AMERICAN SOCIOLOGICAL ASSOCIATION

Sociology of Education
2019, Vol. 92(3) 293–322

DOI: 10.1177/0038040719851879
journals.sagepub.com/home/soe
SAGE

Beyond Tracking and Detracking: The Dimensions of Organizational Differentiation in Schools

Thurston Domina[1], Andrew McEachin[2], Paul Hanselman[3], Priyanka Agarwal[3], NaYoung Hwang[3], and Ryan W. Lewis[4]

Abstract

Schools use an array of strategies to match curricula and instruction to students' heterogeneous skills. Although generations of scholars have debated "tracking" and its consequences, the literature fails to account for diversity of school-level sorting practices. In this article, we draw on the work of Sørensen and others to articulate and develop empirical measures of five distinct dimensions of within-school cross-classroom tracking systems: (1) the degree of curricular differentiation, (2) the extent to which sorting practices generate skills-homogeneous classrooms, (3) the rate at which students enroll in advanced courses, (4) the extent to which students move between tracks over time, and (5) the relationship between track assignments across subject areas. Analyses of longitudinal administrative data following approximately 20,000 eighth graders enrolled in 23 middle schools through the 10th grade indicate that these dimensions of tracking are empirically separable and have divergent relationships with student achievement and the production of inequality.

Keywords

tracking, middle schools, ability grouping, curricular differentiation, organizational differentiation

Schooling may be a "great equalizer" (Downey, von Hippel, and Broh 2004; Mann [1848] 1957; Raudenbush and Eschmann 2015; von Hippel, Workman, and Downey 2018), but at the organizational level, schools are deeply implicated in the production, maintenance, and legitimation of educational inequality. Schools repeatedly sort students, and in the process, they allocate opportunities, resources, and status distinctions unequally across groups (Dreeben and Barr 1988; Kerckhoff 1995; Oakes 1985). Sociologists and other social scientists investigate the mechanisms through which schools generate and perpetuate social inequalities by studying *tracking*, an umbrella term that refers to a broad array of practices associated with grouping students into distinct courses of study.

Several influential studies explore the ideological, political, and technical pressures that lead educators to group students for instruction (Dreeban and Barr 1988; Hallinan 1992; Oakes 1985; Oakes and Guiton 1995; Rosenbaum 1976). Much of this work situates tracking historically.

[1]University of North Carolina-Chapel Hill, Chapel Hill, NC, USA
[2]RAND, Santa Monica, CA, USA
[3]University of California-Irvine, Irvine, CA, USA
[4]WestEd, Sacramento, CA, USA

Corresponding Author:
Thurston Domina, University of North Carolina-Chapel Hill, Peabody Hall, Chapel Hill, NC 27599-3500, USA.
Email: tdomina@email.unc.edu

In the U.S. case, this literature suggests social Darwinist ideas about race and native intelligence enabled the construction of academically differentiated high schools (Cremin 1965; Katz 1971; Kliebard 1975; Oakes 1985), and the civil rights movement's ascendance relaxed and reformed school tracking systems (Lucas 1999; Loveless 2011; Powell, Farrar, and Cohen 1985). But underneath these broad social trends, school-level analyses reveal that the phenomenon we refer to as "tracking" actually consists of a dynamic and diverse set of school practices that change over time and space (Dreeban and Barr 1988; Metz 2003; Oakes 1985; Rosenbaum 1999), often amid considerable uncertainty and debate (Lewis and Diamond 2015; Oakes et al. 1997; Rickles 2011; Watanabe 2006; Wells and Oakes 1996).

The relationship between tracking and the distribution of educational opportunities remains the most persistent question in this literature (Argys, Rees, and Brewer 1996; Carbonaro 2005; Kerckhoff 1986; Oakes 1985). Some of this research takes a macrolevel approach, comparing outcomes across national educational systems that take different approaches to educational stratification (Buchmann and Park 2009; Hanushek and W ößmann 2006; Shavit and Blossfeld 1993). Other recent work takes a microlevel approach, studying individual students' locations in stratified curricular systems and their consequences (Carbonaro and Gamoran 2002; Gamoran and Nystrand 1994; Kelly and Carbonaro 2012; Van Houtte 2004). The sociology of education provides a rich conceptual framework for understanding tracking and its effects, but we argue that both macro- and microlevel work on tracking and its consequences has been hindered by a failure to adequately measure the diverse ways academic tracking manifests in schools. Some studies compare tracked and untracked schools; others compare students placed in different track locations or courses of study within schools (Betts and Shkolnik 2000; Carbonaro 2005; Duflo, Dupas, and Kremer 2008; Figlio and Page 2002; Gamoran et al. 1995; Kelly and Carbonaro 2012). As a result, the existing empirical literature provides limited information about how specific organizational practices associated with tracking matter for student outcomes and inequality.

