Tracking in Middle School

A Surprising Ally in Pursuit of Equity?

Prepared for the Thomas B. Fordham Institute's Education for Upward Mobility Conference, December 2, 2014. Draft and not for citation without author's permission.

Tom Loveless

Introduction

Education can always use an upbeat story. The past decade's expansion of the Advanced Placement (AP) program is surely one of those. The graduating class of 2013 included, for the first time ever, more than one million AP examinees. Participation doubled in only ten years. Among low-income kids, the expansion has been amazing. In 2003, 58,489 low-income students sat for AP tests. In 2013, the number was 275,864—a 370 percent increase. AP started as an elite high school curricular option, designed to serve the nation's top students, and it was democratizing. Students from privileged backgrounds had always been well represented in AP classes. Now academically talented students from poor families were joining them.[1]

But concerns about the success of the expansion have emerged. It's one thing to take an AP class; it's quite another to learn the material and to do well on AP exams. A 2013 story in the *Baltimore Sun* described the plight of thousands of low-income students who take AP classes, receive As and Bs in the courses—and then fail AP exams. They don't fail by just a little bit. They fail by a lot, receiving a score of 1 or 2 when at least a 3 is required by some colleges to receive credit. Not only are these low-income kids denied college credit for their AP courses, but they also show up at college, the *Sun* reports, "with skills so low that they must take remedial classes." The *Sun's* summary of the local situation in Baltimore is stunning: "In at least 19 high schools throughout the Baltimore region, more than half of the students earning an A or B in an AP class failed the exam." The nineteen schools all serve low-income neighborhoods.[2]

Experts and AP officials pinpointed the same culprit: students' lack of preparation for AP classes. Too many low-income students arrive in high school with sterling academic records from middle school, but in reality are light years away from knowing what they need to know to be successful in an AP class.

This Paper's Argument

This paper argues something unthinkable to a large number of analysts who consider themselves equity-minded: that tracking—the assignment of students to different classes on the basis of ability or achievement—may be a means of better preparing disadvantaged students for AP classes. This will require a mind-shift from policies emphasizing equal access to advanced courses for all or most youngsters to policies emphasizing talent development for high-achieving students. Middle schools prepare students for high school. By adequately preparing more impoverished middle school students for the academic demands of AP, tracking can serve as a tool for greater fairness.

The paper's objective is to convince readers that this is a plausible hypothesis—and to do so empirically. It will not produce causal evidence that will prove or disprove the assertion. Such evidence does not currently exist.

Evidence does exist, however, that supports three propositions. When logically connected, the three propositions bolster the likelihood of the hypothesis being true.

1. Poor, Hispanic, and black middle school students are less likely to be enrolled in tracked classes than students who are socioeconomically better off, white, and Asian.
2. Middle schools serving predominantly disadvantaged students are less likely to offer tracked classes than schools serving advantaged populations.
3. Research on tracking is mixed, but studies focusing on its distributional properties—that is, how tracking differentially affects different kinds of students—generally show a positive effect for high-achieving students. That is particularly true for classes that group academically talented students together and offer an enriched or accelerated curriculum.

The upshot of these propositions is that high-achieving eighth graders in socioeconomically disadvantaged communities are denied an opportunity that high-achieving peers in advantaged communities enjoy. Kids from middle- or upper-middle-class families are more likely to attend schools with tracked, high-achieving classes that prepare them for AP courses in high schools. They are more likely to have access to middle school classes that challenge them and allow them to excel. That's not fair.

The Current Study

Following this introduction, the second section of this paper describes the academic achievement of students in poverty. A promising program that boosts the scores of high-achieving, disadvantaged youngsters through tracking also receives attention. The term "tracking" refers to assigning students to different curricula, classrooms, and teachers based on ability or prior achievement. Why isn't tracking used by more schools serving disadvantaged students? Section three answers that question by briefly describing the recent history of tracking reform, offering clues as to how we got to where we are now. The fourth section digs into the latest NAEP data to describe the demographic characteristics of students and schools who are tracked or untracked. The fifth section looks at the research on tracking, concluding that although mixed in determining tracking's overall effect, the preponderance of evidence indicates that tracked classes are beneficial to high achievers. The sixth section concludes by offering several design characteristics that a tracking-for-equity program in mathematics should feature, operating in concert with the Common Core State Standards in math.

Academic Achievement and Students in Poverty

As Sean Reardon has documented, fifty years ago the test score gap between whites and blacks was much larger than the gap between rich and poor.[3] Today that situation has reversed. Children who were born into poor families in the 1950s, 1960s, and 1970s scored about 0.9 standard deviations lower on achievement tests than children from wealthy families. The gap has expanded to 1.4 standard deviations among children born in the 1990s. Poverty has a depressing saliency to achievement that race had prior to the civil rights movement in the mid-twentieth century.

