schools varied dramatically in the growth of their high achievers; in other words, high- and low-growth schools could be found among high- and low-poverty schools alike. Attending a low-poverty school improves the average high achiever's prospects for growth by very little; it appears that factors other than poverty control the growth of high achievers within a given school.[1]

1 Due to the limited number of schools available for the school factor analyses, we did not have a representative sample of all American schoolchildren in these grades. Our sample contained proportionally fewer high-poverty schools and urban schools. Note that our key finding—that a school's poverty rate is not a strong predictor of success for high achievers—might be less robust given a more balanced sample. Further, because student mobility within schools is likely to affect the average growth rates observed by those schools, a longitudinal design such as ours essentially disregards the potential impact of mobility on student growth. Thus, our findings must be considered preliminary and not conclusive. For more information on this line of analysis, see Appendix II.

CONCLUSION

The future prosperity of our nation rests not only on our ability to improve the performance of our lowest-achieving students, but also on our ability to support and advance the performance of our highest-achieving students. As this study shows, many students—about two in five—who were high-achieving in early grades had lost that status just four years later. While the vast majority of these Descenders were never *low* performers—on average, they declined from above the 90th percentile to just below the 80th percentile—their decline is likely to have a substantive impact on their long-term education outcomes. Students performing at or above the 90th percentile are more likely to have access to gifted programs in elementary school and honors or advanced placement courses in high school. If they maintain their achievement and grades, their performance is likely to qualify them for selective colleges and universities and for higher tiers of merit aid than other students. These Descenders should be of great concern for educators. Once a student's capacity for high achievement is established, the school's objective should be to ensure that that student maintains an upward trajectory. After all, these are bright, highly capable individuals. Every casualty among this group is a loss in potential human capital, and schools need to find and implement strategies that effectively stem performance losses among students who show promise in the early grades.

It is important to note, however, that this out-migration from the top decile is surpassed by in-migration; in other words, a larger number of students achieve high-flyer status over time than lose it. Like those students who lost their high-achieving status, these Late Bloomers never performed far below the 90th percentile—the majority consistently ranked in the top 30 percent of students or above. They demonstrate that growth into high-achiever status is not just possible, but common for above-average students. The progress of these students begs the question: How can we improve the achievement of *other* students performing in the 70th and 80th percentiles?

Perhaps the least expected finding of this study is that, while one finds fewer high achievers in high-poverty schools, school poverty rates have little relation to the academic growth of high-achieving students. This challenges the widespread belief that schools in wealthy suburbs will produce the greatest gains in student achievement, at least among the highest achievers. Our study shows instead that while wealthy, suburban schools do tend to house more high-performing students, their students don't show substantively greater gains than high achievers in high-poverty schools. These findings, albeit preliminary, suggest that placing a high-performing student in a high-growth school is largely a lottery. The schools within the sample varied greatly in the growth they produced for high-performing students, rendering the odds that a low-poverty school would produce high growth at only slightly over 50 percent.

If we are truly serious about providing excellence in education for all students, then we should consider changing accountability systems to place emphasis on the growth of low-, middle-, and high-achieving students alike. Our results suggest that this type of accountability would subject some wealthy, underperforming suburban schools to fair and welcome scrutiny.

Many of America's future leaders in business, in politics, and in service to humanity will almost certainly be the high achievers in our schools today. While these children are not in short supply, this study suggests that we are not doing everything we could to nurture and sustain their promise, to increase their numbers, and to assure that high-achieving minority students and students in high-poverty schools have every opportunity to reach their goals. The primary reasons these promises remain unfulfilled are unrelated to the resources available to schools. Were that the case, low-poverty schools would produce substantively better gains than others. Instead, the problem seems to be one of consistency in both policy and practice. Educational policy in recent years has focused more on low-performing students than high achievers, and the curricula and instructional practices adapted toward these students produce inconsistent and idiosyncratic results. We can and must do better if we want to secure a future for our children that reflects the opportunities that past generations have enjoyed.

REFERENCES

Ballou, Dale and Matthew Springer. 2008. *Achievement Trade-Offs and No Child Left Behind*. Nashville, TN: Peabody College of Vanderbilt University. http://www.vanderbilt.edu/schoolchoice/documents/achievement_tradeoffs.pdf.