In this article, we draw on seminal research by Sørensen (1970) and others (Gamoran 1992; Kelly 2007; Lucas 1999; Lucas and Berends 2002) to articulate several dimensions of school-level academic tracking systems. Focusing on middle school mathematics and English courses, we hypothesize that school-level tracking systems differ in at least five important ways: (1) the extent to which schools use distinct courses to differentiate curricula, (2) the degree of within-classroom skills-homogeneity school tracking practices create, (3) the proportion of students who enroll in high-track courses, (4) the amount of between-track mobility that occurs as students move from middle to high school, and (5) the extent to which course placements are related across subjects. Our project contributes to the literature on tracking and its consequences by highlighting a handful of distinct ways schools sort students for instruction. Just as recent work on between-firm variation in workplace inequality highlights the role that local processes play in the production of inequality (e.g., Stainback, Tomaskovic-Devey, and Skaggs 2010), studying school-level tracking systems highlights specific choices schools make and their implications for the production of educational inequality.

We use a unique set of administrative data from three ethnically and economically diverse school districts to measure the dimensions of school tracking systems and study their relation to students' academic skills development. These data capture 20,000 eighth graders in 23 public middle schools. In contrast to the national probability sample data widely used elsewhere in the tracking literature, our data provide detailed longitudinal achievement, demographic, and transcript information for all students enrolled in sample schools. We observe students and all their peers in middle school math and reading classrooms, and we collect supplemental qualitative data on the political and administrative contexts in which these classrooms were situated. In doing so, we generate nuanced and time-varying measures of the dimensions of tracking in each sample school, and we link these measures with subsequent student outcomes.

Our analyses reveal substantial between-school and within-school (over time) variability, with independent variation in the separate dimensions of tracking. We use between-school and temporal variation in our five measures of tracking to test their effects on student outcomes, allowing us to account for persistent unobserved differences between schools. Our findings indicate that different dimensions of school tracking systems have independent (and occasionally counteracting) consequences for student achievement and student achievement inequality. Furthermore, we find suggestive evidence that school-level tracking systems may exacerbate achievement inequalities within schools by providing a boost for high achievers relative to their lower-achieving peers.

TRACKING AND ITS IMPLICATIONS

The study of tracking and its consequences is central to understanding the role education plays in the construction of social inequality. Several studies suggest that students in rigidly tracked schools demonstrate no greater academic achievement, on average, than students in untracked schools (Hoffer 1992; Kerckhoff 1986; Slavin 1988). However, students in high-track classes enjoy a wide range of educational advantages relative to their peers in low-track classes, including access to high-achieving peers, high educator expectations, and rigorous instruction (Carbonaro and Gamoran 2002; Gamoran and Nystrand 1994; Kelly and Carbonaro 2012, Van Houtte 2004). These educational advantages translate to higher levels of educational achievement, greater access to postsecondary education, and higher levels of educational attainment (Attewell and Domina 2008; Long, Conger, and Iatorola 2012). Furthermore, poor students, students whose parents have relatively low levels of educational attainment, and students of color are all less likely to enroll in high-track classes. Accordingly, much of the research literature suggests school tracking practices have negligible average effects on student achievement, but these practices contribute to achievement inequalities by providing relative educational advantages to the already-advantaged students who enroll in high-track classes. (For a more thorough review of this literature, see Gamoran 2010.)

