

TRACKING AND DETRACKING:

HIGH ACHIEVERS IN MASSACHUSETTS MIDDLE SCHOOLS



TOM LOVELESS



FOREWORD BY CHESTER E. FINN, JR. AND AMBER M. WINKLER
THOMAS B. FORDHAM INSTITUTE
DECEMBER 2009

Contents

| | |
|---|----|
| Foreword..... | 1 |
| Executive Summary..... | 5 |
| Introduction..... | 7 |
| Research on Tracking and Achievement..... | 11 |
| Methods and Research Questions..... | 15 |
| Results From the 2009 Survey of Massachusetts Middle Schools..... | 17 |
| Summary and Discussion..... | 26 |
| Conclusion..... | 28 |
| Appendix..... | 32 |

The Thomas B. Fordham Institute is a nonprofit organization that conducts research, issues publications, and directs action projects in elementary and secondary education reform at the national level and in Ohio, with special emphasis on our hometown of Dayton. 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, DC 20036. This report is available in full on the Institute's website. The Institute is neither connected with nor sponsored by Fordham University.

Cover image ©Imageplus/Corbis

Foreword

Chester E. Finn, Jr. and Amber M. Winkler

By 2011, if the states stick to their policy guns, all eighth graders in California and Minnesota will be required to take algebra. Other states are all but certain to follow. Assuming these courses hold water, some youngsters will dive in majestically and then ascend gracefully to the surface, breathing easily. Others, however, will smack their bellies, sink to the bottom and/or come up gasping. Clearly, the architects of this policy have the best of intentions. In recent years, the conventional wisdom of American K-12 education has declared algebra to be a “gatekeeper” to future educational and career success. One can scarcely fault policy makers for insisting that every youngster pass through that gate, lest too many find their futures constrained. It’s also well known that placing students in remedial classes rarely ends up doing them a favor, especially in light of evidence that low-performing students may learn more in heterogeneous classrooms.

Yet common sense must ask whether *all* eighth graders are truly prepared to succeed in algebra class. That precise question was posed in a recent study by Brookings scholar Tom Loveless (*The 2008 Brown Center Report on American Education*), who is also the author of the present study. He found that over a quarter of low-performing math students—those scoring in the bottom 10 percent on NAEP—were enrolled in *advanced* math courses in 2005. Since these “misplaced” students are ill-prepared for the curricular challenges that lie ahead, Loveless warned, pushing an “algebra for all” policy on them could further endanger their already-precarious chances of success.

When American education produced this situation by abolishing low-level tracks and courses, did people really believe that such seemingly simple—and well-meaning—changes in policy and school organization would magically transform struggling learners into middling or high-achieving ones? And were they oblivious to the effects that such alterations might have on youngsters who were already high-performing?

This study examines tracking in Massachusetts—arguably the leading state in “reforming” that traditional practice of U.S. schools—and the changes that occurred there between 1991 and 2009. (By “tracking,” we mean the practice of grouping students into separate classes based on achievement, not the tracking of yesteryear that placed high school students into academic, general, or vocational tracks based primarily on their IQs, origins, or career aspirations. Further, the study targets within-subject within-grade tracking at the middle school level, a practice that is not nearly as draconian or immutable as the more antiquated tracking policies of the past.)

Of particular interest was how tracked and untracked schools measured up when it came to producing high-achieving students.

The short answer is...tracked schools did better. But there aren’t many of them anymore. Loveless finds that middle schools have mostly done away entirely with tracking in English language arts

(ELA), history, and science, but that this practice endures, albeit with fewer tracks than two decades ago, in mathematics. Further, “detracking”—reducing the number of subject-area courses offered in a given grade in a given school—may adversely affect high-achieving youngsters in math (but not ELA; history and science achievement were not analyzed). Middle schools with more tracks have significantly more math pupils performing at the *advanced* and *proficient* levels and fewer students at the *needs improvement* and *failing* levels. What’s more—and this is important—a mirror image of this trend appears for detracked schools. In other words, they also have more *failing* and *needs improvement* math students than schools with two or three tracks. In fact, a declining number of failing math students is linked to each additional track in the school, i.e., schools with one track have the most failing math students (26 percent), schools with two tracks fewer (20 percent), and schools with three or more tracks fewer failures still (14 percent).

Nor is that all. Loveless dug further and found that, when schools’ socioeconomic status is held constant, each additional track in eighth-grade math (up to three) is associated with a 3 percentage-point rise in students scoring at the *advanced* level. That means the advantage for a school offering three tracks instead of one is associated with a 6 percentage-point gain in the number of students performing at high levels.

That may not look like much to you, but it could have considerable impact. The average middle school in Massachusetts can boast just 18 percent of its students at the advanced level in math—and many schools have *no* students who reach such altitudes. Increasing the proportion of high achievers by six percentage points—that’s eighteen kids in an eighth-grade class of 300—could have a dramatic impact on other students, as well as on the school’s culture. At a time when the United States needs every academic high-flyer we can find or produce to buttress our domestic and international prospects, that’s certainly worth attention.

An analysis such as this cannot *prove* that tracking is the direct cause of the observable difference in schools. But the association is clear: More tracks, more high-performing kids and fewer failures. Fewer tracks, fewer high-performing kids and more failures.

Which schools track and which do not? The evidence reported here indicates that urban schools serving mostly poor children are more likely to have diminished or abolished tracking while suburban schools serving children from more prosperous backgrounds are more apt to have retained it, despite pressure from “experts” and ideologues to do away with it.

This is a point worth pausing over. Schools that eliminated or reduced curricular tracking thought they were doing this to benefit needy and minority kids. They succumbed to the accusation that they and their evil old tracking policies had been complicit in harming such youngsters. In fact, Kevin Welner, Professor of Education at the University of Colorado-Boulder and leading opponent of tracking, says that it “consigns children to substantially unequal education opportunities.” Further, he urges schools to adopt a vigilant and proactive position against tracking:

Whatever the underlying motives for detracking, moving forward successfully with the reform requires direct and conscious confrontation of the normative and political opposition likely to arise...the political majority in many communities will defend inequalities, even in the face of powerful evidence of unfairness that would sway an unbiased observer. Left unopposed, they will exercise political power sufficient to end or undermine the reform. Equity-minded reforms like detracking often implicate core values and self-interests and therefore give rise to overt political action.¹

THE NEW TRACKING?

It's not hard to see why tracking became taboo. (Loveless details this history below as well as the mixed results that followed.) But as the old saying has it, the more things change, the more they stay the same. Even as tracking diminished, American education seems to have picked up more politically acceptable alternatives in the form of school choice, individualized learning, and differentiated instruction.

Is school choice another form of tracking? Think about it. The choice movement rests in significant part on the family's right to choose the best education for its children and, more and more, on the *school's right* to make certain academic demands on children and families. Consider how some of our most-esteemed "brand name" charter outfits (e.g., KIPP, Achievement First, Uncommon Schools, YES-Prep) operate. They have extended school days and years, rigorous behavioral expectations, homework requirements, parental contracts, etc. No, they're not exactly "selective." But neither are they schools for the faint of heart. And while their mission is to provide disadvantaged students with a top-notch education—a mission the schools named above are generally fulfilling—they do cater to the *motivated* disadvantaged. This could be viewed as tracking by school, not within school or by subject. Note, though, that school choice is parent-centered and volitional on the part of the learner, whereas old-style "tracking" was usually involuntary, determined by educators and administrators.

Whereas choice involves selecting schools, differentiated instruction happens within a school. A teaching theory that says instructional approaches should vary in relation to individual students' needs, capacities, learning styles, and prior academic achievements, it ends up in much the same place as within-class ability grouping. The University of Virginia's Carol Tomlinson, author of *The Differentiated Classroom* (1999), describes it thusly:

In differentiated classrooms, teachers...[do not] assum[e] one student's road map for learning is identical to anyone else's....They work diligently to ensure that struggling, advanced, and in-between students think and work harder than they are meant to....[They] use time flexibly, call upon a range of instructional strategies, and become partners with their students to see that both what is learned and the learning environment are shaped to the learner.²

Ask teachers if differentiation is a worthwhile use of students' time and most will enthusiastically agree, even though they have little time themselves to craft such lessons and not all of them can pull it off. Still, differentiation is one of today's politically-acceptable alternatives to tracking. Students reap the benefits of heterogeneous and diverse classrooms while their individual needs are addressed via ability-based clusters, tiered assignments, varied texts or resources, and other adaptations.

In the end, one truth survives the test of time. Even if schools *say* they've gotten rid of tracking, they haven't gotten rid of the problems that gave rise to it nor have they abolished the need to solve those problems for the children's own sakes. Most Americans recognize that children are distinct individuals who bring very different backgrounds, temperaments, cognitive attainments, and earlier academic achievements with them to school. They deserve schools that honor and respond to such differences.

Bottom line number 1: American education needs to care more about taking students to the next level and less about how we get them there. Anna Penny, a former teacher in New York City, said as much in the *New York Daily News* this past summer: "Anyone who has ever taught knows that kids progress at dramatically different speeds in different subjects. When our schools resist tracking even when it's clearly needed, they wind up valuing [heterogeneous] classrooms over effective ones."³

Bottom line number 2: In the name of equity, gap closing, political correctness, and leaving no child

behind, American education has been a bit too willing to neglect its higher-performing students and the school arrangements that best meet *their* needs. A recent report by the National Association for Gifted Children finds that eighteen states can't even tell us how many children have been identified as gifted within their borders. Further, the vast majority of gifted children are placed in regular classrooms (no surprise, given Loveless's findings), where most teachers are not trained in gifted education. In fact, thirty-six of forty-seven responding states don't require regular teachers to have training in gifted education at any point in their careers, nor do most teacher-preparation programs include coursework on gifted learners. That's more than unfortunate for high-achieving youngsters and the ill-equipped teachers who teach them, but it's also damaging to our long-term national interest.

Progressive friends, forgive us for borrowing your phrase, but you were right. When it comes to educating children as when selling shoes, one size does not fit all.

This is the fourth in a multi-part investigation of the educational state of high-achieving students in the No Child Left Behind era conducted by the Thomas B. Fordham Institute. The project is supported by the John Templeton Foundation and our sister organization, the Thomas B. Fordham Foundation. The first two studies, which included both an analysis of National Assessment (NAEP) data and a national teacher survey, were released in June 2008. Jointly titled *High-Achieving Students in the Era of NCLB*, they found that, although the lowest-achieving students made rapid gains from 2000 to 2007, the progress of top students during the same period was, as author Loveless put it, "languid." Further, the survey results, as distilled by the Farkas Duffett Research (FDR) Group, found that today's teachers pay considerably more attention to their lowest-performing students than to their highest. The third study, also conducted by FDR, surveyed Advanced Placement (AP) teachers and found that, even though they believe the quality of the AP program is holding up in the face of tremendous expansion, they also see troubling signs from students who overestimate their abilities and from parents who are overly zealous about enrolling their daughters and sons in AP courses.