We know much more about race gaps in achievement than about gaps related to income. In a series of fascinating studies, Eric A. Hanushek and Steven G. Rivkin analyzed Texas data

that documented students' achievement trajectories as they progressed through school. The black-white test score gap ballooned among higher-achieving students, growing much more than among students who were of average or below-average achievement. Although the pattern is evident in test-score changes from third to fifth grade, it is even more pronounced in fifth to eighth grade changes. The researchers divided students into sixteen groups based on the students' scores as kindergartners. As the students moved through school grades, the black-white achievement gap changed. From fifth to eighth grade, the gap actually contracted in all but one of the seven bottom-achievement groups, expanded by at least 0.05 standard deviations in all but one of the next six groups, and expanded by at least 0.09 in the top three groups. For the entire third-to-eighth-grade span, the achievement gap increased two and a half times more in the top three achievement groups than in the bottom seven (0.25 vs. 0.10 standard deviations).[4]

Hanushek and Rivkin's analysis applied to high-achieving black students, not to high-achieving, low-income students. Nevertheless, it is not a stretch to believe that some of the reasons put forth for race gaps—negative peer influences, attending segregated and under-resourced schools, exposure to inexperienced or ineffective teachers, parents who aren't savvy about getting the most out of schools—also apply to the achievement gap between wealthy and poor students.

David Card and Laura Giuliano analyzed data from special classes serving gifted elementary youngsters in a large eastern U.S. school district. In 2004 the district mandated that its elementary schools set up separate classrooms for gifted students in fourth and fifth grades. Pursuant to state law, the district used an IQ test score of 130 as criterion for being identified as gifted, with retesting allowed (including by a psychologist hired by the parents) if students scored at least 127. The following year, concerns about underrepresentation in the program by low-income and minority students, particularly English language learners, led to universal screening using a non-verbal ability test and more flexible IQ cutoff (these were called "Plan B" gifted students).[5]

The district also adopted an innovative policy to promote gifted classes in schools with large underrepresented populations. Every school in the district was required to provide a special gifted class if it had even a single gifted student. Classes were to be filled out with high achievers—students who, although not certified "gifted," had scored the highest in the school on the previous year's state assessment. Because classes are staffed at twenty to twenty-four pupils, that requirement forced schools serving predominantly low-income students to place their highest-achieving poor kids into gifted classes. The policy had its intended effect. Disadvantaged high achievers represent about 40 percent of students in the district's gifted classes. In very poor schools, they make up the vast majority of participants.[6]

Card and Giuliano exploited this unique policy to evaluate whether separate gifted tracks benefit gifted students as traditionally identified, Plan B students, or high achievers. They found no effect on the reading or math achievement of gifted students, regardless of how giftedness is identified. But they uncovered "positive and relatively large effects on the achievement of the non-gifted high achievers who fill the remaining seats in the class, concentrated among free/reduced price lunch participants and black and Hispanic

students." The authors conclude, "Our findings suggest that a comprehensive tracking program that establishes a separate classroom in every school for the top performing students could significantly boost the performance of the most talented students in even the poorest neighborhoods, at little or no cost to other students or the District's budget."[7]

Despite these positive findings, tracking is more likely to be found in upper-middle-class, suburban schools than in schools serving low-income students. Why is that?

The Recent History of Tracking Reform

Tracking is one of education's most controversial practices. Jeannie Oakes's 1985 book, *Keeping Track*, charged that schools inflict great harm on poor and minority students through tracking. Measures of achievement, which are the basis of most track placements, are highly correlated with students' socioeconomic characteristics. So the book's finding that high tracks are disproportionately populated by white, Asian, and wealthier students and low tracks by black, Hispanic, and poorer students is not surprising.[8]

But Oakes's book went a step further. Tracking's defenders had always argued that the practice is educationally sound, providing students with a curriculum better matched to individual needs compared to randomly grouping students into classes and then giving them all the same curriculum to study. Oakes repudiated the claim, asserting that when students are separated into classes matched to a hierarchy of course topics in mathematics (algebra, geometry, etc.) or assigned to honors, regular, and remedial courses in English language arts, students placed in the top classes are challenged and thrive academically; those in the bottom classes languish. Moreover, Oakes's indictment embraced a political dimension. She charged that schools were catering to the interests of social elites by using tracking to perpetuate the past racial and economic inequities of American society. Tracking was racist and profoundly unfair to disadvantaged students. And it was not that way by accident.