However, the literature is by no means unanimous on tracking's effect on achievement and achievement inequality. From a teacher's point of view, differentiation is a technical response to pedagogical challenges associated with schooling large and heterogeneous student populations (Hallinan 1994b; Rosenbaum 1999). Some forms of tracking might thus help teachers target instruction to students' needs, yielding positive effects for a broad range of students, particularly those with unique needs (e.g., special education students and English-language learners). Consistent with this hypothesis, a handful of studies using experimental and quasiexperimental methods indicate that sorting students into skills-homogeneous classes has positive achievement effects for students across the skills distribution (Betts and Shkolnik 2000; Duflo et al. 2008; Figlio and Page 2002). By contrast, large-scale policy efforts to create more skills-heterogeneous classroom assignments often have unintended negative consequences for high- and low-achieving students alike (Allensworth et al. 2009; Domina, McEachin et al. 2015).

Conceptualizing Tracking

Past research conceptualizes tracking in nuanced ways, but the often-simplistic measurement and operationalization of tracking in most empirical studies may contribute to this literature's mixed and ambiguous findings. Scholars have utilized a variety of measures to operationalize tracking, including principals' reports of school differentiation practices and written school policies related to course assignments (Betts and Shkolnik 2000; Hoffer 1992; Kelly 2007; Kelly and Price 2011), student reports of track assignment (Gamoran and Mare 1989), teacher reports of classroom composition (Argys et al. 1996), and transcript-verified measures of student course assignments (Lucas 1999). In many cases, these measures impose simplistic categorizations on school tracking systems, classifying schools as "tracked" or "untracked" or dividing students between "vocational" and "academic" tracks. This literature demonstrates the importance of tracking for educational achievement and inequality, but it largely fails to address how tracking systems likely differ and the consequences of these differences for student outcomes. For example, Duflo and colleagues (2008) estimate the effects of an isolated change in one dimension of a school tracking system—the degree to which students are grouped by ability into separate classrooms for instruction—but provide little evidence regarding the relations among this change and other dimensions of school tracking systems. Investigating these relations is essential to understand the social organization of schooling and designing effective and equitable instructional practices.

A handful of studies have attempted to operationalize a more nuanced view of school tracking systems. Hallinan (1992, 1994a) documents between-school variation in the number of academic tracks as well as between-school variation in the relationship between placement in high-track courses and student outcomes. Using school course catalogues and assignment policies to measure several dimensions of tracking systems, Kelly and Price (2011) find that schools with wide variation in student skills are most likely to develop highly differentiated academic tracking systems. Lucas (1999) uses student-level data from the nationally representative High School and Beyond (HSB) to measure the flexibility of secondary-school tracking systems, demonstrating that despite the dissolution of an overarching track system, the curricular experiences of U.S. high school students remain highly stratified by race and class. Using the same data, Gamoran (1992) shows that different dimensions of school tracking systems have different consequences for students, demonstrating that the achievement effects of enrolling in high-track courses varies across schools. In particular, Gamoran finds that relatively flexible tracking systems are associated with high levels of mean student achievement and low levels of cross-track achievement inequality.

These studies point to the potential for a more nuanced view of tracking practices in understanding schools' role in the production and reproduction of social inequality. However, this work all faces substantial data limitations. Lacking access to student-level data, Kelly and Price (2011) were unable to test the relationship between tracking systems and student outcomes. Lucas (1999), Gamoran (1992), and other scholars who used NCES cohort-based studies (e.g., the HSB, National Education Longitudinal Study of 1988 [NELS], Education Longitudinal Study of 2002 [ELS], High School Longitudinal Study of 2009 [HSLS]) are limited by the paucity of available contextual data (Argys et al. 1996; Betts and Shkolnik 2000; Figlio and Page 2002; Gamoran 1992; Lucas 1999). These panel studies generally provide detailed data on 20 to 50 students sampled from each of approximately 500 secondary schools. This stratified sampling scheme provides data on a nationally representative sample, but it situates the student as the unit of analysis and includes limited direct data on the emergent institutional structures in which students are situated. In particular, these panel studies provide limited data about the range of courses schools offer, the ways schools sort students across those courses, and the demographic and skill composition of students' classroom peers. Moreover, because these studies collect information on only one cohort of students per school, they do not allow consideration of within-school variation in elements of tracking over time. By contrast, our data provide student-level measures as well as school and classroom indicators for every eighth grader enrolled in 23 schools in three districts over three school years. These data make it possible to observe several highly salient dimensions of school tracking systems that are unobservable in widely utilized nationally representative panel datasets.