Now we're back with study number four, which examines tracking at the middle-school level in Massachusetts as well as the performance of high-achieving students in tracked and untracked schools. To do the heavy lifting, we again enlisted Tom Loveless, Senior Fellow in Governance Studies at the Brookings Institution. Not only do we think highly of his scholarship, but some of his earlier work in Massachusetts (and California) paved the way for this analysis.

This report, like those that came before it, benefited from the expert counsel of an independent review committee that included Cynthia Brown, Director of Education Policy at the Center for American Progress; Paul Gross, Professor Emeritus at the University of Virginia; Frederick Hess, Director of Education Policy Studies at the American Enterprise Institute; Stephanie Pace Marshall, Founding President and President Emerita of the Illinois Mathematics and Science Academy; and Joyce Van Tassel-Baska, Executive Director of the Center for the Gifted and Talented and Smith Professor of Education at the College of William and Mary. The panel reviewed the study methods and results and provided Loveless and Fordham with helpful feedback. However, not every suggestion made by every reviewer could be accommodated. (And, of course, the views expressed in this foreword may not reflect those of the review committee or Dr. Loveless.)

We'd also like to thank the Fordham team for diligently seeing this project to completion and dissemination, particularly Associate Editor and Policy Analyst Stafford Palmieri, media and public affairs staffers Amy Fagan and Laura Pohl, Vice President Michael Petrilli, research intern Jack Byers, and designer Alton Creative.

Executive Summary

This study examines tracking—the practice of grouping students into separate classes or courses based on their prior academic achievement—at the middle-school level, and the percentage of high-achieving students in tracked and untracked schools. It focuses on Massachusetts, a leader in “reforming” tracking, and the changes that have occurred there over eighteen years (1991-2009).

Specifically, the study addresses the following:


- 1 How has tracking changed since the 1990s?
- 2 Why have some schools detracked while others are resisting such reform?
- 3 Are there differences in the number of high-achieving students associated with tracked and untracked schools? Specifically, do tracked and untracked schools produce similar percentages of students reaching the “advanced” level on the Massachusetts Comprehensive Assessment System (MCAS)?

Key findings:

- 1 **Tremendous change has occurred in tracking since the 1990s.** Nearly twenty years ago, eighth-grade students attended tracked classes for most of the day. They now spend most of their day in detracked classes. An eighth grader in the early 1990s attended middle schools offering at least two distinct tracks in (each of) English language arts, history, and science. Mathematics courses were organized into three or more tracks. The eighth grader of 2008, however, attended schools with much less tracking. English language arts, history, and science are essentially detracked, i.e., schools typically offer a single course that serves students at every level of achievement and ability. Mathematics usually features two tracks, often algebra and a course for students not yet ready for algebra.
- 2 **Several factors influence tracking policy.** Why do some schools continue to track while others have detracked in favor of heterogeneous grouping? Several factors are influential. Schools serving predominantly poor populations are more likely to have stopped tracking. Those serving students from higher socioeconomic backgrounds are more apt to have retained tracking. A school’s grade configuration also matters. Mimicking high schools, middle schools serving grades 7 and 8 are more likely to embrace tracking compared with their grade 5-8 and 6-8 counterparts. Finally, schools in which parents wield greater influence tend to keep tracking in place, as do schools in communities where local school boards have discussed the topic.
- 3 **Detracking is more prevalent in urban, high poverty schools.** Urban schools with children of lower socioeconomic status (SES) are more likely to detrack than suburban schools with children of higher SES. Consequently, the risks associated with detracking are concentrated in urban schools serving large numbers of poor, low-achieving children.

- 4 Detracking Carries Risks for High-Achieving Students.** The study compared the percentage of students achieving at the advanced level on the MCAS in tracked and untracked schools. There was no difference in English language arts. However, with school-level SES held constant, each additional track level in eighth-grade math (up to three) is associated with a 3 percentage-point gain in students scoring at the advanced level. That means a school with 200 eighth graders that offers at least three levels of math is typically attended by twelve more students scoring at the advanced level than a detracked school of similar size and SES status. The study cannot link tracking policy causally to this outcome but, combined with previous research on the effects of detracking, it serves as a caution to schools and policymakers that detracking may adversely affect high-achieving students.

Introduction

 This study traces the evolution of tracking in middle schools and examines the performance of high achieving students in tracked and untracked schools. It focuses on Massachusetts, one of the leading states in “reforming” tracking, and the changes that have occurred there over eighteen years (1991-2009). And it documents a dramatic transformation: Middle schools have significantly altered the way in which they group students into classes for instruction, an issue of particular interest to those concerned with the educational fate of gifted and talented youngsters.

Tracking is the practice of grouping students into separate classes based on achievement. Tracking policy is typically made at the local level and therefore does not attract much national media attention. Yet it annually affects more than *14 million* young people in the middle grades alone.⁴ Further, it is the type of policy that controls critical aspects of education—the classes students take, the curricula they are taught, the peers with whom they learn, and the teachers who instruct them. These elements of education largely define a child’s experience in school—and all of them have changed significantly because of tracking reform.

BACKGROUND

In the middle of the twentieth century, traditional tracking systems were rigid and deterministic. Students and families had little input regarding classroom placements. Based predominantly on IQ scores, schools assigned youngsters to tracks⁵—academic, general, or vocational—that cut across all subject areas, ignored students’ individual strengths and weaknesses in particular subjects, made unjustifiable assumptions about children’s destinations in life, and systematically discriminated against pupils from disadvantaged backgrounds.⁶

That system eventually gave way to a more open form of tracking—one allowing, for example, a precocious reader to take advanced ELA classes while still enrolling in a less challenging math class. Reliance on IQ tests fell by the wayside and track placement came to be based on past performance and achievement test results. Students who did well in a particular subject could take a more advanced class the following year. Parents could challenge course placements and insist that their children take the classes that they deemed best.

But this more flexible form of tracking also came under fire, most notably in Jeannie Oakes’ 1985 book, *Keeping Track*. Even less rigid tracking, she insisted, made distinctions among students that often reflected their socioeconomic backgrounds. (To be sure, allowing parents a say in track assignment favored school-savvy families.) Oakes and others charged that the changes that had been made to tracking were largely cosmetic and that its fundamental unfairness had not been ameliorated. Citing the social reproductionist theories of Samuel Bowles and Herbert Gintis, among others, Oakes argued that tracking exists to maintain the existing distribution of power and privilege in society, one

that is stratified to meet the demands of capitalism. Summarizing her study of twenty-five schools' tracking systems, Oakes claimed that "Track levels in schools, reflective of social and economic groupings in society, were provided differential access to school knowledge in such a way that the children of more powerful societal groups had greater access to the kind of knowledge that may, in turn, permit them greater access to social and economic power." And she demanded that schools "relinquish their role as agents in reproducing inequities in the larger society."⁷

By 1990, a push to abolish tracking was underway across the land. Condemnations of tracking came from such powerful groups as the National Governors Association, the ACLU, the Children's Defense Fund, the Carnegie Corporation, the College Board, and the NAACP Legal Defense Fund.⁸ The middle school movement, as it was called, also gained traction during this time. It embraced a long list of reforms reflecting progressive educators' longstanding desire to alter the education of young adolescents. These reforms included converting 7th-9th grade junior-high schools into 5th-8th or 6th-8th-grade middle schools; balancing adolescents' social, emotional, and academic needs in the school curriculum; recruiting into middle schools teachers with elementary training, who would presumably possess more child-centered philosophies than their subject-specialist counterparts in high schools; and promoting project-based learning and other student-led forms of pedagogy over traditional, teacher-led instruction.⁹ Given the overlap between progressivism and egalitarianism, it is no wonder that tracking reformers found fertile ground in middle schools.

Their cause was bolstered by several high-profile publications in the late 1980s and early 1990s that promoted middle school reform. These included the Carnegie Corporation's *Turning Points* (1989) and two state policy documents, California's *Caught in the Middle* (1987) and Massachusetts' *Magic in the Middle* (1993). About this same time, a battle over detracking erupted in Cambridge, Massachusetts. The *Boston Globe* ran an editorial cartoon illustrating the fear that high achieving children would receive instruction far below their capabilities. The cartoon showed Einstein at a chalkboard, sadly writing $1 + 1 = 2$ and $3 + 3 = 6$ on the board, with the caption: "Panel Recommends End of Tracking in the Cambridge Schools."¹⁰

Opposition to detracking came from groups representing the parents of gifted and talented students. They argued that bright students should have advanced classes.¹¹ The National Association for Gifted Children (NAGC) passed a resolution in support of grouping for advanced students.¹² But these arguments were swept aside by the charge that singling out high achievers in honors or accelerated classes is inherently elitist and equity can only be attained when all students receive the same curriculum. Many middle schools—but not all—steadily reduced the number of track levels offered in academic subjects.

In most schools, today's middle school parents, like their predecessors fifty years ago, have few choices regarding the courses their children take. For most academic subjects, students are placed in a single, heterogeneously-grouped class offering the same curriculum and pace of instruction to all. The tracking that still occurs is generally confined to mathematics, but even options in math have been curtailed. Some middle schools have resisted the trend towards detracking, but they are in the minority, and in many cases, they must buck state and district recommendations to maintain their tracked systems.

What is behind this historic controversy about how young people are prepared for high school? Why is it that some middle schools have detracked while others continue tracking? Is there any evidence that one approach is associated with better school performance than the other, specifically when it comes to high achievers?

ABILITY GROUPING AND TRACKING: WHAT'S THE DIFFERENCE?

Tracking and ability grouping refer to ways of grouping students for instruction. Used interchangeably, the two terms cause confusion since they can refer to several different types of grouping strategies and have changed meaning over the past couple of decades. Let's start by defining terms and describing how schools use tracking and ability grouping.