The detracking movement had an impact on policy. By the mid-1990s, many schools—in particular, middle schools—were reducing or eliminating tracking.

A distinct pattern soon emerged as to where detracking was, and was not, implemented.[9] Schools responding positively to tracking reform tended to serve students who were allegedly harmed by the practice—urban or inner-city schools serving predominantly black, Hispanic, or poor populations. These schools began to detrack, abolishing both honors and remedial classes. The schools that resisted tended to be suburban, with predominantly white, Asian, middle- or upper-middle-class students.

By the end of the decade, the distribution of reformed tracking policies—how they mapped over schools with different demographic profiles—mirrored the political appeal of the anti-tracking crusade. An intriguing twist arose. The elites who were supposedly perpetrating tracking for the purpose of maintaining their privileged status—in particular, suburban whites—were imposing that unfair regime on their own children's schools. Disadvantaged and minority children, on the other hand, now were more likely to attend detracked schools with heterogeneously grouped classes.

The Demographics of Tracking and Detracking

Let's examine the latest and most reliable data on the demographics of tracking and detracking. Tables 1–4 present demographic data from NAEP on students who are in detracked classes.

Frequency of Detracking: Student Characteristics

The first table reports on eighth graders who are eligible for free and reduced-price lunch (FRL). In 2013, 25 percent of students eligible for FRL were in detracked math classes compared to 20 percent of students not eligible (from wealthier families). Disadvantaged students are more likely to attend detracked classes than students from more advantaged families. Note that the difference narrowed from 2000 to 2013. Most of it was due to a decline in the frequency of detracked FRL students. In 2000, 34 percent of FRL students attended heterogeneously grouped classes, compared to 21 percent of students not eligible for FRL. Although stigmatized as harming disadvantaged children throughout the 1990s, tracking has staged a partial comeback since 2000. Poor kids were more likely to be tracked in 2013 than in 2000, but were still less likely to attend tracked classes than their peers from wealthier families.

Table 1: Percentage of Detracked Eighth Graders
By Eligibility for Free and Reduced Lunch

Year	ELIG-FRL	NOT ELIG-FRL
2013	25	20
2011	28	21
2009	28	20
2007	31	20
2005	33	23
2003	34	21
2000	34	21

Source: Compiled by author from NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/naepdata/)

Most of the narrowing occurred after 2007. It is possible that recent changes in eligibility for free and reduced-price lunch (there have been several) are skewing the data. Policy innovations, such as allowing schools to offer free and reduced-price meals to all students if 50 percent or more qualify for the program, might draw students into the FRL program whose families have higher incomes than the FRL threshold.[10] Free and reduced-priced lunch data are losing their cache as an indicator of poverty. Tables 2 and 3 provide a check on FRL's weakening validity, and by examining other student characteristics, also allow a more complete picture to emerge of the demographics of tracking.

Table 2 looks at tracking and parent education. The columns are arranged left to right by increasing levels of education, meaning that, given the strong relationship between education and income, the columns also correlate with family socioeconomic status.

Table 2: Percentage of Detracked 8th Graders

By Parent Education

	NO HS DIPLOMA	HS GRAD	SOME POST-HS ED	COLL GRAD
2013	24	26	24	20
2011	30	27	25	22
2009	28	27	24	20
2007	30	28	26	22
2005	32	30	28	24
2003	32	29	27	22
2000	28	27	25	22

Source: Compiled by author from NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/naepdata/)

In 2013, 24 percent of students whose parents lacked a high school diploma were in detracked classes compared to only 20 percent of students whose parents were college graduates. Consistent with the FRL data in the previous table, Table 2 shows that disadvantaged students are more likely to be detracked than advantaged students. There is also a hint of narrowing from 2000–2013 as the 4 percentage-point difference between "No HS Diploma" and "College Grad" in 2013 is the smallest of the period. The 10-point difference in 2003 is the largest.

Table 3 examines detracking by students' race and ethnicity. Black students are consistently more likely to attend detracked classes than white and Asian/Pacific Island students. In 2013, 29 percent of blacks were in detracked classes versus 22 percent of Hispanics, 19 percent of Asian/Pacific Islanders, and 21 percent of whites. But also note that by 2013, Hispanic students had closed what was a huge gap in 2000 (14 percentage points compared to whites). In 2013, the frequency of Hispanic students studying math in detracked classes was statistically indistinguishable from whites.

