

The Tracking and Ability Grouping Debate

by [Tom Loveless](#)

07/01/1998



by Tom Loveless

Section Three: The Research

The research on tracking and ability grouping is frequently summarized in one word: inconclusive. This pronouncement is accurate in that nearly a century's worth of study has failed to quantify the impact of tracking and ability grouping on children's education. It doesn't necessarily mean, however, that the gallons of ink spilled on these issues have been much ado about nothing. A non-effect in educational research is quite common. It can mean that the practice under study is truly neutral vis-a-vis a particular outcome. But it can also mean that the practice has off-setting negative and positive effects, that positive effects are produced under some conditions and negative effects under others, or that effects occur that researchers either don't measure, because they're measuring something else, or can't measure, because of inadequate methods or expertise.

Non-findings must be interpreted with great care, especially when looking for policy guidance. In 1966, a federal report was released that many scholars consider the single most famous study in the history of education, *Equality of Educational Opportunity*, otherwise known as the Coleman Report for its primary author, the famed sociologist James Coleman. The Coleman Report was widely interpreted as finding that schools themselves have no significant effect on student learning. Fortunately, policymakers did not rush out to close schools and turn them into car washes or something else more useful.

I will review the research on tracking and ability grouping by first surveying what is known about its effect on academic achievement and then examining the evidence on five of the most serious charges leveled against tracking.

Achievement

Tracking's non-effect on achievement can be appreciated by contrasting the conclusions of two prominent analysts, Robert Slavin, a critic of tracking, and James G. Kulik, a defender of some forms of tracking and ability grouping. Both have conducted meta-analyses of tracking and ability grouping. A meta-analysis is essentially a study of studies. The analyst pools the existing studies that meet certain criteria for quality and statistically summarizes their conclusions. As an indication of the massive amount of material with which a meta-analysis on this topic begins, Kulik's initial search uncovered over 700 studies on tracking and ability grouping.

First, the points of agreement. Slavin and Kulik agree that studies of within-class ability grouping are positive, with Slavin's support largely resting on the benefits uncovered for grouping in mathematics in the upper grades of elementary school. They also agree that cross-grade ability grouping boosts achievement in elementary schools. The most popular form of this approach, the "Joplin Plan," originated in Joplin, Missouri in the 1950s. In short, Slavin and Kulik validate the most widely used forms of ability grouping at the elementary level. Ability grouping promotes achievement, and no particular group of children—high, middle, or low ability—misses out on the gain (see the Appendix).

The analysts diverge on between-class grouping, or tracking. Because the national debate, like the Slavin-Kulik debate, focuses so intensely on tracking, and because there are several facets of the practice that are controversial, I will confine the remainder of this discussion to the tracking research. Slavin and Kulik basically agree that "XYZ" grouping plans have no significant effect on learning. A species of tracking, this scheme gets its name from Detroit's XYZ program, which began in 1919. In most of the XYZ studies, schools ranked students by IQ test or some other omnibus test of ability, grouped the students into separate classes (in Detroit, labeled X, Y and Z), and taught an identical curriculum to all three groups. The XYZ students' achievement was then compared to that of similar students in ungrouped classes. Taken as a whole, the best XYZ studies show no difference between ability grouped and ungrouped students.³⁶

Slavin concludes from this evidence that tracking has no effect on achievement. Kulik points out that XYZ bears little resemblance to the way most schools use tracking in the real world. Schools typically use tests that measure achievement in specific disciplines to ability group students in each subject. They no longer use IQ tests or the other omnibus measures that were used to form XYZ groups. And students are assigned to tracks for the express purpose of adjusting the curriculum to students' ability. Since all levels of XYZ typically studied an identical curriculum, Kulik argues that its negligible effect on achievement is not surprising.³⁷

Pursuing this line of inquiry, Kulik finds that tailoring course content to ability level yields a consistently positive effect on the achievement of high ability students. Academic enrichment programs produce significant gains. Accelerated programs, where students tackle the curriculum of later grades, produce the largest gains of all. Accelerated gifted students dramatically outperform similar students in non-accelerated classes. Slavin omits studies of these programs from his analysis. He argues that the gains, though large, may be an artifact of the programs' selection procedures, that schools admit the best students into these programs and reject the rest, thereby biasing the results.³⁸

Three things are striking about the Slavin- Kulik debate. First, the disagreement hinges on whether tracking is neutral or beneficial. Neither researcher claims to have evidence that tracking harms achievement, of students generally or of students in any single track. Second, accepting Slavin or Kulik's position on between-class grouping depends on whether one accepts as legitimate the studies of academically enriched and accelerated programs. Including these studies leads Kulik to the conclusion that tracking promotes achievement. Omitting them leads Slavin to the conclusion that tracking is a non-factor.