Ability Grouping

Ability grouping is the practice of organizing students into small groups for instruction. In contrast to tracking, students are grouped *within* a class, not in separate classes. Ability grouping is quite common in American elementary classrooms, especially in ELA. (Many readers may recall “bluebirds” and “redbirds” and such from their early schooling.) Based mainly on the results of diagnostic tests, students are assigned to reading groups that vary in both level and pace of instruction. Students who are better readers are placed in groups that read more difficult texts and progress more quickly through material; students who are struggling receive a slower paced instruction pitched at a lower level. The teacher rotates among the groups. Groups that are not receiving instruction typically work on practice or extension activities at their seats. Sometimes students will continue working together in cooperative groups.¹³ Although ability grouping has detractors, it is not as controversial as tracking.¹⁴

Tracking

Tracking usually takes place in middle and high schools. Implemented on a subject-by-subject basis, it refers to grouping students into separate classes based on achievement. The system in Montgomery County, Maryland provides a good example of the organization of high school ELA classes. Ninth graders are placed into four levels:

- 1 English 9A/B¹⁵ (Honors)
- 2 English 9A/B (grade level)
- 3 Developmental Reading (on or below grade level)
- 4 Basic Reading (two years or more below grade level)

Tenth grade has the same structure, as do the eleventh and twelfth grades, except both of the latter add an Advanced Placement (AP) option in English Language and Composition.

Across the country, mathematics classes are usually grouped by topic, particularly in high school. Courses are organized hierarchically, meaning that a student's placement largely depends on the course taken in the previous year. A high-achieving eighth grader who successfully completes Algebra I will usually take Geometry as a ninth grader, Algebra II in tenth grade, and Trigonometry, Math Analysis, or Pre-Calculus as a junior. The sequence culminates in Calculus, perhaps Advanced Placement Calculus, in the senior year. Less-advanced pupils may complete only parts of that sequence. Because of the sequential nature of math courses, critics of tracking rightly note that where a student starts—as early as seventh grade—often determines where he or she finishes at the end of high school.

Middle School Tracking

Middle school is where tracking begins, providing a bridge between the heterogeneously grouped classes of elementary school and the tracked classes of high school. Middle schools span sixth through eighth grades, although some offer fifth grade. Many middle schools heterogeneously group students

in sixth grade and do not track at all until seventh or eighth grades and then only in mathematics.

If courses in ELA or history are tracked, an “honors” or “advanced” offering is typically included as an option for high achievers. Everyone else takes a class called “8th Grade English” or “English 8.” Some schools may try to remove status-laden distinctions between the more and less difficult classes by calling one “advanced” and the other “honors,” but real curriculum differences remain. Classes for remedial students, once quite common, fell out of favor after the detracking movement. Critics argued that they stigmatized students and were often taught by the least capable teachers. Further, with student behavior problems concentrated in remedial classrooms, those classes became extremely difficult to manage.¹⁶

In mathematics, offering an eighth-grade algebra course for high achievers is becoming the conventional practice, with students not yet ready for algebra taking pre-algebra or a grade level course (frequently called “8th grade math”). In fact, more eighth graders now take algebra than any other math course: About 16 percent of eighth graders were enrolled in Algebra I in 1990; that nearly doubled to 31 percent in 2007.

Further, algebra courses have become the focal point in contemporary detracking efforts—so much so that California and Minnesota have mandated universal algebra for eighth graders beginning in 2011.¹⁷ Advocacy groups favoring “algebra for all” argue that eighth-grade algebra serves as a gatekeeper, allowing some students to proceed on a glide path towards college while keeping others out. They completely oppose tracking in mathematics, which, as will be shown in this study, continues to resist tracking reform.

Research on Tracking and Achievement

Research on tracking focuses on two questions. The first is, What is the effect of tracking on average achievement? Here, analysts compare mean achievement in tracked and detracked settings in an effort to determine whether tracking has an overall effect on pupil learning. From a policy perspective, the issue is whether the test scores of an entity—nation, state, district, or school—are higher or lower because of its tracking policy.

The second question is, Do high, middle, or low ability students benefit from tracking? In this case, analysts examine the distributional effects of tracking, i.e., whether a particular group of students benefits from or is harmed by the practice. The same question about distributional effects can be applied to *detracking*—who does well or poorly when schools group students heterogeneously?

EFFECT ON AVERAGE ACHIEVEMENT

Researchers generally agree that tracking has little effect on average achievement.¹⁸ The best evidence comes from randomized experiments. Mosteller, Light, and Sachs (1996) found fifteen studies with random assignment of students to grouped and ungrouped conditions.¹⁹ No significant differences were detected between the two treatments. The same conclusion has been drawn from summaries of studies that meet minimum standards for research quality. In the two most widely cited of these meta-analyses, James Kulik reports an effect size of +0.03 and Robert Slavin reports an effect size of +0.00. In other words, no discernible effect. A school that switches from a tracked to a detracked curriculum or vice versa should not experience any significant change in average test scores.

Only one randomized experiment has been conducted since the review by Mosteller and colleagues: an interesting study in Kenya involving first graders in 121 schools.²⁰ A class-size reduction program permitted all of the schools in a particular region to hire new teachers mid-year, doubling the number of first grade classrooms from one to two. Half of the schools (randomly selected) created upper- and lower-ability classes. Class assignment in the second semester was based on students' grades in the first semester. The other schools created two heterogeneously-grouped classes. The program continued for eighteen months, through second grade. Students in the tracked schools gained 0.14 standard deviations on achievement tests compared to students in the detracked schools.²¹ The gains were similar for high, middle, and low performing students, and the advantage persisted on tests administered a year after the experiment concluded.

The Kenya study has limitations. First and second grades may be unique in regards to tracking effects and, as noted above, tracking is primarily practiced in secondary schools in the U.S., not in primary grades. Also, the classes in the Kenyan schools held over thirty students, even after the class size reduction. Heterogeneity in such large classes may create challenging conditions for teachers that are more intense than in smaller classes. Analysts conducting the study speculated that teachers found it difficult to deliver the appropriate level of curriculum to students in the heterogeneous classes. But no

measures of curriculum or instruction are provided.

In most of the experimental studies reviewed by Mosteller et al. both high and low groups received the same curriculum. That makes sense in controlling for curriculum effects, but as any educator who favors tracking will argue, the point of tracking is to vary the curriculum—to give highly able students more challenging work while struggling students get the work that will help them catch up.²² It is not surprising that tracking's effects appear muted when the same curriculum is offered to all groups, an out-of-date system called XYZ grouping.²³ Finally, entirely detracked schools have always been rare so there are very few studies comparing large numbers of tracked and detracked schools over time.

EFFECT ON THE DISTRIBUTION OF ACHIEVEMENT

Researchers generally agree that tracking increases inequality, primarily by boosting the achievement of students who are in a high track while having a negative effect on those in a low one. The finding is based on comparing the learning of students who differ only in track assignment but have the same initial test scores. Since the positive and negative effects offset each other, the net outcome is consistent with the finding that tracking has no impact on average achievement.

Note, though, that this inference is made primarily from studies analyzing differences within tracked systems. Studies rarely compare the distribution of achievement within tracked and detracked schools because detracked schools are so rare. Still, it appears that low-performing students may learn more when placed in heterogeneously grouped classes instead of low tracks.²⁴

Does detracking pose risks for high achievers? Kulik and Kulik's 1992 meta-analysis detected significant benefits for high-performing students in classes with an enriched or accelerated curriculum.²⁵ Students in accelerated classes outpaced similar pupils in non-accelerated classes by 0.87 standard deviations (twenty-three studies), a whopping amount for any educational intervention. Students in enriched classes (twenty-five studies) gained 0.41 standard deviations compared to similar students in mixed-ability classes studying the regular curriculum.

Critics of these studies argue that selection effects render the comparisons invalid. Students were not randomly assigned to treatments. Instead, bright students who were likely to succeed with an accelerated or enriched curriculum were probably placed in the special classes; those who were less apt to benefit, even if they did have high test scores, were placed in the regular classes. Indeed, the problem of selection plagues all non-experimental studies comparing students in different tracks, including those that control for initial test scores and find that students benefit in high-ability tracks and suffer a loss in low-ability tracks. Researchers do the best they can to control for initial test-score differences but educators consider more than test scores when assigning students to track levels. Work habits, motivation, engagement with the subject matter, prior academic record, and parental wishes may also be considered. These variables typically go unmeasured in the databases used by tracking researchers.

International studies of tracking run into additional difficulties. In a widely cited study, Hanushek and Woessmann (2005) compare nations that track early (before age fifteen) with those that track later.²⁶ The authors conclude that nations with early-tracking practices exhibit wider distributions in achievement—in other words, greater inequality—and lower test scores in high school. However, the study confounds different forms of tracking. European and Asian tracking takes place primarily among schools at the secondary level, with students grouped heterogeneously through eighth or ninth grade, at which point exam results determine the high schools that they will attend. In other words, students are tracked by schools—some offering very different programs—not by classes within

schools. Most but not all nations practice this type of tracking. As discussed above, the dominant U.S. model is for youngsters to attend comprehensive middle and high schools serving students of all abilities. Tracking occurs within schools and varies by subject area. It is debatable whether the effects of these two types of tracking are comparable.²⁷

Two studies of American students have directly compared the distribution of achievement in tracked and detracked classes. They suggest that detracking is not without costs. Rees, Argys, and Brewer (1996) analyzed data from the National Educational Longitudinal Study of 1988 (NELS: 88) and concluded that high achievers pay a price for placement in heterogeneous math classes. Their achievement falls by approximately 8.4 percent. Low performing math students do get a boost from heterogeneous classes, but their gain of 8.6 percent is nearly the same as the achievement loss suffered by high-ability students. Students of average ability lose a little, too. Their achievement dips by about 2.0 percent. Greater equality is attained—the gap between high and low achievers shrinks—but it is accomplished by depressing achievement at the top.²⁸

Allensworth et al. (2008) investigated Chicago's 1997 effort to make college preparatory classes mandatory for all ninth graders. In math, the policy meant the elimination of remedial courses and placement in algebra or higher for all students. In ELA, all students were required to take English I instead of remedial ELA. Researchers estimated the impact of the policy on reducing inequities associated with tracking. They analyzed nearly a decade of data via an interrupted time series cohort design.²⁹ And they found one benefit: More students received credit in Algebra I and English I, not surprising considering that that all ninth graders were now compelled to enroll in the two courses.

However, several alarming effects were also uncovered. Failure rates in the courses increased, grades declined, and test scores were unchanged. Students who would otherwise have taken remedial classes were no more likely to obtain advanced math credits beyond Algebra II, nor were they more likely to graduate from high school or to attend college. Absenteeism increased among average and high ability students. The authors questioned whether the benefits suggested by previous research actually materialize when detracking is implemented in a system as large as Chicago. They also expressed concerns that detracking may depress the achievement of middle and high achievers.