Table 3
Percentage of Detracked Eighth Graders
By Race/Ethnicity

	HSP	BLK	ASIAN/PI	WHITE
2013	22	29	19	21
2011	29	30	22	21
2009	25	32	15	21
2007	29	33	21	21
2005	30	35	20	25
2003	32	35	20	22
2000	34	33	24	20

Source: Compiled by author from NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/naepdata/)

Frequency of Detracking: School Characteristics

Tables 1–3 reveal how student characteristics are related to attending detracked classes. Students from impoverished backgrounds are less likely to be tracked, and the data on both parent education and race/ethnicity confirm the tendency. Looking at student

characteristics only goes so far, however, in explaining why some students are more likely to be tracked than others. School characteristics are also important. It's reasonable to assume that some students are tracked or detracked because of the schools they attend, not necessarily because of their own individual characteristics.

Table 4 shows the percentage of detracked students by school wide statistics on free and reduced-price lunch. This is a measure of concentration of poverty, and the columns are arranged left to right with the level of school poverty descending.

Table 4: Percentage of Students Detracked By School's Percentage of Students Eligible for Free and Reduced Lunch

	99-76%	75-51%	50-26%	25-1%
Year	High Poverty <----------> Low Poverty			
2013	30	23	22	16
2011	35	27	22	16
2009	38	27	20	15
2007	42	27	23	16
2005	43	32	26	17

Note: Author's calculations (narrower categories of NAEP data have been aggregated). Schools with 100% and 0% FRL omitted.
Source: Compiled by author from NAEP Data Explorer (http://nces.ed.gov/nationsreportcard/naepdata/)

Schools in the farthest left-hand column (99–76 percent eligible for free lunch) are high-poverty schools; schools in the farthest right-hand column (25–1 percent eligible) are low-poverty (serving students from wealthier families). Schools with 100 percent of students eligible for FRL have been omitted. They only comprise 3 percent of schools, and omitting them helps dampen any skewing caused by recent FRL policy changes that artificially boosted their numbers. Schools with 0 percent FRL have also been omitted. They do not represent enough schools in the NAEP sample to generate reliable data.

The two extreme categories offer a telling comparison.[11] In 2013, high-poverty schools were about twice as likely to have detracked math classes as low-poverty schools (30 percent vs. 16 percent). But again, there is a clear trend toward increased tracking over the past decade, especially in high-poverty schools. In 2005, 43 percent of high-poverty schools and only 17 percent of low-poverty schools were detracked. The percentage of detracked schools has remained stable in schools serving advantaged communities but is clearly falling in schools serving children in poverty.

Research on the Effects of Tracking

The literature on tracking goes back to the early twentieth century. Most tracking studies unfortunately are vulnerable to selection effects; that is, they cannot untangle the effects of different tracks from the characteristics of students populating those tracks. When high-track students outgain low-track students on achievement tests, is it because of advantages produced by enrollment in a high track? Or would they have outscored low-track students

under any conditions, considering that some students are assigned to high tracks because of their strong academic abilities and other students to low tracks because of their academic weaknesses?

Even studies that attempt to control for initial achievement (e.g., use test scores prior to track placement as covariates) can overlook bias introduced by omitted variables. Schools may place two students with the same initial test scores into different tracks because one student has better attendance, consistently puts forth greater effort, or has a highly involved parent who aggressively lobbies the school for assignment to the high track. If such variables are unmeasured or unaccounted for in estimating track effects—and if they are correlated with the outcome variable of interest (student achievement)—the results of statistical models may be biased.

Experimental studies featuring random assignment and quasi-experimental studies with strong designs try to mitigate both of these shortcomings and offer the best evidence on tracking's effects. Meta-analyses of experimental and quasi-experimental studies have been conducted by Kulik and Kulik (1982), Slavin (1990), and Mosteller, Light, and Sachs (1996).[12] These meta-analyses reach remarkably similar conclusions about tracking's effect on mean student achievement—that it's statistically insignificant.[13] Schools can choose to track or to group students heterogeneously and the decision will have no impact on average achievement.

Does Tracking Benefit High Achievers?

Where the meta-analysts disagree is on the distribution of effects—whether some students benefit or are harmed by either practice. Slavin's review of twenty-nine studies of tracking found tiny, insignificant effect sizes for high achievers (+0.01), students near average achievement (-0.08), and low achievers (-0.02). Mosteller, Light, and Sachs identified only ten experiments that evaluated the distribution of effects from tracking (which they call "skill grouping"). Noting that the small number of studies left the analysis underpowered (i.e., a larger number of studies may have detected an effect), Mosteller, Light, and Sachs found statistically insignificant effects but described "a slight tilt toward skill grouping being more favorable for high-skill than for medium- and low-skill students. The estimates of average effect sizes were 0.08 for high, -0.04 for medium, and -0.06 for low-skill groups."