Chudowsky, Naomi, Victor Chudowsky, and Nancy Kober. 2009. *State Test Score Trends Through 2007-08: Are Achievement Gaps Closing and Is Achievement Rising for All?* Washington, D.C.: Center on Education Policy. http://www.cep-dc.org/publications/index.cfm?selectedYear=2009.

Cronin, John, Gage G. Kingsbury, Martha S. McCall, and Branin Bowe. 2005.*The Impact of the No Child Left Behind Act on Student Achievement and Growth: 2005 Edition*. Portland, OR: Northwest Evaluation Association. http://greatlakescenter.org/docs/early_research/pdf/NCLBImpact_2005_Brief.pdf.

Dee, Thomas S. and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30: 418–446.

Duffett, Ann, Steve Farkas, and Tom Loveless. 2008. *High-Achieving Students in the Era of NCLB*. Washington, D.C.: Thomas B. Fordham Institute. http://www.edexcellence.net/publications-issues/publications/high-achieving-students-in.html.

Kober, Nancy, Jennifer McMurrer, and Malini Silva. 2011. *Is Achievement Improving and Are Gaps Narrowing for Title I Students?* Washington, D.C.: Center on Education Policy. http://www.cep-dc.org/publications/index.cfm?selectedYear=2011.

Neal, Derek and Diane Whitmore Schanzenback. 2010. "Left Behind by Design: Proficiency Counts and Test-Based Accountability." *Review of Economics and Statistics* 92: 263-283.

Northwest Evaluation Association. 2008. *RIT Scale Norms*. Portland, OR: Northwest Evaluation Association. http://www.nwea.org/sites/www.nwea.org/files/resources/2008%20RIT%20Scale%20Norms%20Study.pdf.

Reback, Randall. 2008. "Teaching to the Rating: School Accountability and the Distribution of Student Achievement." *Journal of Public Economics* 92: 1394-1415.

APPENDIX I METHODOLOGY

Two cohorts were created for the main study. (The sample for the secondary analyses, highlighted in the sidebar titled *A Closer Look at High Flyers in High-Poverty Schools,* is explained in Appendix II.) They consisted of public-school students in grades three through eight (Cohort 1) and in grades six through ten (Cohort 2). Among these cohorts, high achievers were those students who performed at or above the 90th percentile, based on NWEA's 2008 norms, on their third-grade (Cohort 1) or sixth-grade (Cohort 2) Measures of Academic Progress (MAP) tests.

Students in **Cohort 1** were followed from third grade in the 2004-05 school year until the end of eighth grade in the 2009-10 school year (Table A-1). The elementary/middle school cohort consisted of 81,767 students in math and 93,182 students in reading, and was drawn from more than 1,500 schools in thirty states.

Students in **Cohort 2** were followed from sixth grade in the 2005-06 school year through tenth grade in the 2009-10 school year.[12] This cohort consisted of 43,423 students in math and 48,220 students in reading, and was drawn from more than 800 schools in twenty-eight states.

Each cohort included only students who had data in both the initial and final grades studied. Students were not required to test in each intervening grade as a condition for inclusion in the cohort, because this requirement would have decreased the size of the sample dramatically.

TABLE A-1

Cohorts 1 and 2 by Grade and Year

Cohort	2004-05	2005-06	2006-07	2007-08	2008-09	2009-10
Cohort 1 (Elementary/Middle School)	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cohort 2 (Middle/High School)	N/A	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10

The proportions of high achievers in the study varied by cohort and by subject, though every group saw its proportion of high achievers grow over the course of the study (Table A-2). The percentage of high flyers in math at the elementary/middle level, for instance, grew from 12.4 percent of all students in third grade to 14.1 percent in eighth grade.

Demographic characteristics of the full sample, by cohort and subject, are shown in Table A-3.[13] Parallel demographic breakdowns of the high-achieving groups within each cohort and subject are shown in Table A-4 (initial year of study) and Table A-5 (final year of study). A quick comparison of the two tables reveals that females were underrepresented among high achievers in math, and both minority students and students in high-poverty schools were underrepresented in all categories of high achievers.

12 We did not track students through higher grades due to the fact that smaller numbers of students participate in testing at those grades.

13 Minority students were defined as children from traditionally disadvantaged ethnic groups and included African American, Hispanic, and Native American students. Non-minority students included Anglo and Asian students. Low poverty was defined as schools in which less than 50 percent of students received free or reduced-price lunch, while high poverty refers to a school in which more than 50 percent did so.