Let's summarize the research on tracking and achievement. The best research has found no effect of tracking on average achievement. The literature on tracking's effect on the distribution of achievement is less settled. A consensus exists that tracking increases inequality by benefitting high achievers and having negative effects on low achievers. Several studies have suggested that detracking would help low-achieving students, meaning they would learn more in heterogeneously grouped classes than in low-level tracks. This finding, however, is rarely based on direct comparisons of detracked and tracked systems. Two studies making such comparisons found that any gains by low achievers are offset by negative effects for middle- and high-performing students.

THE TRACKING WARS

Despite the mixed evidence in its favor, the anti-tracking movement was a powerful political force in the 1990s. I spent most of that decade studying detracking in middle schools, primarily in California and Massachusetts, where the state education departments pushed middle schools to detrack. I was a doctoral student at the University of Chicago and tracking was the subject of my dissertation.

As a former classroom teacher in California, I knew that some schools in my home state were enthusiastically abolishing tracking. Others were digging in their heels to keep multi-leveled systems intact. A

simple question—Why do some schools track while other schools don't?—struck me as an interesting line of inquiry. The question had been ignored in empirical research. Researchers were pre-occupied with studying tracking's educational effects, not its policy origins.

When I began my research, I suspected that tracking policy decisions could be explained by micro-level factors, i.e., the bundle of characteristics that makes each school a unique place. I hypothesized that the political arguments swirling around the issue could shape which schools detrack. Surely educators serving large numbers of poor and minority students—those that critics claim are victimized by tracking—would be more hospitable to the reform. Similarly, the middle school movement had made detracking part of its agenda and detracking part of the grade 5-8 and 6-8 brand names. I also hypothesized that other factors would come into play, including the size of a school, whether the issue was debated publicly before a school board, and whether parents were involved in deciding policy.

The dissertation analyzed survey data from 373 California middle schools and interviews with 175 educators at twenty-nine case study schools. After it was completed in 1992, I conducted two more surveys, one in Massachusetts (1995) to which 134 schools responded and another in California (1994) that elicited 166 responses. I made twelve more case study visits in 1996-97 and conducted eighty-nine additional interviews. All of the research was published as *The Tracking Wars* in 1999.³⁰

In short, I found that schools are influenced by their local settings when responding to state policy recommendations on tracking. They do not blindly follow the state's wishes. Detracking was taking root in schools serving students who, according to tracking's critics, were victimized by the practice—in particular, low-performing, urban schools attended by large numbers of students from disadvantaged backgrounds. Suburban schools, high-achieving schools, and schools serving wealthier families, on the other hand, were more likely to stay with tracked systems or adopt minor modifications.³¹

Methods and Research Questions



The current study updates the earlier research and includes a special focus on high-achieving students. What has happened over the past two decades? Have poorer schools continued to detrack while their wealthier counterparts continue to resist? Are decisions around tracking still made at the local level?

A survey of Massachusetts middle schools was conducted in the 2008-2009 academic year using the *Tracking Wars* questionnaire with a few minor modifications. Principals were asked to report the number of distinct tracks that their school offered in ELA, math, history, and science over the last five years.³² The response rate was 128 out of 295 schools or 43.4 percent. In addition to survey data, the study also collected achievement data for 1995, 2005, and 2008. The key questions guiding the analysis are:

- 1 How has tracking changed since the 1990s?
- 2 Why have some schools detracked while others resist the reform?
- 3 Are there differences in the number of higher-achieving students associated with tracked and untracked schools? In other words, do tracked and untracked schools produce similar percentages of students reaching the “advanced” level on MCAS?

Before turning to Massachusetts survey data, let’s quickly examine tracking data from the National Assessment of Educational Progress (NAEP). These data can show us whether Massachusetts schools are outliers from national trends or typical of what is going on across the country.

NAEP DATA ON TRACKING

The eighth-grade NAEP data on tracking are based on a question in the school administrator survey: “Are eighth-grade students typically assigned to mathematics [or English language arts] classes by ability so that some classes are higher in average ability than others?” The response produces a dichotomous indicator of tracking policy. It distinguishes between strictly untracked schools and those with *any* tracking, but it cannot register policy changes within existing tracked programs (e.g., schools going from three to two tracks or vice versa in a particular subject). The question has been posed to school administrators intermittently over the last fifteen years. In the 1990s, when ELA and math were sometimes assessed by NAEP in different years (they’ve been administered together since 2003), the question accompanied a particular subject. Keep in mind, too, that the NAEP sample is not a perfect match with the current study’s schools, because schools with eighth grades are not necessarily middle schools.³³

Table 1 shows the percentage of tracked schools in 1992-2007. Mathematics appears consistently more tracked than ELA, with about three-quarters of schools employing tracking in math (75 percent in 2007). Tracking in math also appears more stable over time. Tracking in ELA starts at 48 percent in 1992, drops sharply to 31 percent in 1996, and then rebounds to 43 percent in 2007. A possible

Table 1: NAEP National Tracking Data
(Percentage tracked, 8th grade)

| Mathematics | 1992 | 1996 | 2007 |
|------------------------------|-------------|-------------|-------------|
| Schools | 73 | 71 | 75 |
| Black students | 66 | 64 | 66 |
| White students | 75 | 73 | 78 |
| Hispanic students | 63 | 61 | 70 |
| English language arts | 1992 | 1996 | 2003 |
| Schools | 48 | 31 | 43 |
| Black students | 54 | 39 | 51 |
| White students | 46 | 33 | 39 |
| Hispanic students | 49 | 38 | 59 |

Table 2: NAEP Massachusetts Tracking Data
(Percentage tracked, 8th grade)

| Mathematics | 1992 | 1996 | 2007 |
|------------------------------|-------------|-------------|-------------|
| Schools | 88 | 85 | 83 |
| Black students | 62 | 72 | 71 |
| White students | 90 | 88 | 86 |
| Hispanic students | 80 | 63 | 72 |
| English language arts | 1992 | 1996 | 2003 |
| Schools | 60 | 41 | 38 |
| Black students | 41 | 27 | 31 |
| White students | 61 | 44 | 40 |
| Hispanic students | 58 | 35 | 34 |

confounding element in these data could be classes offered for English language learners (ELL). If principals included such classes in calculating their responses to the NAEP question, that might explain the rebound in tracking after 1996, particularly in ELA. From the mid to late 1990s, the nation experienced a sharp increase in immigration and the number of ELL students surged.

Immigration would also affect schools attended by Hispanic students more than schools attended by white or black students. As shown in Table 1, the post-1996 increase in the percentage of Hispanics at schools that track in ELA (from 38 percent to 59 percent) suggests that ELL may be a factor. Nevertheless, the possibility that a substantial amount of “re-tracking” has taken place is a surprising development that must be seriously considered. The tracking reform movement of the 1990s may have run out of steam late in the decade or even have been reversed. All racial/ethnic subgroups experienced more tracking in ELA (2003) and math (2007) than was the case in both subjects in 1996.

Table 2 reports NAEP tracking data for Massachusetts. Consistent with

findings reported in *The Tracking Wars*, both subjects detracked between 1992 and 1996. Massachusetts schools look as if the detracking trend continued from the 1990s until 2007 but at a slower pace. For all three time periods, math is far more amenable to tracking than ELA, and white eighth graders are more likely than black or Hispanic youngsters to attend schools with tracked classes. Like the national data, tracking of Hispanic students in Massachusetts in math shows an increase from 1996 to 2007. Examining school-level data from Massachusetts will allow us to dig deeper into these trends.

Results From the 2009 Survey of Massachusetts Middle Schools

Massachusetts middle schools have experienced significant detracking in the past two decades. The early 1990s were a time of intense reform. Across all academic subjects, the number of tracks in the typical school declined, on average, by about one-half track (or level) from 1991 to 1995 (see Table 3). The pace of detracking subsequently slowed; nevertheless, schools shed an additional one-half track in most academic subjects—a little less than that in mathematics, though tracking changes were minor from 2005 to 2009.

Table 3: Mean Number of Track Levels in 8th Grade

| | 1991 | 1995 | 2005 | 2009 |
|---------|------|------|------|------|
| Math | 2.99 | 2.59 | 2.36 | 2.30 |
| English | 2.38 | 1.86 | 1.39 | 1.33 |
| History | 2.11 | 1.66 | 1.20 | 1.14 |
| Science | 2.16 | 1.72 | 1.23 | 1.18 |

As the 2009 data indicate, math is quite different from other subjects. Here substantial tracking still remains. In 1991, the typical Massachusetts school featured approximately three separate tracks (2.99) in 8th-grade math. By 1995, that average had shrunk to 2.59 levels. In 2009, it had fallen to 2.30. Usually these are an algebra course for high achievers and a pre-algebra or 8th grade math course for students not yet ready for algebra.

In ELA, history, and science, by contrast, the dominant practice in 2009 is to offer only a single, heterogeneously-grouped course to eighth graders. That is a significant change from 1991, when most schools had two or more levels in those subjects.

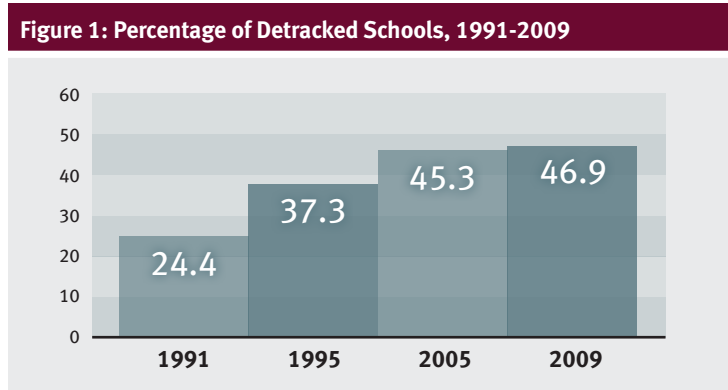
Table 4 provides additional detail on how detracking has unfolded within subjects. The number of schools with only one level in math was static from 1995 to 2009. In other words, today only one in six middle schools in Massachusetts offers a single, heterogeneously-grouped course in math, much as in 1995. The math detracking that has occurred has been primarily schools with three or more levels scaling down to two. Schools with three or more levels in any subject are now dinosaurs. The demise of that convention is most evident in ELA, history, and science. Schools with three or more

Table 4: Percentage of Middle Schools with Tracking in Academic Subjects, 1995 and 2009

| | 1995 | | | 2009 | | |
|-----------------------|---------|----------|------------|---------|----------|------------|
| | 1 level | 2 levels | 3 + levels | 1 level | 2 levels | 3 + levels |
| Math | 15.2 | 30.3 | 54.5 | 15.6 | 49.2 | 35.2 |
| English language arts | 55.1 | 14.9 | 30.0 | 72.7 | 22.7 | 4.7 |
| History | 67.7 | 9.4 | 22.9 | 89.8 | 7.0 | 3.1 |
| Science | 62.2 | 14.2 | 23.6 | 86.7 | 9.4 | 3.9 |

Note: The n-size for schools is 128.

levels in those subjects were already in the minority by 1995. Now they are practically extinct. In ELA, for example, schools with three or more ability levels declined from 30 percent in 1995 to less than 5 percent in 2009.