These reviews focused on XYZ style tracking, named for a program that began in Detroit in 1919. The school system administered IQ tests to all incoming first graders and then divided students into three separate classes: X for the top 20 percent, Y for the middle 60 percent, and Z for the bottom 20 percent. The three classes all received the same curriculum.

Kulik and Kulik's review of XYZ studies found a small positive effect for high achievers, but not large enough to represent a substantive, real-world benefit. But Kulik and Kulik went a step further by separately analyzing programs in which curricular differentiation occurred. They argued that modern tracking, at least since the 1950s, exists to match students with an appropriate curriculum, not just to reduce the heterogeneity of ability within classes. Here they discovered a huge benefit. They found that talented students who were tracked into classes with an enriched curriculum made significant gains (mean effect size 0.41)

compared to students of equal ability who were untracked. High-ability students who were placed in tracks featuring an accelerated curriculum, meaning the curriculum covered topics from later grades, benefitted the most, gaining a whopping 0.87 SDs (nearly a year of learning) compared to untracked peers.

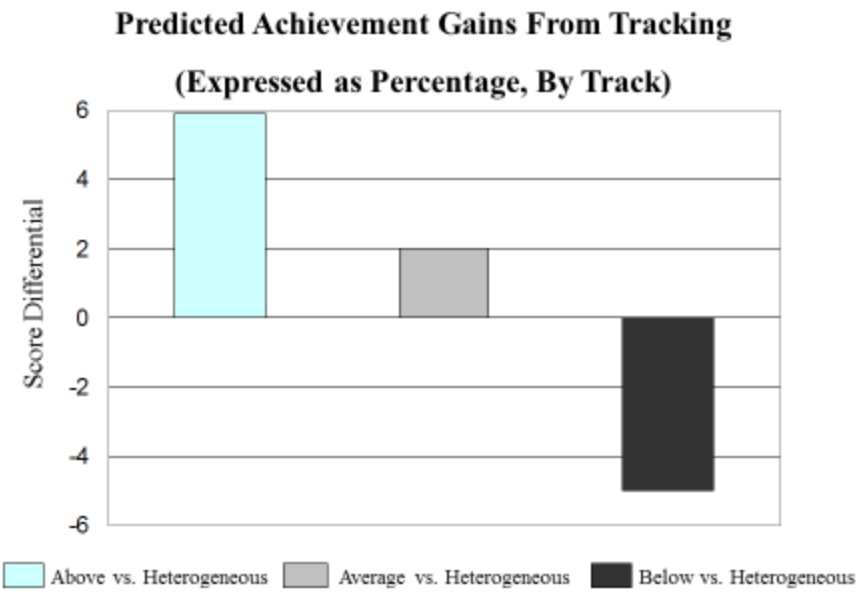
The last experimental study in the Mosteller, Light, and Sachs meta-analysis was conducted in 1974. The next experiment to evaluate tracking was conducted in Kenya in 2005. Like most experiments in education, a propitious set of circumstances allowed for the random assignment of students to treatment groups. Schools in western Kenya received extra funds to hire first-grade teachers. At the time, 121 of these schools had only one teacher teaching first grade, meaning that they would now go from one to two first-grade classes. The schools were randomly assigned to either: a) a tracked condition, in which students above the mean were assigned to one class and students below the mean to the other, or b) an untracked condition, in which students were randomly assigned to the two classes.

The experiment ran for eighteen months. All students in the tracked classes benefitted, with a mean advantage of 0.18 standard deviations after controlling for baseline scores. Students in the top half of the pre-assignment distribution (i.e., higher achievers) gained 0.19 standard deviations, and those in the bottom half (i.e., lower achievers) gained 0.16 standard deviations. Students in the middle of the pre-assignment distribution—near the cut point for assignment to either the top or bottom class—gained as much as either the top or bottom students.

Students were tested again one year after the experiment ended. The positive effect persisted, with students in the tracked condition continuing to score 0.18 standard deviations higher than students in the heterogeneously grouped classes. The authors speculated that teachers were better able to tailor instruction to students' academic needs in the tracked classes, regardless of whether a student attends a high- or low-ability class. Data on instruction were not collected in the study.

Argys, Reese, and Brewer analyzed data from the National Education Longitudinal Study of 1988 (NELS:88). They looked at tenth-grade math scores and measured the effect of students being placed in a heterogeneously grouped class as opposed to an above-average, average, or below-average tracked class. From this analysis, they were able to predict the advantage or disadvantage that tracking provides to students in different tracks (see Figure 1). The main finding of the study is that tracking creates winners and losers. High- and average-track students are both winners under tracking. Being placed in an above-average track produces a 6 percent gain in achievement compared to being placed in a heterogeneously grouped class. Being placed in an average track yields a 2 percent gain. Low-track students, on the other hand, pay a price under tracking, with a loss of about 5 percent in math achievement. The gains and losses average out to a 1.7 percent gain, the magnitude of achievement benefit expected if a detracked school switched to tracking.