TABLE A-2

High Achievers in Initial and Final Years

	Total Number of Students in Cohort	Number of High Flyers in Initial Year	Percentage of High Flyers in Initial Year	Number of High Flyers in Final Year	Percentage of High Flyers in Final Year	Change in High-Flyer Percentage
COHORT 1 (ELEMENTARY/MIDDLE SCHOOL)						
Math	81,767	10,116	12.4%	11,544	14.1%	+1.7%
Reading	93,182	10,925	11.7%	12,429	13.3%	+1.6%
COHORT 2 (MIDDLE/HIGH SCHOOL)						
Math	43,423	2,912	6.7%	4,779	11.0%	+4.3%
Reading	48,220	4,394	9.1%	4,677	9.7%	+0.6%

TABLE A-3

Demographic Characteristics of Total Sample of Students

	Gender		Ethnicity		School Poverty Status	
	Female	Male	Minority	Non-minority	High Poverty	Low Poverty
Cohort 1 Math	49.7%	50.3%	23.3%	76.7%	30.6%	69.4%
Cohort 1 Reading	49.3%	50.7%	24.4%	75.6%	29.4%	70.6%
Cohort 2 Math	49.2%	50.8%	25.1%	74.9%	31.0%	69.0%
Cohort 2 Reading	48.9%	51.1%	23.5%	76.5%	30.1%	69.9%

TABLE A-4

Demographic Characteristics of High Achievers (Initial Year of Study)

	Gender		Ethnicity		School Poverty Status	
	Female	Male	Minority	Non-minority	High Poverty	Low Poverty
Cohort 1 Math	41.9%	58.1%	8.2%	91.8%	19.4%	80.6%
Cohort 1 Reading	51.7%	48.3%	9.0%	91.0%	13.5%	86.5%
Cohort 2 Math	39.0%	61.0%	7.3%	92.7%	18.1%	81.9%
Cohort 2 Reading	49.8%	50.2%	6.7%	93.3%	16.6%	83.4%

TABLE A-5

Demographic Characteristics of High Achievers (Final Year of Study)

	Gender		Ethnicity		School Poverty Status	
	Female	Male	Minority	Non-minority	High Poverty	Low Poverty
Cohort 1 Math	44.0%	56.0%	8.2%	91.8%	16.1%	83.9%
Cohort 1 Reading	53.0%	47.0%	9.4%	90.6%	13.4%	86.6%
Cohort 2 Math	41.7%	58.3%	7.8%	92.2%	15.3%	84.7%
Cohort 2 Reading	52.6%	47.4%	7.3%	92.7%	14.7%	85.3%

Beyond the composition of the two cohorts, there are five key technical issues and/or limitations that require clarification:

1. **Possible misclassification attributable to error on the assessment.** The issue of measurement error is relevant to this study because it is possible that students whose measured MAP scores placed them just under or over the 90th percentile threshold for high achievement might "truly" belong on the other side. In other words, it is possible that the true scores of some high achievers may fall below the 90th percentile and the true scores of some non-high achievers may fall above the cut line. (For individual students, the misclassification probability is 50 percent for students whose scores are exactly at the threshold, but that probability decreases quickly as students' measured scores rise higher above or fall further below the high-achievement cut off).

   To account for measurement error, tests often provide a standard error of measurement (SEM) along with an observed score. Standard errors indicate how much potential measurement error may exist. For example, if a student's measured score was 110 and the standard error of measurement was two, then that student's "true" score most likely falls within the range of 108 to 112, and almost certainly within the range of 106 to 114. With computerized adaptive tests such as MAP, standard errors of measurement are roughly constant for all observable MAP scores; typical SEMs on the test are about three points.

   Because this study examines the attrition rate among students who perform at or above the 90th percentile on NWEA's norms, it is important to address questions regarding the likely proportion of students scoring at that level who may have been misidentified due to SEM. Hence, we calculated the "maintenance rate" that would be expected for the high-achieving group if measurement error were taken into account. The calculation depends on the score distribution among the sample of high achievers (how close initial high achievers scored to the cut off) and the measurement error associated with each score. Based on this information, we estimated the expected maintenance rates were 83.4 percent for math and 83.3 percent for reading for the elementary/middle school cohort. Put another way, we would expect about 83 percent of the high-achieving math group to remain intact if the entire group of high achievers were immediately retested, and the attrition in the group would be explained by the measurement error associated with the test. The report indicates that the actual maintenance rate for this group was considerably lower; only 57 percent of the elementary/middle math cohort remained high-achieving. That means that the expected attrition, at about 17 percent, was much higher in reality, at 43 percent. The difference between the anticipated attrition and the observed attrition is perhaps the best representation of the actual attrition within the high-achieving group.[14] In other words, a total of 26 percent of students became Descenders for reasons other than measurement error.