Detracked classes are the norm for ELA, history, and science instruction in Massachusetts middle schools. Three-fourths of schools or more do not provide classes grouped by ability in those subjects. But math still contains vestiges of tracking. To model a school’s entire tracking policy, we constructed a variable that allows a “fudge factor” for mathematics. Based on 8th-grade groupings, we categorized as “detracked” those

schools with no more than one level in ELA, history, and science and no more than two levels in math. Schools with more than one level in ELA, history, or science or more than two levels in math were classified as “tracked.” True, most of the “detracked” schools in the analysis are only partially so, but the definition offers a simple whole-school metric for tracing a change in policy over the past two decades. Figure 1 illustrates that policy change. The percentage of detracked schools nearly doubled from 1991 to 2009, from 24.4 percent to 46.9 percent.

INFLUENCES ON TRACKING POLICY

How tracked and detracked schools differ on key characteristics tells an interesting story. The following analysis examines the association of demographic and organizational characteristics, as well as the influence of key actors on school tracking policy.

Demographic Characteristics of Tracked and Detracked Schools

Table 5 displays key demographic characteristics of tracked and detracked schools. The proportions of black and white students are similar under the two policies. The percentage of Hispanic students differs in tracked (6.8 percent) and detracked (11.7 percent) schools, but the difference falls just shy of statistical significance ($p = .06$). This pattern is consistent with my earlier studies of tracking reform. The only race/ethnicity statistic with a statistically significant difference then was the percentage of

Asian students. Schools with large Asian populations were resistant to tracking reform. That finding came from California, where Asian students are a much larger proportion of the population than in Massachusetts.³⁴

Table 5: Demographic Characteristics of Schools by Tracking Status

| Student Characteristic | Tracked | Detracked |
|------------------------|---------|-----------|
| Black | 5.7 | 4.7 |
| Hispanic | 6.8 | 11.7 |
| White | 82.1 | 78.7 |
| F/R Lunch | 20.5 | 29.7* |

*F-statistic = 4.74, 1 df, $p < .05$

Note: The n-size for schools is 128.

On the other hand, the percentage of students qualifying for free and reduced-priced lunch, a proxy for poverty, stands out. About 30 percent of students at detracked schools are from poor households, compared

to 21 percent at tracked schools. Indeed, the main finding from Table 5 is that the poverty level of a school is linked to tracking policy even though race is not. The argument that tracking discriminates against disadvantaged students apparently persuades many educators that detracking benefits students from impoverished backgrounds. But that explanation is not ironclad. A counter explanation reverses the direction of causality—that schools attract or repel students based on tracking policies. In fact, Figlio and Page (2000) found some evidence that students from high socioeconomic classes leave schools that detrack.³⁵

Table 6: Percentage of Urban, Suburban, and Rural Schools by Tracking Status

| School Community | Tracked | Detracked |
|------------------|---------|-----------|
| Urban | 35.7 | 64.3 |
| Suburban | 63.6 | 28.6 |
| Rural | 28.6 | 71.4 |

F statistic: 4.77, 2 df, p= 0.01

Note: The n-size for schools is 97.

demographic characteristics discussed so far, location is fixed. Here, the demographic characteristic is definitely driving policy, not the other way around. It appears that urban schools that serve large numbers of poor students are receptive to the claim that tracking harms their clientele; this claim is less relevant for suburban schools. Rural schools are every bit as detracked as urban schools, however, so such reasoning does not explain everything.

Organizational Characteristics of Tracked and Detracked Schools

In the earlier studies, I detected a strong relationship between tracking policy and school size. Large schools tracked while small schools didn't. Interestingly, school size mattered in California but not in Massachusetts. Staffing ratios—the minimum number of students necessary to form a class staffed by a full-time teacher—may have differed between the two states. With a sufficient number of exceptionally high-achieving students, for example, large schools can establish a separate advanced class to accelerate these students. Small schools, on the other hand, may have advanced students but not enough of them to justify a separate class.

Table 7: Percentage of 8th Grade Population by Tracking Status

| Number of 8th-Grade Students | Tracked | Detracked |
|------------------------------|---------|-----------|
| 1-199 | 46.9 | 53.1 |
| 200-299 | 60.0 | 40.0 |
| 300+ | 57.9 | 42.1 |

Note: The n-size for schools is 128.

gregates schools by the grades they serve. As noted above, tracking reform is a central tenet of the middle-school movement. Another change that the middle-school movement advocates is for middle schools to take in more elementary grades, a visible way to make them organizationally distinct from high schools.

Table 6 adds to the demographic profiles of tracked and detracked schools by indicating the neighborhoods in which schools are located. Detracking is largely an urban and rural phenomenon. Approximately two-thirds of their middle schools are detracked. Resistance to tracking reform comes from the suburbs where only one-third of the middle schools are detracked. Unlike the

demographic characteristics discussed so far, location is fixed. Here, the demographic characteristic is definitely driving policy, not the other way around. It appears that urban schools that serve large numbers of poor students are receptive to the claim that tracking harms their clientele; this claim is less relevant for suburban schools. Rural schools are every bit as detracked as urban schools, however, so such reasoning does not explain everything.

Table 7 shows the tracking policies for schools of various sizes in the current study. The number of 8th graders is used to model school size. Here, small schools stand out, but the difference is not statistically significant.

Another organizational feature of schools that affects tracking policy is grade configuration. Table 8 disag-

This reconfiguration has taken place. The junior high schools of yesteryear typically served grades 7-9. With the rise of middle schools in the 1980s and 1990s, ninth grade migrated toward the high school. The 7th- and 8th-grade schools that remain—some still call themselves “junior highs”—are more likely to track than the 5-8 or 6-8-grade schools. As Table 8 shows, schools with an earlier starting grade are associated with detracking, but the p-value of .06 is just shy of statistical significance.

Table 8: School Grade Configuration 2009, Percentage by Tracking Status

| Grades Served | Tracked | Detracked |
|---------------|---------|-----------|
| 5th-8th | 42.9 | 57.1 |
| 6th-8th | 52.0 | 48.0 |
| 7th-8th | 77.8 | 22.2 |

F statistic: 2.87, 1 df, p= 0.06 for 7th-8th versus other two combined

Note: The n-size for schools is 121.

with a single teacher—since the teacher is usually trained in elementary education and instructs her/his pupils in all subjects. But seventh grade offers more classes that are grouped by ability than sixth grade, and eighth grade more than seventh. In the current study, principals reported slightly fewer levels in seventh grade than in eighth for every subject: math (2.01 vs. 2.30), ELA (1.27 vs. 1.33), history (1.13 vs. 1.14), science (1.14 vs. 1.18) (data not tabled).

Tracking also differs among grades within the same school. More specifically, it increases as students age, perhaps as a result of the differences among students also increasing, or perhaps because prerequisite knowledge becomes more important each year. Fifth and sixth grades, for example, are almost always self-contained—i.e., students are heterogeneously grouped within a single class

Actors Who Influence Tracking Policy

The survey data show that authority over tracking policy continues to rest with district or school officials. In fact, 44.7 percent of principals reported that tracking was primarily a school decision, and 55.3 percent reported that it was primarily a district decision. This marks a slight shift away from schools and toward districts since 1995, when the split was 52-48 percent favoring school-based policies (data not tabled).

Table 9: Principals' Ratings of Influential Actors in Tracking Policy, Percentage by Influence Level

| Actor | No Influence | Moderate Influence | Considerable Influence | Great Influence |
|------------------------|--------------|--------------------|------------------------|-----------------|
| State Policy Makers | 54.7 | 25.8 | 12.5 | 7.0 |
| District Policy Makers | 14.1 | 22.7 | 28.1 | 35.2 |
| School Administration | 3.9 | 10.2 | 42.2 | 43.8 |
| Teachers | 6.3 | 23.4 | 40.6 | 29.7 |
| Parents | 22.7 | 46.9 | 21.9 | 8.6 |
| Community Groups | 78.1 | 18.8 | 2.3 | 0.8 |
| Research Literature | 10.2 | 28.1 | 38.3 | 23.4 |
| Students | 44.5 | 33.6 | 15.6 | 6.3 |
| NCLB | 36.7 | 22.7 | 25.0 | 15.6 |

Note: The n-size for schools is 128.

Table 10: Actor Influence at Tracked and Detracked Schools

| Actor | Tracked | Detracked |
|------------------------|---------|-----------|
| State Policy Makers | 1.72 | 1.72 |
| District Policy Makers | 2.85 | 2.83 |
| School Administration | 3.34 | 3.17 |
| Teachers | 3.06 | 2.80 |
| Parents | 2.32 | 1.98* |

*F-statistic = 4.95, 1 df, p < .05

Note: This table is based on a survey scale of 1-4, where respondents were asked to rate actor influence such that 1 was “no influence,” 2 was “moderate influence,” 3 was “considerable influence,” and 4 was “great influence.”

The n-size for schools is 128.

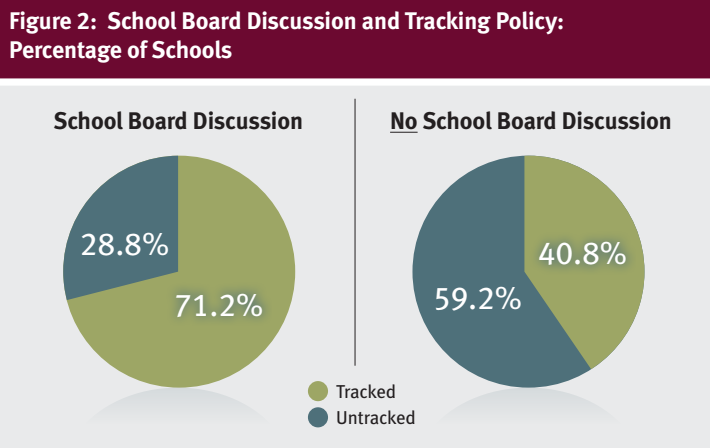
17.9 percent of principals rated district policy makers as having “great influence” on tracking policy. In 2009, the figure is 35.2 percent. Since tracking is controversial, it is more likely to involve district officials and to be resolved at the district level. Urban schools are where district authority is strongest and about 70 percent of the principals in those schools reported that their tracking practices were dictated primarily by district policy (data not tabled).