Figure 1
Achievement Gains from Tracking
By Track Membership
(Argys, Rees, and Brewer, 1996)



Source: Argys, Rees, & Brewer, "Detracking America's Schools: Equity at Zero Cost?" *Journal of Policy Analysis and Management*, 15 (Autumn 1996), pp. 623-645.

Elaine Allensworth investigated a detracking policy in Chicago. In 1997, the school system adopted an "Algebra for All" policy for ninth graders. Remedial math was abolished and mixed-ability algebra classes were created. Using a discontinuity design, the researchers found that the detracking effort increased the number of students receiving algebra credits but had had no effect on the test scores of average and below-average students who were intended to benefit from the policy. Failure rates in math classes increased for average students. Takako Nomi examined the effect of the policy on high-achieving students. Although these students were not the targets of the policy—they would have taken algebra anyway regardless of the policy change—they were affected by the change in classroom composition of Algebra I classes. Their classes now held low-achieving peers. Detracking had the unintended effect of driving down the scores of high achievers.

In 2009, I conducted a survey of tracking practices in Massachusetts middle schools. In addition to finding the same demographic pattern as reported above in NAEP data—high-poverty, urban schools were less likely to offer tracked classes than suburban schools serving mostly advantaged students—regression analysis controlling for SES characteristics found that detracked schools lagged in achievement, especially when it came to high achievers. Schools featuring three eighth-grade math tracks were associated with a six percentage-point gain in the number of students scoring at the "advanced" level in mathematics, compared to detracked schools offering heterogeneously grouped math classes. Considering that only 18 percent of eighth graders scored "advanced" at the average middle school in the state, a six percentage-point gain would represent a significant real world impact on school achievement.

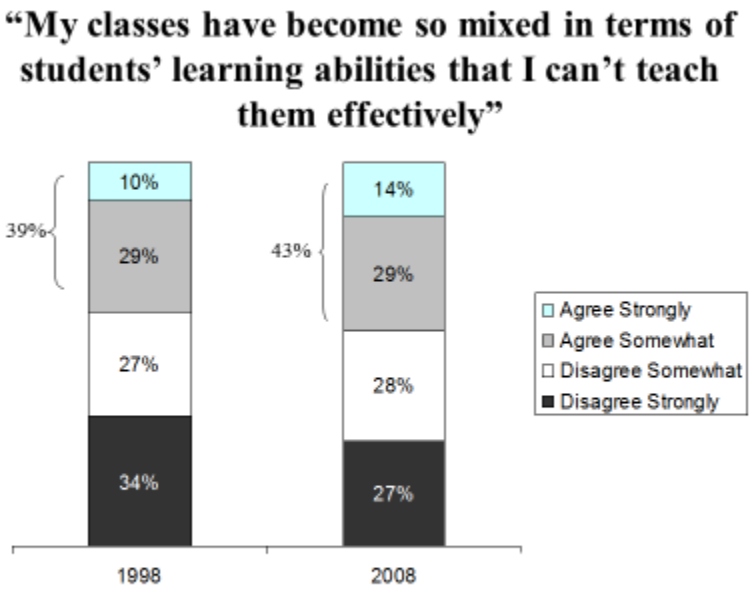
These data did not allow for causal conclusions. Were tracked schools creating more advanced students with their high tracks—or were families with advanced students simply

more likely to attend tracked schools? David Figlio and Marianne Page were able to tease out causal effects by investigating tracking and its relationship to school choice. They found that tracked schools attract and retain families with greater wealth. And after controlling for confounding factors with an instrumental variable strategy, Figlio and Page also discovered that disadvantaged students make larger test-score gains in tracked rather than detracked schools. They concluded, "We can find no evidence that detracking America's schools, as is currently in vogue, will improve outcomes among disadvantaged students. This trend may instead harm the very students that detracking is intended to help."[14]

Conclusion

Teachers find teaching classrooms with vast heterogeneity in ability a hindrance to effective instruction. A 2007 survey by the National Math Advisory Panel reported that a majority of algebra teachers viewed mixed-ability grouping as a moderate (28 percent) or serious (23 percent) problem. And the problem seems to be getting worse. The Met Life Survey of the American Teacher asked about the impact of class heterogeneity in 1998 and 2008 (see Figure 2).

Figure 2
Teachers' Reactions to Classes Heterogeneous in Ability
(Met Life Surveys, 1998 & 2008)



Source: MetLife, Inc, *The MetLife Survey of The American Teacher* (MetLife, 2008).