DIMENSIONS OF TRACKING

We understand school tracking systems as the culmination of an array of school-level processes related to the provision of differentiated academic course work and the allocation of students among the available courses. We thus measure track structures as school-level variables. Building on Sørensen's (1970) theoretical work, as well as prior efforts to measure dimensions of tracking, we develop a framework for thinking about and measuring school tracking systems. We articulate and measure five conceptually distinct dimensions of school tracking systems, trace the relationship between these dimensions of school tracking systems and student achievement growth, and investigate the extent to which the effects of school tracking systems vary with students' prior achievement. Table 1 provides a summary of the five dimensions of school tracking systems we identify and measure.

Degree of Curricular Differentiation

Sørensen (1970:355) defines organizational differentiation as "the division of a school's student body into subgroups of a permanent character." Some form of organizational differentiation is nearly ubiquitous in contemporary schools. The U.S. public education system sorts children into schools by neighborhood and parental preferences. These schools then sort children by age into grades. Beyond these basic forms of differentiation, schools vary considerably in the degree to which they differentiate curriculum and instruction. Schools may differentiate curriculum and instruction *horizontally* by providing students with various learning environments in which they are exposed to different bodies of knowledge, like when a university offers a wide range of

Table 1. Five Dimensions of Tracking and Strategies to Measure Them Using Educational Administrative Data.

Dimension	Definition	School-level measure
Degree of curricular differentiation	The number of distinct curricular positions in the organization	Number of different courses available to eighth graders (math and ELA)
Classroom skills homogeneity (ability grouping)	Degree to which organizations assign students to different settings based on salient observed characteristics	Seventh-grade test-score homogeneity within eighth-grade classrooms (math and ELA)
Track exclusiveness	Extent to which access to high-status positions is restricted	Percentage of students in low-status courses (math and ELA)
Track stability	Extent to which organizational positions persist over time	Percent of students in same track level in eighth and ninth grade (math and ELA)
Track scope	Extent to which organizational position in one domain predicts position in other domains	Correlation between math and ELA track location

Note: ELA = English language arts.

graduate seminars focusing on distinct topics. In addition, schools may differentiate curriculum and instruction *vertically*, by creating different learning environments that expose students to similar bodies of knowledge but at different paces, with different levels of rigor, or with differing degrees of social status. In our conceptualization, schools that offer students a broad range of classes—whether vertically or horizontally differentiated—display a high degree of differentiation (as measured by the number of course offerings), whereas schools that offer few classes display a low degree of differentiation (Hallinan 1992). All else being equal, one would expect curricular differentiation to have positive consequences for student achievement, because it allows educators to develop subject-matter and skill-level specializations, and students can find classes that match their academic interests and instructional needs. Given that our focus is on tracking in a middle school context, schools are unlikely to use horizontal differentiation.

Classroom Skills Homogeneity (Ability Grouping)

By sorting students across learning environments according to their measured skills, many tracking strategies attempt to simplify the task of instruction. Whereas teachers in skills-heterogeneous (or ungrouped) classrooms may struggle to deliver instruction at the appropriate level for a wide range of students (Rosenbaum 1999), skills-homogenous (or grouped) classrooms may allow teachers to provide instruction more appropriately tailored to their students (Eccles and Roeser 2011). Schools vary in the extent to which their assignment processes generate skills-homogeneous classrooms. Some schools attempt to assign students to courses exclusively on the basis of their prior test scores (Dougherty et al. 2015; Kelly 2009). However, scheduling constraints and limited resources often restrict educators' discretion over students' classroom assignments. Many schools allow teacher recommendations as well as parent and student preferences to influence classroom assignments (Oakes and Guiton 1995; Rickles 2011). As a result, even in otherwise "tracked" schools, students with very different skills levels may sit in the same academic classrooms (Mickelson 2001). Conversely, even in explicitly "untracked" schools, informal pathways may develop that lead students to be grouped based on skills levels across classrooms (Horvath 2015; Watanabe 2006). Building on Sørensen's (1970) notion of "selectivity," we conceptualize the degree to which schools assign students to skills-homogeneous classrooms as a distinct dimension of tracking systems.[1]

This dimension of tracking systems likely has mixed consequences for students. Skills-homogeneous classroom assignments may allow teachers to target their instruction to student skills, but such grouping strategies may broaden skills gaps by exposing high-achieving students to positive peer effects and low-achieving students to negative peer effects (Becker 1987; Epple, Newlon, and Romano 2002; cf. Marsh 1987; Zimmer 2003). Furthermore, skills-homogeneous classroom assignments may create status hierarchies in schools, leading to inequalities in learning opportunities and academic expectations across high- and low-achieving classrooms (Domina, Penner, and Penner 2017; Carbonaro 2005; Dreeben and Barr 1988; Kelly and Carbonaro 2012; Metz 1978; Nystrand and Gamoran 1997; Oakes 1985; Page 1991).