Parents are a special kind of political actor in school affairs and illustrate an important political lesson. When compared to other potential influences, their power to formulate tracking policy is minimal. When they do exercise influence, however, it appears to make a difference (see Table 10). Parental influence is rated significantly higher, in fact, in schools that are tracked than in those that are not. Teacher influence is also more pronounced at tracked schools, though here the difference falls short of statistical significance ($p = .08$).

Other surveys have reported that parents and teachers favor tracking.³⁶ In the twenty-nine schools that served as my earlier case studies, parents and teachers often forged an alliance to resist district-led tracking reform. I described one school where disgruntled math teachers, who felt they were being forced to detrack by district and state officials, took the case for resistance to the school’s parent council. By a unanimous vote, the council passed a resolution opposing changes to the school’s tracking system.

Fast forward to February 2009, when the *Chicago Tribune* reported that Evanston Township High School initiated mixed-level classes in freshman humanities and senior English language arts. Not surprisingly, the article cites subsequent opposition by the school’s parent-teacher organization (PTO).

In fact, the Evanston PTO took its opposition to the school board, the main forum for channeling opinions on education policy. As shown in Figure 2, school board discussion of tracking



Note: The n-size for schools is 128.

policy matters. An overwhelming number of schools are tracked when school boards have deliberated the topic (71.2 percent). But the percentage falls sharply among schools reporting no school-board discussion of tracking (40.8 percent). When school boards discuss tracking the likely outcome is to keep tracked systems intact (note that the survey question simply asked principals whether tracking was discussed, not the substance of those discussions). Detracking is more likely, then, when policy is decided behind closed doors by district and school administrators.³⁷ (See Sidebar below for discussion of four statistically significant factors associated with tracking policy.)

RELATIONSHIP BETWEEN TRACKING POLICY AND MATH/ELA ACHIEVEMENT

This section examines the relationship between tracking policy and school-level achievement in math and ELA. It would be wonderful to know for sure which policy leads to higher achievement, tracking or detracking. Our present data, however, can gauge the association between policy and achievement—revealing how schools with different policies perform—but not how one affects the other.³⁸

Let's start by examining achievement in ELA. Massachusetts reports the percentage of students scoring at four performance levels on the Massachusetts Comprehensive Assessment System test (MCAS). As displayed in Table 11, there are no significant differences in ELA scores between schools with tracked and detracked ELA programs. The distribution of performance looks similar whether schools offer only one level in ELA or two or more. Under both policies, about 13 percent of students score at the *advanced* level and two-thirds at the *proficient* level. Consistent with the research literature cited above, Massachusetts middle schools with tracked and detracked curricular structures score about the same in eighth-grade ELA.

FOUR SIGNIFICANT INFLUENCES ON TRACKING POLICY

Four factors are statistically associated with tracking policies. They include the grade levels served in a school, the percentage of students on free and reduced-priced lunch, the involvement of parents in tracking discussions, and the involvement of school boards in tracking discussions. (See Appendix for more detail and regression results.) These factors tell a concise (although admittedly speculative) story of why some schools track and other schools do not.

First, we see that the argument for detracking in pursuit of greater equity is heard sympathetically by schools serving predominantly poor populations (i.e., those with higher rates of students receiving free and reduced-priced lunches). They have embraced tracking reform. Schools serving students from higher socioeconomic backgrounds, on the other hand, have resisted detracking. In short, the political argument for eliminating tracking falls on deaf ears in wealthier schools.

Configuration of grades within a school also matters. Mimicking high schools, middle schools serving grades 7 and 8 are more likely to embrace tracking when compared to their grade 5–8 and 6–8 counterparts. The latter are more elementary-oriented in staffing and curriculum and thus may be more hospitable to heterogeneously grouping students for instruction.

Schools in which parents wield influence tend to keep tracking in place as do schools that have had their local school board discuss the topic. As mentioned earlier, detracking typically occurs outside of the spotlight. However, when it is met with resistance, it's often in the form of organized parental opposition. Parents of high achieving youngsters are particularly vocal in resisting efforts to abolish honors or advanced classes. Such confrontations may take place in the open forum of a school board meeting.

Table 11: Distribution of Achievement in 8th Grade ELA

| Number of ELA tracks | % Advanced | % Proficient | % Needs Improvement | % Failing |
|----------------------|------------|--------------|---------------------|-----------|
| 1 (n=91) | 13.2 | 65.5 | 16.0 | 5.2 |
| 2+ (n=35) | 13.1 | 67.7 | 14.9 | 4.4 |

Math is a very different story (see Table 12). The difference is especially pronounced at both tails of the distribution. Schools with more track levels have significantly more pupils performing at the *advanced* and *proficient* levels and fewer students at the *needs improvement* and *failing* levels. Schools with two tracks in math have more advanced kids than schools with only one track (18.6 percent versus 15.8 percent). And schools with three tracks have the most advanced students (26.6 percent). A mirror image of this trend appears for the number of failing students, with detracked schools having more failing students (26.2 percent) than schools with two tracks (20.7 percent) or three tracks (14.8 percent).

Table 12: Distribution of Achievement in 8th Grade Math

| Number of math tracks | % Advanced | % Proficient | % Needs Improvement | % Failing |
|-----------------------|------------|--------------|---------------------|-----------|
| 1 (n=17) | 15.8 | 29.3 | 28.8 | 26.2 |
| 2 (n=66) | 18.6 | 31.9 | 28.7 | 20.7 |
| 3+ (n=43) | 26.6** | 34.5* | 24.1** | 14.8** |

*p < .05 ** p < .01

TRACKING AND HIGH ACHIEVERS

Now let's dig deeper into the relationship between tracking and the number of high achievers in mathematics. Since math is the only tracked subject left in eighth grade, it's vulnerable to future detracking efforts. Moreover, the contemporary push for universal eighth-grade algebra, which would effectively reduce eighth-grade math to a single track in every school, is hitting its stride. The presumptive benefit of such a policy is an increase in the number of high-achieving math students.

The next three tables (Tables 13-15) assess the short-, mid-, and long-term associations between school tracking policy and the percentage of math students scoring at advanced levels. For all three time intervals, the association is modeled in ordinary least squares regressions predicting the 2008 percentage of advanced students in eighth-grade math.³⁹

Table 13 shows the short-term association. The coefficient for math tracks (6.05) indicates that each additional track level in a school's mathematics program—up to a maximum of three levels—is associated with about a 6 percentage-point increase in the number of advanced students. That is a large gain. (Approximately 19 percent of eighth graders in Massachusetts score at this level.) Tracking is more prevalent in suburban, high-SES schools, however, so adding a variable to control for poverty (see the column "+ Lunch") reduces the coefficient for mathematics tracks to about 3 percentage points (2.98). In other words, half of the 6 percentage-point gain associated with an additional track

Table 13: Short-Term Association: Predicting Percent Advanced 2008 Using 2008 Tracking Policy

| | No Controls n=126 | + Lunch n=125 |
|---------------------------|----------------------|---------------------|
| Math Tracks (2009) | 6.05*** (1.43) | 2.98** (1.26) |
| Lunch | -- | -25.72*** (3.46) |
| Constant | 7.58** (3.29) | 20.63*** (3.24) |
| <i>Summary statistics</i> | | |
| F statistic | 17.97 | 39.72 |
| Degrees of freedom | 1 | 2 |
| P-value | <.001 | <.001 |
| Adjusted R-square | 0.12 | 0.38 |

*p<.10 ** p<.05 ***p<.01

Table 14: Mid-Term Association: Predicting Percent Advanced 2008 Using 2005 Tracking Policy

| | No Controls n=126 | +2005 Ach. n=121 | + Lunch n=120 |
|---------------------------|----------------------|---------------------|---------------------|
| Math Tracks (2005) | 4.61*** (1.38) | 0.91 (0.90) | 0.27 (0.87) |
| 2005 Percent Advanced | -- | 0.88*** (0.06) | 0.74 (0.07)*** |
| Lunch | -- | -- | -11.82*** (3.06) |
| Constant | 10.58*** (3.25) | 6.26*** (2.01) | 12.46*** (2.53) |
| <i>Summary statistics</i> | | | |
| F statistic | 11.15 | 120.99 | 93.96 |
| Degrees of freedom | 1 | 2 | 3 |
| P-value | .001 | <.001 | <.001 |
| Adjusted R- square | 0.08 | 0.67 | 0.70 |

*p<.10 ** p<.05 ***p<.01

is due to schools in poor neighborhoods detracking and schools in wealthier neighborhoods continuing to track. But even with a school's socioeconomic status held constant, a 3 percentage-point increase in the number of high achievers remains associated with each additional track. That is *still* large. It means that a school with 200 eighth graders that offers at least three or more levels of math is typically attended by twelve more students scoring at the advanced level than a detracked school with 200 eighth graders of similar socioeconomic status.

Calculations for the mid-term association are displayed in Table 14. Each additional math track that schools employed in 2005 is associated with a 4.6 percentage-point increase in the number of advanced math students three years later (2008 test scores are predicted). Adding the percentage of advanced students in 2005 as a control variable (second column, "+2005 ach.") eliminates most of the tracking effect (about 1 percentage point remains, 0.91). By adding this control, the equation now predicts the change in the percentage of advanced students from 2005 to 2008. Adding a control for school SES further reduces the coefficient for math tracks to a trivial figure (0.27). The results suggest that once the association of tracking policy and achievement is "baked into the cake," it does not change much over time. The results also confirm the remarkable stability of school test scores. (A powerful predictor of a school's percentage of advanced students in 2008 is the school's percentage of advanced students in 2005.)