The percentage of teachers agreeing either agreeing somewhat or agreeing strongly to the statement "My classes have become so mixed in terms of students' learning abilities that I can't teach them effectively" grew from 39 percent in 1998 to 43 percent in 2008. Considering the stark wording of the prompt—the consequence of instructing students of mixed abilities being "that I can't teach them effectively"—the fact that the statement resonates with more than 40 percent of teachers is indeed troubling.

The Common Core State Standards (CCSS) in mathematics offer an excellent opportunity to experiment with tracking for equity. Implementing CCSS is itself an experiment. The CCSS math standards recognize that differentiation may have to take place as schools implement curriculum. The CCSS math standards are only written through Algebra II in the junior year of high school. Some advanced math students will be able to move faster than that.

Appendix A of the CCSS math standards suggests "compacting" in middle school as an acceleration strategy, defining it as:

> A "compacted" version of the Traditional pathway where no content is omitted, in which students would complete the content of 7th grade, 8th grade, and the High School Algebra I course in grades 7 (Compacted 7th Grade) and 8 (8th Grade Algebra I), which will enable them to reach Calculus or other college level courses by their senior year.

Appendix A goes on to state the rationale for compacting:

> Based on a variety of inputs and factors, some students may decide at an early age that they want to take Calculus or other college level courses in high school. These students would need to begin the study of high school content in the middle school, which would lead to Precalculus or Advanced Statistics as a junior and Calculus, Advanced Statistics or other college level options as a senior.[15]

Design Features of Tracking for Equity in Mathematics

As the literature review above suggests, tracking entails risks. Opponents of tracking are justifiably concerned about the effects of tracking on students in low tracks. Adam Gamoran warns,

> "… it is important to acknowledge that most studies of ability grouping and curriculum tracking have found that high-achieving students tend to perform better when assigned to high-level groups than when taught in mixed-ability settings. Proponents of tracking tend to emphasize the benefits of high-level classes for high-achieving students with little attention to implications for inequality, while critics tend to focus on the inequality without acknowledging the effects for high achievers. As a result, proponents and critics are apt to talk past one another with little chance for resolution…"[16]

With that warning in mind, I propose design features of a tracking-for-equity program in middle school mathematics, a program that would offer an accelerated math track. These features are intended to maximize the benefits of tracking for high achievers while minimizing any potential negative impact on low-achieving students. The program offers an accelerated option for states following the Common Core. It costs next to nothing. A tracking-for-equity program could be adopted by most low-income schools tomorrow.

1. Create an accelerated track for seventh and eighth graders that would complete three years of mathematics—meeting the standards for seventh grade, eighth grade, and Algebra I—in two years.
2. Implement the program in low-income schools first. An analysis of students scoring at the highest levels on the NAEP math test discovered a huge pool of untapped talent among disadvantaged eighth graders. In 2005, approximately one-quarter of black, Hispanic, and low-income eighth graders were enrolled in a course lower than Algebra I (pre-algebra or regular eighth-grade math) despite scoring at the 90th percentile on NAEP.[17]
3. Track by performance on subject-specific assessments. Research on omnibus tracking programs in secondary schools is mixed. By "omnibus" tracking, I am referring to programs that identify students with an IQ test or comprehensive test of academic abilities and then track students into all or most subjects. Indeed, this is how most U.S. high schools tracked until the early 1970s, as students were placed into honors, regular, or vocational tracks that dictated how they were grouped for all subjects. Omnibus systems may not adequately match students to curriculum. Students who are great at math but struggle with reading will not receive a mathematics curriculum appropriate to their talents if the placement assessment is heavily weighted toward verbal skills. English language learners who are talented at math may also be overlooked.

   Omnibus tracking between schools describes the tracking programs of most European and Asian nations. Students typically take a high-stakes test at the end of junior secondary school (i.e., middle school in the United States) and the test score on that exam decides whether they attend an academic, technical, or vocational high school. Omnibus tracking also describes American exam schools. Research on their effectiveness is rare, but a couple of recent, high-quality studies found that exam schools have no effect.
4. Alter the curriculum to match the track. Tracking appears to produce its greatest benefits when curriculum is altered to correspond to the levels of students. Compacting three years of CCSS into a two-year program would provide the necessary acceleration for talented math students.
5. Attend to the lowest track. Offering an accelerated track in low-income schools means that some students will not be accelerated, with the risk that students in non-accelerated classes will encounter low expectations or that the classes will deteriorate into curricular dead ends. In addition, teachers of low tracks in middle schools must be skilled in classroom management, because low-achieving students often exhibit a constellation of anti-learning behaviors by the time they reach middle school. Chicago addressed this problem by providing a double-dose of algebra—two periods daily—to students who scored below the national median on incoming math assessments.