Track Exclusiveness

Over the past several decades, policy makers and educators have undertaken a concerted effort to intensify academic curricula in U.S. schools. Nonetheless, schools vary in the extent to which they expose students to high-level academic content (Domina and Saldana 2012; Penner, Domina et al. 2015; Stein et al. 2011). Some schools enroll all students in courses previously reserved for relatively high-achieving students; other schools allocate relatively advanced or academically rigorous instruction to some students and less advanced and rigorous instruction to others (Domina, Hanselman et al. 2016). We label this dimension of school differentiation systems "track exclusiveness." Our conceptualization of exclusiveness focuses on the relative size of the lower track.

If enrolling students in a more advanced course exposes them to a more rigorous curriculum, one might expect track exclusiveness to limit student achievement. However, track exclusiveness has the potential to positively affect student outcomes. If, for example, inclusive systems expose some students to material beyond their current capabilities, exclusivity could protect overmatched students' learning. Alternatively, if teachers adjust instruction in high-track classes to accommodate underprepared students, decreased exclusivity could undermine high-achieving students' learning (see Cronbach and Snow 1977; Snow 1989). Furthermore, decreases in track exclusivity might depress achievement for students left in low-track classes by creating new stigmas associated with these classes (Gamoran 1992).

Track Stability

School tracking systems also vary in the extent to which they create opportunities for students to move between tracks over time. We describe this dimension of school tracking as "track stability," and we distinguish between schools in which track placements are fairly permanent whereby students have few opportunities to move up or down in a track system and schools in which track placements are relatively fluid over time. Rosenbaum's (1976) classic portrayal of tracking at "Grayton High" provides an example of a "tournament-style" track system, in which few students move from low-track to high-track courses and upward track mobility is thus exceedingly rare. Less rare, however, is downward mobility, or the phenomenon of students moving from high-track to low-track courses. Subsequent analyses suggest this description may not always hold, indicating that some schools provide opportunities for both upward and downward track mobility (Hallinan 1996; Lucas 1999; Lucas and Good 2001; McFarland 2006).

Systems that allow for high degrees of track mobility may be particularly effective at matching students with instruction. If so, exposure to a low-stability track system may boost student achievement. These positive effects may be less common in tournament-style track systems, where upward mobility is rare and downward mobility is common. Tournament mobility systems, however, might boost achievement by facilitating an appropriate match between students and instructional offerings and motivating students. Alternatively, a high degree of tournament mobility might depress student achievement and broaden inequalities by stigmatizing track mobility and associating it with failure.

Track Scope

The tracking system common in U.S. secondary schools throughout the first half of the twentieth century sorted students into vocational, general, and college preparatory tracks, which typically defined students' secondary school curricula. One distinguishing characteristic of this system, as well as the between-school tracking systems common in secondary education in much of Europe and Asia, is that it places students into overarching tracks such that students exposed to high-level instruction in one subject tend to receive high-level instruction in all areas (Hanushek and Wößmann 2006; Lucas 1999). As such, this system has a high degree of "scope." As Lucas (1999) documents, U.S. schools dismantled this overarching track system during the 1960s and 1970s, creating a system that theoretically allows students to take high-track classes in some subjects and low-track classes in others. Lucas's analyses suggest track scope remained high in U.S. high schools through the 1980s, but he shows that track scope varies considerably across schools. Furthermore, Lucas demonstrates that track placements in socioeconomically diverse schools tend to be higher scope than in homogeneous schools. We consider "scope" as a fifth dimension of contemporary tracking systems.

One might expect scope to relate negatively with student achievement, if schools with high degrees of track scope find it difficult to match students with instruction appropriate for their course-specific skills (Hallinan 1994b; Sørensen 1970). High-scope tracking systems may also intensify a tendency toward social closure—or cliquishness—in student peer networks, because they limit students' opportunities to socialize in class with peers outside their academic track (Hallinan and Sørensen 1985). The resulting social processes may increase the extent to which students identify with their academic track position, exacerbating the association between track assignments and achievement inequality.