Is there an association between tracking policy in 1995 and a school's percentage of advanced students thirteen years later?⁴⁰ As shown in Table 15, each additional track in math in 1995 is associated with

Table 15: Long-Term Association: Predicting Percent Advanced 2008 Using 1995 Tracking Policy

| | No Controls n=105 | +1995 Ach. n=105 | + Lunch n= 105 |
|-----------------------------------|----------------------|---------------------|--------------------|
| Math Tracks (1995) | 6.36*** (1.86) | 1.11 (1.02) | 0.93 (1.00) |
| Percent at top 2 levels (1995) | -- | 1.14*** (0.07) | 0.95*** (0.10) |
| Lunch | -- | -- | -10.90** (4.41) |
| Constant | 5.01 (4.79) | -4.93 (2.56)* | 1.99 (3.75) |
| <i>Summary statistics</i> | | | |
| F statistic | 11.65 | 161.55 | 115.12 |
| Degrees of free- dom | 1 | 2 | 3 |
| P-value | <.001 | <.001 | <.001 |
| Adjusted R-square | 0.09 | 0.76 | 0.77 |

*p<.10 ** p<.05 ***p<.01

detracked schools do not differ significantly. However, tracked schools have significantly higher test scores in math than detracked schools. In particular, they report more students performing at the advanced level in mathematics. In regression models that control for school SES and measure a contemporaneous relationship in 2008, each additional math track is associated with a 3 percentage-point increase in the number of advanced students. The reverse is also true. A reduction of a single math track is associated with a 3-percent decrease in advanced math students. Regression models of mid-term and long-term associations do not detect such a relationship once controls for initial achievement and SES are included. In other words, the number of students performing at advanced levels remains fairly constant over time.

a 6 percentage-point increase in the number of advanced students in 2008 (6.36), but controlling for the level of achievement in 1995 and school SES explain all but about one point of the gain (0.93). Again, like the mid-term analysis presented in Table 14, the large initial achievement differences between tracked and detracked schools explain most of the difference in achievement in 2008. It is astonishing how consistent school test scores are over a thirteen-year period. The percentage of advanced students attending a school in 1995 is an excellent predictor of how many advanced students attend the school in 2008 (correlation coefficient of 0.87).

To summarize the results of this section, the analysis offers qualified support for the hypothesis that tracking policy is associated with achievement. The ELA scores of tracked and

Summary and Discussion



his study investigated tracking in Massachusetts middle schools. Drawing on survey and achievement data from 2008-2009 as well as the 1990s, it offers four key findings:

- 1. Tremendous change has occurred.** The middle-school student of 1991 attended schools offering at least two distinct tracks in ELA history; science and math classes were then organized into three tracks. Today's middle-school student encounters far less tracking. ELA, history, and science are essentially detracked. In these subjects, schools typically offer a single course serving students who are heterogeneous in ability. Mathematics now features two tracks in eighth grade: algebra and a course for students not yet ready for algebra. Nearly twenty years ago, middle-school students attended tracked classes for most of the day. Now they spend most of their day in detracked classes.
- 2. Several factors are associated with different tracking policies.** Why do some schools continue to track while others have detracked and employ heterogeneous grouping? The survey data reject the idea that tracking is spawned by capitalism or any other single, overarching source. Study results also reject the dramatic narrative of the advocacy literature: that tracking policy pits good versus evil, egalitarianism versus elitism, tolerance versus racism (see Sidebar on page 27). But other distinctions emerge. In fact, the study documents sharp policy differences between math and ELA departments; between grade 7-8 schools and their more elementary-oriented counterparts; between high- and low-SES schools; between schools in which parents are more influential in developing tracking policy and those in which parents are less influential; and between schools whose tracking policies were discussed by school boards and those whose policies were deliberated less publicly.
- 3. Detracking is more prevalent among urban, high poverty schools.** Demographic characteristics are indeed associated with tracking policy, but in a way that tracking reformers would surely applaud. Low-SES urban schools are more likely to detrack than are higher-SES suburban schools. The argument that tracking harms disadvantaged students is being heard and embraced by educators at schools serving students from lower-SES families. That means that the risks associated with detracking are concentrated in schools serving large numbers of low-achieving (and low-income) children.
- 4. Detracking carries risks for high-achieving students.** The present study joins the studies of Reese et al. and Allensworth et al. in cautioning that detracking may adversely affect high-achieving students, but only in mathematics. With school-level SES held constant, each additional track in eighth-grade math (up to three) is associated with a 3 percentage-point increase in the number of students scoring at the advanced level. That means the advantage for a school offering three tracks instead of one is about six percentage points. The data do not allow us to test whether tracking is the direct cause of this difference. Although the effect is stable over time, one cannot determine whether selection is taking place that would skew the results—for instance, whether parents are selecting schools based on tracking policy or educators are selecting policies based on their schools' achievement.

The study also raises cautionary flags concerning the contemporary push to detrack mathematics by offering algebra to all eighth graders. Schools with a single math course have fewer advanced students—the very students who are best prepared for algebra in eighth grade. These are also the students who typically serve as academic role models in heterogeneously grouped classes.


DOES PARENT INFLUENCE AND SCHOOL BOARD ACTIVITY SUGGEST THAT TRACKING SERVES THE PRIVILEGED AND IS DEFENDED BY THEM?

The survey data may indeed reflect that parents in higher socioeconomic classes use their influence to maintain tracking. Certainly one way of exercising influence is by lobbying school boards. We ran several analyses to explore the question. Particular ratings of parent influence on tracking are not associated with school SES. In other words, similar parent ratings are reported by high- and low-SES schools. But when schools are divided into SES quartiles, high-SES schools tilt towards tracking when parents are rated as influential. In fact, in low-SES schools, parent influence seems to favor neither tracking nor detracking. Also among low-SES schools, whether school boards discussed tracking policy made little difference in how schools approached tracking. Among high-SES schools, school board discussion favors tracking over detracking.⁴¹

This pattern is consistent with the argument that privileged parents use their resources to defend tracking. Parent influence and the impact of school board discussion on tracking policy are greater in wealthier communities. No doubt these are reasons why tracking is more apt to persist in high-SES schools. Despite this, the charge that these activities are motivated by a desire to maintain privilege remains doubtful. The social-reproductionist argument hinges on tracking producing bad results for low-track kids. But math achievement is higher in tracked schools and ELA achievement is about the same. In addition, schools are not well integrated by socioeconomic status. That means the low tracks in high-SES schools are almost certainly populated with high-SES kids. Why aren't their powerful parents rising up to demand detracking? It could be that they see remedial courses as valuable.

The main reason why parents are suspicious of heterogeneous grouping may not be sinister. Parents have particular educational interests; for example, they desire a curriculum that closely matches their child's needs and interests. Since the latter are somewhat unique to children, this parental impulse naturally transcends race and class. The Public Agenda Foundation asked parents, for instance, whether racism and privilege lay at the heart of parental opposition to detracking. The authors concluded that "Some proponents of heterogeneous grouping, professional educators and others, have suggested that parental opposition to it is a camouflage for racial prejudice—the fear of white parents that their children will be put in classes with 'underachieving' African-American students, but opposition to heterogeneous grouping is as strong among African-American parents as among white parents, and support for it is generally weak."⁴²

Conclusion

 This report supplements earlier studies on high-achieving students by the Fordham Institute. *High-Achieving Students in the Era of No Child Left Behind* (2008) reported teachers' concern that, as they scramble to help struggling students gain proficiency in academic subjects, the needs of gifted/high-achieving youngsters are neglected. A companion report in the same study used national data for black, Hispanic, and poor eighth graders scoring at the 90th percentile on the National Assessment of Educational Progress (NAEP) and showed that these high achievers do not have access to the same educational resources—provided by homes, schools, or teachers—as white and higher-SES high achieving students. It also found that high achievers who were black, Hispanic, or poor were more likely to attend detracked schools.

The present study helps illumine how that happened. Based on good intentions but scant empirical evidence, detracking unfolded slowly over many years. Math is now the only subject with a significant amount of tracking in middle schools. Whether detracking eventually helps the youngsters it is primarily intended to help—or looms as another barrier to their success—remains a question that detracking's advocates, despite success in the policy arena, have failed to answer.

Endnotes

1. Burns, C.C. and Kevin Welner. 2006. Alternative Approaches to the Politics of Detracking. *Theory into Practice*, 45(1): 90-99.
2. Tomlinson, Carol Ann. 1999. *The Differentiated Classroom: Responding to the Needs of All Learners*. Alexandria, VA: Association for Supervision and Curriculum Development. 2.
3. Penny, Anna. 2009. Get kids on track: Stop packing quicker, slower students in the same class. *New York Daily News*. Editorial. August 30.
4. National Center for Education Statistics. 2009. *Digest of Education Statistics 2008*. NCES 2007017, Table 34.
5. The words “track” and “level” are used interchangeably throughout this report.
6. Rosenbaum, James. 1976. *Making Inequality: the Hidden Curriculum of High School Tracking*. New York: John Wiley & Sons. For a history of tracking see Loveless, Tom. 1998. *The Tracking and Ability Grouping Debate*. Washington, D.C.: Thomas B. Fordham Institute. Lucas, Samuel. 1999. *Tracking inequality: Stratification and mobility in American high schools*. New York: Teachers College Press.
7. Oakes, Jeannie. 2005. *Keeping Track: How Schools Structure Inequality, Second Edition*. New Haven, CT: Yale University Press. 202.
8. See Wheelock, Anne. 1993. *Crossing the Tracks: How “Untracking” Can Save America’s Schools*. New York: The New Press. xi.
9. For a critique of the middle school movement, see Yecke, Cheri Pierson. 2005. *Mayhem in the Middle*. Washington, D.C.: Thomas B. Fordham Foundation.
10. The cartoon appears in *Boston Globe*, November 19, 1993, pg. 19.
11. Carnegie Corporation. 1989. *Turning Points: Preparing American Youth for the 21st Century. The Report of the Task Force on Education of Young Adolescents*. New York: Carnegie Corporation. California Department of Education. 1987. *Caught in the Middle: Educational Reform for Young Adolescents in California Public Schools*. California Department of Education. Massachusetts Department of Education. 1993. *Magic in The Middle: A Focus on Massachusetts Middle Grade Schools*. Malden, MA: Massachusetts Department of Education. For description of parental response see Oakes, Jeannie, Amy Stuart Wells et al. 1996., *Beyond the Technicalities of School Reform: Policy Lessons from Detracking Schools*. Los Angeles: UCLA Graduate School of Education & Information. and Welner, Kevin. 2002. *Legal Rights, Local Wrongs: When Community Control Collides with Educational Equity*. Albany: State University of New York Press.
12. The NAGC policy is quoted in Oakes (2005) note 13, pg. 316.
13. Success for All, founded by education psychologist Robert Slavin, is a noted literacy program that employs ability grouping in ELA instruction. Students are assessed frequently and re-grouped to fine tune the match between evolving skills and reading materials.
14. For a study critical of ability grouping, see Condron, Dennis J. 2008. An Early Start: Skill Grouping and Unequal Reading Gains in the Elementary Years. *The Sociological Quarterly* 49:363-94. The study documents a gain for students placed in high groups and a loss for placement in low groups, net of other variables.
15. The A and B designations refer to semesters.
16. Page, Reba. 1991. *Lower-track classrooms: A curricular and cultural perspective*. New York: Teachers College Press.
17. Loveless, Tom. 2009. *The 2008 Brown Center Report on American Education: How Well are Students Learning?* Washington, D.C.: Brookings Institution Press.
18. Slavin, Robert E. 1990. Achievement Effects of Ability Grouping in Secondary Schools: A Best Evidence Synthesis. *Review of Educational Research*, 60(3): 73-77. Kulik, Chen-Lin and James Kulik. 1982. Effects of Ability Grouping on Secondary School Students: A Meta-Analysis of Evaluation Findings. *American Educational Research Journal* 19(3): 415-428. Ferguson, Ronald F. 1998. Evidence that Schools Can Narrow the Black-White Test Score Gap. In *The Black-White Test Score Gap*, ed. Christopher Jencks and Meredith Phillips. Washington, D.C.: Brookings Institution Press.