2. **Risk of regression toward the mean.** Fixed-form tests have a relatively high risk of score regression toward the mean, partly due to ceiling effects and partly because they exact a relatively high penalty on inadvertent errors. Because fixed-form tests have to assess performance across the entire spectrum of achievement, they provide a limited number of items that can be targeted to any one group. Thus, the number of items used to measure the performance of high achievers tends to be small, generally only five to ten in a fifty item test. This contributes to a ceiling effect. In addition, because so few items discriminate among high achievers, a high achiever who inadvertently misses an item (forgets to "carry a one" on an addition problem, for example) often finds his score takes a large (and unrecoverable) penalty. Adaptive tests such as MAP, however, have lower risk of regression to the mean because they offer more appropriately targeted items to high-performing students and exact relatively small penalties for inadvertent errors. Further, in an adaptive test, the primary "penalty" for missing an item is that the student receives an easier item; in fact, students are expected to miss approximately 50 percent of the items on the test. Thus, they have many opportunities to "recover" from an inadvertently missed item.

14 While this addresses the likelihood of overclassification—i.e., classifying students as high-performing who are not—there are also students not included in the initial high-achieving cohort because they were underclassified due to measurement error. In a normal distribution, the number of underclassified students is likely to be slightly largely than the number overclassified because of their position in the distribution (there are fewer students just above the 90th percentile than there are just below the 90th percentile). Were measurement error to be completely resolved—an impossibility, unfortunately—the actual estimate of high achievers might be slightly higher than what we report.

3. **Use of a normative cut score to define eligibility for the sample.** This study used a normative standard, the 90th percentile on the 2008 NWEA norms, to define the cut point for high achievement. While the cut point was normative, the use of that standard did not limit the number of students in the study sample who could be identified as high-achieving, nor did it limit the number of students who could enter or leave high-achieving status. In other words, we did not design the study so that any student moving above the 90th percentile norm had to replace another student. In fact, the overall count of high-achieving students increased over time within the study group. Thus while a normative cut score was applied, its application did not create a zero-sum result for the study sample.

4. **Limiting the sample to students with scores at the beginning and ending grades.** We limited the sample to students who had test records at both the beginning and ending grades. We considered restricting the sample further to require students to have scores at every grade, but rejected that option because it would have excluded the vast majority of students from the sample. Limiting the sample in this way could impact measurement error as discussed above. However, it should be noted that in the secondary analyses (Appendix II) we used data points in the intervening years to measure growth in the hierarchical linear model (HLM).

5. **Measurement of student improvement.** To estimate improvement rates, we compared high achievers (students performing at or above the 90th percentile on NWEA norms) with low achievers (students performing below the 10th percentile) and middle achievers (students performing between the 45th and 54th percentiles, inclusive). We used all available data points to depict their trend lines, but the calculation was not based on individual growth trajectories. For the measure of growth (see *A Closer Look at High Flyers in High-Poverty Schools* on page 15), we estimated growth trajectories for the school-defined groups via HLM. For more information on this analysis, see Appendix II.

APPENDIX II
METHODOLOGY SPECIFIC TO *A CLOSER LOOK AT HIGH FLYERS IN HIGH-POVERTY SCHOOLS*

While the main findings of the report focused on the analysis of Cohorts 1 and 2, the sidebar titled *A Closer Look at High Flyers in High-Poverty Schools* introduced a separate line of inquiry. This analysis used a new definition of high flyers, and centered on two distinct cohorts (Cohorts 3 and 4). Here, a high achiever was defined as a student who performed in the top 10 percent *of his particular grade and school*.

Students in **Cohort 3** were followed from third grade through fifth grade. This cohort consisted of 235,709 students in math, of whom 21,291 were high flyers, and 250,550 students in reading, of whom 22,868 were high flyers. The students were drawn from 952 schools in thirty states.

Students in **Cohort 4** were followed from sixth grade through eighth grade. This cohort consisted of 184,674 students in math, of whom 17,425 were high flyers, and 210,577 students in reading, of whom 20,309 were high flyers. The students were drawn from 410 schools in twenty-nine states.