RELATIONS AMONG DIMENSIONS OF TRACKING

These five dimensions of tracking are conceptually linked, in the sense that each can be understood as a contributing factor to school tracking systems. Consider, for example, a school that offers a highly differentiated curriculum, in which students are grouped into skills-homogeneous classrooms, all but the highest-achieving students are excluded from high-track classes, track placements are highly stable over time, and track scope is high. One would clearly describe such a school as highly tracked. In practice, however, the dimensions of tracking need not vary together across schools or over time. As Oakes (1985:43) observed in her landmark study of tracking in 25 high schools, "tracking in schools is *not* an orderly phenomenon in which practices, even within a single school, are consistent or even reflective of clearly stated school or district policies." Because tracking is a collection of practices that schools accumulate, debate, reject, and reform over time, the same school may rank high on one dimension of tracking and low on another. For example, a school can offer a highly differentiated curriculum composed of a wide array of distinct courses even as it places students into highly-skills-heterogeneous classrooms. Another school could maximize track scope by having students spend the entire school day with the same set of peers even as it minimizes track stability by changing students' location in the track system year after year.

In this article, we view the relations among the dimensions of tracking, as well as the extent to which these dimensions are stable within schools over time, as an empirical question. We measure these dimensions, explore their correlation across schools and over time, and investigate their independent links to student achievement and achievement inequality.

DATA

We operationalize the five dimensions of tracking using administrative panel data gathered from all students enrolled in three Southern California public school districts. These data follow more than 20,000 students enrolled as eighth graders during the 2010–11, 2011–12, and 2012–13 school years in 23 Southern California middle schools from their 7th-grade-through-10th-grade years. We strategically selected these three districts for inclusion in this study because they enroll a diverse group of students, have distinct histories related to mathematics and English language arts (ELA) course placement practices, and were willing to participate in qualitative and quantitative primary data collection. These data include student demographics (gender, race-ethnicity, language status, free/reduced lunch eligibility); students' scores on California Standards Tests (CST) in mathematics and ELA administered in the spring of 7th, 8th, 9th, and 10th grades; transcript information on students' middle and high school math and ELA course assignment and performance; course title, teacher ID, and course period for students' middle and high school courses; and California High School Exit Exam (CAHSEE) scores, which provide a standardized measure of students' math and ELA achievement in the spring of 10th grade. We supplement these data with qualitative data gathered in interviews with administrators from each district and approximately 25 teachers who teach eighth-grade mathematics courses in the three districts.[2]

Table 2 provides a summary of the longitudinal student-level administrative data we collected from our sample schools during the 2009–10, 2010–11, and 2011–12 school years. Our sample is by no means nationally representative; in particular, our sample schools enroll a disproportionately large number of Latinx and Asian American students and a correspondingly small number of white and African American students. However, the sample is racially, ethnically, and economically diverse. Districts A and B, which are among the 10 largest public school districts in California, are situated in inner-ring suburban communities that include middle-class and relatively poor neighborhoods. District C spans an affluent beach community as well as a considerably poorer inland city. The share of students in our sample eligible for the federal Free and Reduced Lunch Program, based on their family incomes, roughly matches the state average (55 percent in 2010–11).[3]

METHODS

These data provide a unique opportunity to develop nuanced measures of school tracking systems. Because our three partner districts provide a census of transcript, achievement, and demographic data for three cohorts of students enrolled as eighth graders in 23 schools, including teacher and period identifiers, we can identify the classrooms in which students took core academic courses and all of their peers in these classrooms. We draw on school and district course listings and academic policy documents, as well as interviews with educators at the school and district levels, to contextualize these transcript and administrative data. Our analyses use these quantitative and qualitative data to measure the degree of curricular differentiation, classroom skills homogeneity, track inclusiveness, track mobility, and track scope for eighth graders in 23 schools across three school years.

We first analyze these measures at the school/year level (N = 69). To explore the extent to which "tracking" as implemented in contemporary schools is a single practice or a collection of independent practices, we estimate a correlation matrix for our measures of the dimensions of tracking. If tracking is best conceptualized as a single institutional practice, the dimensions of tracking should correlate highly across schools and over time. Alternatively, weak correlations among these dimensions suggest tracking may be better conceptualized as a diverse set of structural elements and practices that are realized in different ways across schools and over time.