Third, in terms of policy, Slavin and Kulik are more sharply opposed on the tracking issue than their other points of agreement would imply. Slavin states that he is philosophically opposed to tracking, regarding it

as inequalitarian and anti-democratic. Unless schools can demonstrate that tracking helps someone, Slavin reasons, they should quit using it. Kulik's position is that since tracking benefits high achieving students and harms no one, its abolition would be a mistake.

More meta-analyses will not resolve this philosophical dispute. Furthermore, the XYZ studies that Slavin and Kulik are scrutinizing fall short of providing a clear policy direction. The studies vary on critical dimensions, and important variables go unreported.³⁹ Actually, we don't know a lot about the education children received in the studies—in either the grouped or the ungrouped settings. Moreover, virtually all of the studies that Slavin and Kulik reviewed were conducted before 1975. The structural changes that have occurred in tracking since then are not represented in the XYZ literature. School people can't search through this mass of research, find a school that has similar practices to their own, and figure out whether their school's tracking system is good or bad or how it can be improved.

National Data

In the last two decades, researchers have also analyzed large national surveys to evaluate tracking. High School and Beyond (HSB) is a study that began with tenth graders in 1980. The National Education Longitudinal Study (NELS) started with eighth graders in 1988. These two studies followed tens of thousands of students through school, recording their academic achievement, courses taken, and attitudes toward school. The students' transcripts were analyzed, and their teachers and parents were interviewed. The two massive databases have sustained a steady stream of research on tracking.

Three findings stand out. High track students in HSB learn more than low track students, even with prior achievement and other pertinent influences on achievement statistically controlled. Not surprising, perhaps, but what's staggering is the magnitude of the difference. On average, the high track advantage outweighs even the achievement difference between the student who stays in school until the senior year and the student who drops out.⁴⁰

The second major finding is that race and tracking are only weakly related. Once test scores are taken into account in NELLS, a student's race has no bearing on track assignment. In fact, African-American students enjoy a 10% advantage over white students in being assigned to the high track. This contradicts the charge that tracking is racist. Considered in tandem with the high track advantage just described, it also suggests that abolishing high tracks would disproportionately penalize African-American students, especially high achieving African-American students. Moreover, NELLS shows that achievement differences between African-American and white students are fully formed by the end of eighth grade. The race gap reaches its widest point right after elementary and middle school, when students have experienced ability grouping in its mildest forms. The gap remains unchanged in high school, when tracking between classes is most pronounced.⁴¹

Third, NELLS identifies apparent risks in detracking. Low-achieving students seem to learn more in heterogeneous math classes, while high and average achieving students suffer achievement losses—and their combined losses outweigh the low achievers' gains. In terms of specific courses, eighth graders of all ability levels learn more when they take algebra in tracked classes rather than heterogeneously grouped classes. For survey courses in eighth grade math, heterogeneous classes are better for low achieving students than tracked classes.⁴²

These last findings are important because we don't know very much about academic achievement in heterogeneous classes. When the campaign against tracking picked up steam in the late 1980s, tracking was essentially universal. Untracked schools didn't exist in sufficient numbers to evaluate whether abandoning tracking for a full regimen of mixed ability classes actually works. The NELLS studies that attempt to evaluate detracked classes, which thus far have been restricted to mathematics, point toward a

possible gain for low achieving students and a possible loss for average and above average students, but these findings should be regarded as tentative.⁴³

To summarize what we know about ability grouping, tracking, and achievement: The elementary school practices of both within-class and cross-grade ability grouping are supported by research. The tracking research is more ambiguous but not without a few concrete findings. Assigning students to separate classes by ability and providing them with the same curriculum has no effect on achievement, positive or negative, and the neutral effect holds for high, middle, and low achievers. When the curriculum is altered, tracking appears to benefit high ability students. Heterogeneous classes appear to benefit low ability students but may depress the achievement of average and high achieving students.

Fosters Race and Class Segregation?

Critics charge that tracking perpetuates race and class segregation by disproportionately assigning minority and poor children to low tracks and white, wealthy children to high tracks. When it comes to race, the disparities are real, but, as just noted, they vanish when students' prior achievement is considered. A small class effect remains, however. Students from poor families are more likely to be assigned to low tracks than wealthier students with identical achievement scores. This could be due to class discrimination, different amounts of parental influence on track assignments, or other unmeasured factors.⁴⁴

The issue ultimately goes back to whether tracking is educationally sound. Those who complain of tracking's segregative impact do not usually attack bilingual or Title I programs for promoting ethnic and class segregation, no doubt because they see these programs as benefiting students. If low tracks remedied educational problems, the charge of segregation would probably dissipate. Does tracking harm black students? A telling answer is found in African-American parents' attitude toward tracking. A study conducted by the Public Agenda Foundation found that "opposition to heterogeneous grouping is as strong among African-American parents as among white parents, and support for it is generally weak."⁴⁵ If tracking harmed African-American students, one would not expect these sentiments.