To assure that these students reflected their larger school populations, the sample included only those students who attended a school in which 80 percent of enrolled students overall and a minimum of thirty students in each grade were tested with NWEA assessments. In order to increase the sample size, we combined students from three successive years into each cohort (Figure A-1). For example, if a school had eight high achievers in third grade in 2005-06, six high achievers in third grade in 2006-07, and nine high achievers in third grade in 2007-08, we analyzed the total group of students, twenty-three high achievers, as a single entity (Figure A-1).

FIGURE A-1

Cohorts 3 and 4 by Grade and Year

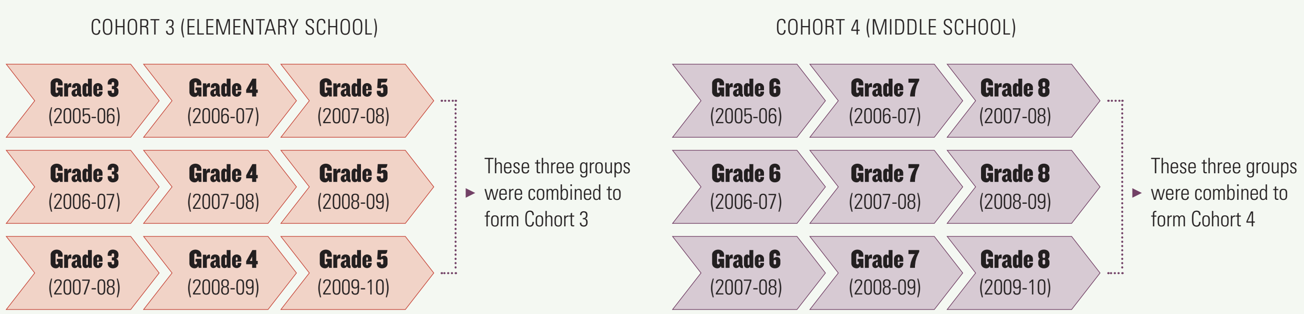


Table A-6 illustrates the gender and ethnic composition of the two cohorts. The proportions of males to females in Cohorts 3 and 4 were similar to the gender compositions of Cohorts 1 and 2, but the proportions of minority students in Cohorts 3 and 4 were higher than the proportions reflected in Cohorts 1 and 2. This is largely due to the two distinct definitions of "high achiever." Because high achievers in Cohorts 3 and 4 comprised those students who performed in the top 10 percent of their individual grades and schools, we expected a higher representation of minority students in those cohorts.

TABLE A-6

Demographics of High Achievers in Cohorts 3 and 4 (Initial Year of Study)

	Gender		Ethnicity		School Poverty	
	Female	Male	Minority	Non-minority	High Poverty	Low Poverty
Cohort 3 Math	39.5%	60.5%	16.2%	83.8%	27.8%	72.2%
Cohort 3 Reading	52.3%	47.7%	18.0%	82.0%	26.8%	73.2%
Cohort 4 Math	40.1%	59.9%	12.5%	87.5%	23.7%	76.3%
Cohort 4 Reading	51.7%	48.3%	13.8%	86.2%	22.9%	77.1%

Hierarchical Linear Modeling

This separate line of analysis employed hierarchical linear modeling (HLM) to evaluate the results for high achievers in high- and low-poverty schools.[15] HLM is an advanced form of linear regression. It is preferred over standard linear regression in circumstances in which data are nested, while linear regression might be preferred in circumstances where data are independent. For example, in a simple study in which one wanted to assess the relationship between the morning calorie intake and stamina of a group of recreational runners, simple regression would normally be sufficient, because the runners are not nested in a set of groupings (like running teams, for example).

These conditions are rarely present in educational data. In this particular study, there is not one single assessment but rather a series of assessments that are *nested*, if you will, within each student. In addition, each student nests within a grade level and school.[16] In these conditions, the individual test events of a student are likely to be highly correlated with one another, and the test events of all students within a school are also likely to be correlated. By accounting for nesting within this analysis, we can draw better inferences about the relationship between our high-achieving students' growth and factors such as gender, ethnicity, or school poverty rate that may influence performance and growth. Another reason for using HLM is that schools vary greatly in maintaining their high achievers' growth. Hence, by accounting for school-level random effects (variability across schools), the estimation of fixed effects (the effects of school poverty and location) are more precise.