   The results from the double-dose intervention are encouraging. Kalena Cortes and Joshua Goodman find that the increased class time devoted to studying

> algebra led to significant positive effects on both short-term math achievement and longer-term outcomes (e.g., course-taking in high school). In this program, as in the many evaluated in the research summarized above, tracking had been effectively enlisted in the cause of equity.[18]

The analysis above supports the hypothesis that tracking may be deployed in the pursuit of equity. Identifying academically talented students in high-poverty middle schools and offering them a curriculum tailored to their needs could create a pipeline of academic excellence running from middle school to high school AP classes. It promises better preparation of disadvantaged students for high school AP courses than that provided by current policies and practices. Low-income middle school students do not have access to high tracks because, in contrast to schools serving predominantly advantaged students, schools in low-income communities are more likely to be untracked. They have adopted an ethos of equal access to all courses; thus, their toughest, most academically challenging classes may be populated with students so heterogeneous in ability that students two or more years above grade level sit alongside students two or more years below grade level—all students studying the same curriculum and receiving the same instruction. Teachers have difficulty teaching such classes.

What middle schools in poor communities need is an ethos of identifying and developing academic talent. This already happens in sports, often to an extreme.[19] It's time to de-stigmatize the word "tracking." A middle school curricular program that offers high-achieving, disadvantaged students the opportunity to excel would advance the cause of educational equity.

---

[1] College Board, *The 10th Annual AP Report to the Nation*, 2014, page 6.
[2] Liz Bowie, "Maryland Schools Have Been Leader in Advanced Placement, but Results are Mixed," *Baltimore Sun*, August 21, 2013.
[3] Sean Reardon, "The Widening Achievement Gap between the Rich and the Poor," in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (New York, NY: Russell Sage Foundation: 2011).
[4] Eric A. Hanushek and Steven G. Rivkin, "School Quality and the Black-White Achievement Gap," NBER Working Paper Series, Working Paper 12651 (Cambridge, MA: National Bureau of Educational Research, October 2006).
[5] David Card and Laura Giuliano, "Does Gifted Education Work? For Which Students?" NBER Working Paper Series, Working Paper 20453 (Cambridge, MA: National Bureau of Educational Research, September 2014).
[6] Card and Giuliano, "Does Gifted Education Work?" 5–6.
[7] Card and Giuliano, "Does Gifted Education Work?" 33–34.
[8] Jeannie Oakes, *Keeping Track: How Schools Structure Inequality* (New Haven: Yale University Press, 1985).
[9] Tom Loveless, *The Tracking Wars: State Reform Meets School Policy* (Washington, D.C.: Brookings Institution Press, 1999).
[10] Sarah D. Sparks, "Popular Child-Poverty Measure Gets Another Look," *Education Week*, August 19, 2014.
[11] Schools with 76 percent to 99 percent of students eligible for FRL were about 19 percent of schools in the 2013 NAEP. Schools with 25–1 percent FRL were about 23 percent of the sample.
[12] A more recent meta-analysis is Hattie (2009), which is essentially a meta-analysis of meta-analyses. I did not include Hattie's estimates in the narrative because he did not limit the meta-analysis to studies meeting criteria for methodological quality. Nonetheless, the effect sizes he reports for ability grouping (0.12) and ability grouping for gifted students (0.30) do not contradict the findings of the other meta-analyses discussed. Hattie uses the term "ability grouping" for tracking.

[13] The earliest study in Slavin's meta-analysis is from 1927.
[14] Figlio and Page, 2012, page 29.
[15] Common Core State Standards for Mathematics, Appendix A, page 3.
[16] Adam Gamoran, WCER, 2009, 8–9
[17] Tom Loveless, *High Achieving Students in the Era of NCLB* (Washington, D.C.: Thomas B. Fordham Institute, 2008). (This refers to Part 1. Part 2 is by Steve Farkas and Anne Duffett.)
[18] Kalena E. Cortes and Joshua S. Goodman, "Ability Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement," *American Economic Review: Papers & Proceedings 2014* 104, no. 5, 400–405.
[19] Adam Himmelsbach and Pete Thamel, "Middle School is Basketball's Fiercest Recruiting Battleground," *New York Times*, June 25, 2012. The University of Washington and Louisiana State University have promised scholarships in exchange for signed letters of commitment from eighth-grade football players. In 2010, David Sills, a thirteen-year-old quarterback in Elkton, MD committed to USC when he was in seventh grade.