Harms Self-Esteem?

Little research indicates that tracking harms students' self-esteem. In fact, the evidence tilts slightly toward the conclusion that low ability students' self-concept is strengthened from ability grouping and tracking, although the effect is insignificant. The public labeling of low track students may cause embarrassment, but the public display of academic deficiencies undoubtedly has a similar effect in heterogeneous classrooms. There, a low ability student's performance is compared daily to that of high-achieving classmates.⁴⁶

Locks Students In?

It would be reprehensible if students were denied the opportunity to move up in track or denied, in the tracking critics' phrase, "access to knowledge," the learning that gets students into college and ultimately better their lives. Data on this issue are difficult to interpret. Mobility rates tell us how much movement occurs, but they don't answer the key question of whether that movement is warranted. For some students, keeping them in the same group year after year may be wise, while for others, moving them up or down in group may be the educationally prudent decision.

How much mobility takes place? A study of transcripts from five Maryland high schools showed 59.9% of

students changed math levels during their high school careers, 65.4% in science. A national survey of high school principals reports substantial movement among tracks, especially upward (see Table 7). But an analysis of NELS data found that only 16.5% of students who were in low-ability classes in 8th grade went on to take either geometry or Algebra II by 10th grade (in comparison to 81.0% of 8th graders in high-ability classes).

Table 7
Change in Track Level After Grade 10, 1993
(% of schools)

Subject	Direction	Almost Never	Rarely	Sometimes	Often
Math	Moved Up	12	27	47	14
Math	Moved Down	16	33	50	2
English	Moved Up	17	23	46	14
English	Moved Down	26	34	39	1

Note: Principals were asked: "How often are students changed to a higher [lower] ability course after completion of 10th grade?"

The results in science were not so dismal, with 61.7% of students in low-ability eighth grade science able to complete biology or chemistry by 10th grade.⁴⁷ These data suggest that substantial movement among groups and tracks occurs. That being said, a disturbing number of students never emerge from the low track. Even where the opportunity to move up and out of low tracks exists, the qualities that one must have to seize this opportunity—strong achievement motivation, independence, and drive—may be lacking in many low track students. Without a push, a lot of students remain in low tracks who are capable of moving up.

High Track Privilege?

Critics of tracking charge that high tracks get more resources than low tracks. Detailed data on school budgets are sparse, and inconsistent expense categories render them almost impossible to compare across schools. It appears that high tracks are taught by better qualified teachers, however, in the sense of having teachers more schooled in content knowledge.⁴⁸ High school principals are inclined to assign teachers who know advanced subject matter to teach advanced subjects. As pointed out by high track defenders, the alternative is unattractive. Does it advance the cause of equity to have teachers with advanced degrees in mathematics teach basic arithmetic while teachers without a single college math course teach calculus? A better solution is to insist that all students take more challenging classes and to staff these classes with well-qualified teachers.⁴⁹

Dead-End Curricula of Low Tracks

Reba Page's 1991 study, *Lower Track Classrooms*, painstakingly reports on the daily activities of eight low track classes, documenting how they often function as caricatures of high tracks, how teachers and students in low tracks make deals to not push each other too hard so that they can cope with their environment. Low tracks may be used as holding tanks for a school's most severe behavior problems. Even under the best of conditions, low tracks are difficult classrooms. The low tracks that focus on academics often try to remediate through dull, repetitious seatwork. This is not to disparage the low track teacher. Research has yet to discover any magic bullets for alleviating tough learning problems or the destructive behaviors that students often exhibit along with them.

Intellectually stimulating low track classrooms do exist, however, and researchers have found the most productive of them in Catholic schools. Margaret Camarena and Adam Gamoran have described low track classrooms where good teaching, lively discussions, and ample learning take place. In 1990, Linda Valli published her study of a heavily tracked Catholic high school in an urban community. The school's course designations publicly proclaimed each student's track level. Textbooks and instruction were adapted for each track. Yet Valli discovered that "a curriculum of effort" permeated the entire school, even the lowest tracks. The school culture centered around academic progress, and the tracking system was but another facet of the school that served this aim. Students of all abilities were aggressively pushed to learn as much as they could. Every year, low track students were boosted up a level. By the senior year, the lowest track no longer existed. A judicious tracking system teaches low track students what they need to know and moves them out of the low track as quickly as possible.⁵¹

Copyright ©2003-2008 The Thomas B. Fordham Foundation. All Rights Reserved.

The Foundation is neither connected with nor sponsored by Fordham University.