The following three-level HLM was used to model the relationship between school poverty rate and school achievement and growth. It applied to Cohorts 3 and 4, separately examining performance and growth rates for elementary school mathematics, elementary school reading, middle school mathematics, and middle school reading. In the model, school poverty rate refers to the percentage of students who are eligible for free or reduced-price lunch at each school.

Level One: Test events (repeated measures)

Level one is an individual growth model of academic achievement at time t for student i in school j.

$$Y_{tij} = \pi_{0ij} + \pi_{1ij}\,(ACADEMIC\ YEAR)_{tij} + e_{tij}$$

15 Additional models that examine characteristics other than school poverty will be described in a forthcoming report.

16 In truth, the student is nested inside a classroom, a grade level, *and* a school. One limitation of the data set used for this study is that we could not consider classroom effects because we did not have reliable data about the particular classrooms to which students were assigned.

Level Two: Students (individual growth trajectory)

We did not include any student-level variables to focus on the relationship between school poverty rate and school achievement and growth.

$$\pi_{0ij} = \beta_{00j} + r_{0ij},$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij},$$

Level Three: Schools

$$\beta_{00j} = \gamma_{000} + \gamma_{001}X_{(FRL\%)j} + \mu_{00j}$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101}X_{(FRL\%)j} + \mu_{10j}$$

ABOUT THE AUTHORS

Yun Xiang is a research specialist at the Northwest Evaluation Association (NWEA). Yun's research interests include growth modeling, program evaluation, and the handling of large data sets. Her recent published works and presentations include an investigation of achievement gaps and summer learning loss, applications of growth modeling in evaluating schools and addressing student mobility, an analysis of school productivity by comparing different value-added models, and research involving the professional satisfaction of teachers. Yun holds a PhD in Educational Research, Measurement, and Evaluation from Boston College, an EdM in Curriculum and Instruction from Boston University, and a BA in English Education from Xiangtan Normal University, China.

Michael Dahlin is a research specialist with the Kingsbury Center at NWEA. He joined NWEA in 2007 as a research associate after working as a research analyst for the Oregon Department of Human Services. He has also worked as an adjunct and visiting psychology professor at Pacific University. Mike's recent NWEA work includes extensive research and reporting to examine the effectiveness of the No Child Left Behind Act. His professional affiliations include the American Psychological Association, the American Psychological Society, the Society for Research in Child Development, the American Educational Research Association, and the National Council on Measurement in Education. Mike holds a PhD in Developmental Psychology from Pennsylvania State University, an MS in Psychology from Western Washington University, and a BA in Biology from Pomona College.

John Cronin has served as director of the Kingsbury Center at NWEA since 2008. As the repository for the largest cross-state longitudinal database of student achievement in America, the Kingsbury Center engages in research and consulting partnerships with organizations such as Vanderbilt University, Notre Dame University, the Wisconsin Center for Educational Research, the Wisconsin Center for Value-Added Measurement, KIPP, Teach for America, and others to bring this resource to bear on important educational issues. John is the primary author of several important studies related to the rigor of educational standards, including two prior studies for Fordham, *The Proficiency Illusion* and *The Accountability Illusion*. John has served as an educator and consultant for thirty-three years and holds a PhD in Educational Studies from Emory University.

Robert Theaker joined the Kingsbury Center in 2011 as a senior research associate. He previously served as director of data analysis for Central Michigan University's Center for Charter Schools, where he was in charge of establishing a data warehouse, school evaluation reporting, and analysis. Prior to joining Central Michigan University, he served as the senior manager of assessment and measurement for National Heritage Academies. In this role, he led one of the most innovative and sophisticated implementations of NWEA's Measures of Academic Progress in the United States. In Bob's thirty-seven years of experience, he has been a statistics professor, teacher, and school administrator, and has presented numerous seminars in data-driven decision making.

Sarah Durant is a senior research associate at the Kingsbury Center at NWEA, where she helps partners understand and use their performance data to make decisions about policy and practice. Prior to joining NWEA in 2009, Sarah worked as a principal performance analyst for Multnomah County and as a research analyst for the Oregon Department of Education. In addition to policy research and data analysis, her skills include performance measurement, program evaluation, project management, and organizational design. Sarah holds an MPP from the University of California at Berkeley and a BA in Linguistics from Boston University.