- Loveless, Tom. 1998. *The Tracking and Ability Grouping Debate*. Washington, D.C.: Thomas B. Fordham Institute.
19. Mosteller, Frederick, Richard Light, and Jason Sachs. 1996. Sustained inquiry in education: Lessons from skill grouping and class size. *Harvard Education Review* 66(4): 797–842.
20. Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2008. Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. NBER Working Paper No. 14475.
21. Recall that a standard deviation is a measure of how much test scores tend to deviate from the mean.
22. Gamoran, Adam and Sean Kelly. 2003. Tracking, Instruction, and Unequal Literacy in Secondary School English. In *Stability and Change in American Education: Structure, Processes and Outcomes*, ed. Maureen. T. Hallinan, Adam Gamoran, Warren Kubitschek, and Tom Loveless, 109-126. Clinton Corners, NY: Eliot Werner Publications.
23. The studies of Joplin Plan grouping, with elementary students grouped for instruction based on a single subject, yield more positive results than XYZ studies. See Slavin, Robert. 1987. Ability Grouping and Student Achievement in Elementary Schools: A Best-Evidence Synthesis. *Review of Educational Research* 57(3): 293-336. And Slavin, Robert. 1996. *Education for All*. New York: Taylor & Francis.
24. Rees, Daniel I., Laura M. Argys, and Dominic J. Brewer. 1996. Tracking in the United States: Descriptive Statistics from NELS. *Economics of Education Review* 15(1): 83-89.
25. Kulik, James A. and Chen-Lin C. Kulik. 1992. Meta-Analytic Findings on Grouping Programs. *Gifted Child Quarterly* 36(2) Spring: 73-77.
26. Hanushek, Eric and Ludger Woessman. 2005. Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries. NBER Working Paper 11124.
27. Assume for a moment that all forms of grouping students by ability produce similar outcomes. Then the treatment of the United States in the Hanushek and Woessmann study raises more questions. It is coded as a late tracker. But in the U.S. within-class ability grouping dominates ELA instruction in the primary grades. By the 4th grade, the American students tested in the Progress in International Reading Literacy (PIRLS), results of which are used as the baseline score for ELA, have been ability grouped for ELA instruction more than most students in the world. American students go on to attend tracked math classes by the end of middle school grade. These experiences are unacknowledged in coding the U.S.; in other words, America is treated as if no ability grouping occurs before age fifteen. New Zealand is another country in which ability grouping for ELA instruction in the elementary grades is a common practice, and yet it too is coded as a late tracker. Such coding decisions may skew the study's finding that early tracking nations produce greater inequality. The U.S. and New Zealand register among the lowest five nations on the study's key inequality indicator, which is the change in standard deviations in test scores from the elementary to high school years. In addition, Eastern European countries, which track early and often, drive much of the finding of inequality in the Hanushek and Woessmann analysis. Eastern European countries score at higher levels on PIRLS and TIMSS than on PISA, probably because their traditionalist curricula are a better match with the content-orientation of PIRLS and TIMSS as opposed to the socioconstructivist orientation of PISA.
28. See Rees, Argys, and Brewer (1996) above.
29. Allensworth, Elaine et al. 2008. *College Preparatory Curriculum for All in Chicago High Schools: Consequences of Ninth-Grade Course Taking on Academic Outcomes*. Chicago: University of Chicago, Consortium on Chicago School Research.
30. Loveless, Tom. 1999. *The Tracking Wars: State Reform Meets School Policy*. Washington, D.C.: The Brookings Institution Press.
31. *The Tracking Wars* also found that schools serving grades 5-8 or 6-8 were more likely to detrack than schools serving grades 7-8. Tracking was more prevalent in schools in which parents exercised influence over policy than in schools without parent influence. And in communities where the tracking issue was debated by the local school board, schools were more likely to reject tracking reform.
32. A web-based survey was also conducted in California but the response rate (about 20 percent) fell well below acceptable levels, so the data were not analyzed.

33. Another difference between the NAEP data and the data collected in the current study is the manner in which principals were asked to report tracks. In the latter (and in *The Tracking Wars*), principals were asked to exclude ELL and special-education classes from the count of separate tracks. However, the NAEP instructions to principals do not include similar guidance. Classes that serve ELL students and special education courses stand outside the tracking debate. Schools offer such programs in response to federal policy and sometimes must create separate courses at the behest of court directives. And although it may seem to be a logical contradiction, tracking's critics are careful to couch their critique of curriculum differentiation in terms that apply only to proficiency differences in subject matter—whether advanced and struggling readers or math students should be taught in separate classes—rather than differences stemming from cognitive difficulties or native language.

34. Lucas and Gamoran found an Asian benefit to tracking. Lucas, Samuel and Adam Gamoran. 2002. Track Assignment and the Black-White Test Score Gap: Divergent and Convergent Evidence from 1980 and 1990 Sophomores. In *Closing the Gap: Promising Strategies for Reducing the Achievement Gap*, ed. Tom Loveless, 171-198. Washington, D.C.: Brookings Institution Press.

35. Figlio, David and Marianne Page. 2002. School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality? *Journal of Urban Economics* 51(3), 497-514.

36. Twohey, Megan. 2009. Narrowing the Gap: Class mergers draw fire as school fights racial divide. *Chicago Tribune*, February 9.

37. Recently parents at a Hughesville, PA school board meeting, for example, “complained of not being aware that ‘Hughesville does not have junior high honors.’”

38. Note that most of the research on tracking and achievement reviewed in this report uses students as the unit of analysis. The current study uses schools as the unit of analysis. When examining test scores from a single year, there is no analytical difference—a school's test score and the test score of the average student attending that school are the same. The analytical difference emerges when examining trends in achievement over time. Changes in a particular school's tests scores are differences in test scores between different groups of students. The students attended the same school but different cohorts in different years. The test score changes that occur are institutional.

39. In the short-term model, two explanatory variables are employed: 2008 tracking policy and a variable to control for socioeconomic status (the percentage of students qualifying for free and reduced-priced lunch). The mid-term association is modeled with the same equation except that a 2005 tracking policy variable is substituted, in effect, asking whether the 2005 tracking policy helps to predict the percentage of advanced students three years later—in 2008. And the long-term model uses a 1995 tracking policy variable to assess whether it contributes to predicting the percent of advanced students thirteen years later—in 2008. Controls for prior achievement are also included in the mid- and long-term models.

40. Evaluating the long-term association presented technical challenges that required a more conservative approach to estimate effects. Of the schools that responded to the 1995 survey of tracking practices, 105 schools had 2008 test score data. The Bay State gave a different test in 1995, called the Massachusetts Educational Assessment Program (MEAP), which reported student achievement at five performance levels instead of the four levels reported in the 2005-2008 data for MCAS. When the percentage of students performing at the highest level in 1995 is used as a control variable, one finds a positive effect for tracking that approaches statistical significance ($p = .11$). This could be a spurious result as only about 5.5 percent of students scored at the top level in 1995. The top two levels combined comprise about 18 percent of students, comparable to the percentage of advanced students in 2008. This more conservative figure is used in the regression results reported in Table 16.

41. The finding on parent influences diverges from the 1990s findings. Loveless, Tom. 1995. Parents, Professionals, and the Politics of Tracking Policy. *Faculty Research Working Paper Series R95-19*. Cambridge, MA: Harvard University. And Loveless (1999), 79-84.

42. Jean Johnson and John Immerwahr. 1994. *First Things First: What Americans Expect from Public Schools*. New York: Public Agenda Foundation. 18-19.

Appendix

A MULTIVARIATE MODEL OF INFLUENCE ON TRACKING POLICY

Table A-1 provides results from a statistical equation that models the influence of various factors on tracking policy. The data are analyzed using logistic regression. Tracking policy is the dependent variable, with school demographic and achievement characteristics, school organizational characteristics, and the influence of actors used to predict tracking policy. Positive coefficients indicate support for detracking; negative coefficients indicate support for tracking. The full model (first column) provides an estimate of each factor's influence on tracking policy holding all the other variables in the model constant. Since many of the variables overlap and are highly correlated, the best fitting model provides a more parsimonious (and superior) explanation for what drives tracking policy.

| Table A-1: Logit Models of Influence on Tracking Policy | | |
|--|------------------------------|----------------------------|
| Factors | Full Model n= 125 | Best Fit n= 127 |
| <i>Demographic and Achievement</i> | | |
| Urban | 0.38 (0.95) | — |
| Suburban | -0.34 (0.69) | — |
| F/R Lunch | 1.36** (0.68) | 0.35* (0.21) |
| Percent nonwhite | -1.23** (0.55) | — |
| Math score | 0.01 (0.02) | — |
| <i>Organizational Characteristics</i> | | |
| Eighth grade population | -0.002 (0.003) | — |
| Grade levels | -0.86 (0.70) | -1.17* (0.62) |

Continued

Four factors exercise a statistically significant influence on tracking policy (see the best fit model in the second column). They include the grade levels served in a school, the percentage of students on free and reduced-priced lunch, the involvement of parents in tracking discussions, and the involvement of school boards in tracking discussions.

Table A-1: Logit Models of Influence on Tracking Policy

| <i>Actors</i> | | |
|-----------------------------|--------------------|--------------------|
| State | -0.12 (0.27) | — |
| District | 0.15 (0.23) | — |
| Principal | -0.42 (0.37) | — |
| Teachers | 0.05 (0.35) | — |
| Parents | -0.27 (0.29) | -0.40* (0.23) |
| Community Groups | -0.92* (0.56) | — |
| School Board Discussion | -1.29*** (0.46) | -1.07*** (0.40) |
| Constant | 3.06* (1.80) | 1.33** (0.54) |
| <i>Summary statistics</i> | | |
| Chi-square | 34.45 | 22.42 |
| Degrees of freedom | 14 | 4 |
| P-value | =.0018 | <.001 |
| Percent Correctly Predicted | 72% | 67% |

* $p < .10$ ** $p < .05$ *** $p < .01$