In this multidimensional conception of tracking, the school-level practices that define the social organization of instruction likely result from time-variant contextually specific technical, political, and cultural factors. As such, it seems likely that different school-level factors predict different dimensions of tracking. To test this notion, we estimate a series of mixed models of the following form:

$$Y_{sdt} = \beta_0 + \beta_1 X_{sdt} + \alpha_t + \alpha_d + u_s + e_{sdt}, \qquad (1)$$

where Y_{sdt} measures the dimensions of organizational differentiation in eighth-grade math and ELA for school s in district d at year t; X_{sdt} is a set of time-varying school-level covariates describing observable characteristics of s at time t, including school enrollment, an index of socioeconomic disadvantage (calculated as the mean of the standardized proportion of African American and Latinx students in the school, the standardized proportion of students who qualify for free and reduced lunch, and the standardized proportion of students who are English-language learners), students' mean prior achievement levels,[4] and dispersion in students' prior achievement; α_t is a vector of year fixed effects; α_d is a district-level fixed effect; u_s represents school-level random effects; and e_{sdt} is the time-varying school-level error term. Although the results of these analyses are limited in external validity, because we can explore the correlates of tracking systems only in the 23 California schools for which we have data, they complement prior research from Kelly and Price (2011) and Oakes (1985) about school-level correlates of tracking systems.

A multidimensional conception of tracking suggests a more nuanced set of answers to historically contentious questions regarding the effects of tracking for student achievement and inequality. If tracking is actually a collection of conceptually and empirically separable practices, it may be possible to develop school structures that realize the potential benefits associated with instructional differentiation while avoiding the costs commonly associated with tracking. To address these

Table 2. Descriptive Statistics, Eighth-Grade Students in Three Southern California Public School Districts, 2009–10 and 2011–12 School Years.

Variable	District A	District B	District C
District administrative information			
Total eighth-grade student enrollment, 2010–12	12,212	7,913	3,714
N traditional schools enrolling eighth graders	9	10	4
N eighth-grade mathematics classrooms[a]	116	103	41
N eighth-grade ELA classrooms	165	80	35
Student demographics (averaged over available cohorts)			
Percentage female	50.6	50.7	47.1
Socioeconomic disadvantage scale (z scored)	0.07	0.03	–0.40
Percentage African American	2.5	0.5	0.9
Percentage Asian	18.1	37.0	6.5
Percentage Latinx	67.1	51.4	44.9
Percentage white	12.3	11.1	47.6
Percentage free or reduced-price lunch	70.7	69.9	50.4
Percentage English-language learners	20.3	28.0	16.9
Percentage reclassified fluent English speakers	43.8	47.1	22.1
Percentage special education	6.2	2.4	12.5
Seventh-grade ELA CST	–0.13 (0.97)	0.18 (0.93)	0.06 (1.15)
Seventh-grade mathematics CST	–0.16 (0.98)	0.22 (0.91)	0.08 (1.13)

Note: ELA = English language arts; CST = California Standards Test.
[a]Classroom counts average over the three study years.

questions, we use student-level data to investigate the relationship between exposure to the dimensions of tracking in 8th grade and students' 10th-grade achievement scores. These models take the following general form:

$$Y_{icsdt} = \gamma_0 + \gamma_1(Tracking_{sdt}) + \gamma_2(X_{icsdt}) + \alpha_t + \alpha_d + u_c + u_s + r_{icsdt} \quad (2)$$

In these analyses, Y_{icsdt} is the 10th-grade math and ELA test scores as measured on the CAHSEE for student i in classroom c, school s, district d, and year t. This exam is administered to all students in the spring of 10th grade. At the time of its administration to the students in our sample, the CAHSEE was a requirement for high school graduation.[5] This high-stakes test was administered in a consistent form throughout the study period to virtually all students regardless of their skill level, postsecondary plans, or course enrollments. Because this test primarily measures skills aligned to sixth-to-eighth-grade standards, it may not fully capture the effects of tracking for high-achieving students, potentially introducing a negative bias in our findings. In supplementary analyses, we test this bias by estimating parallel models predicting eighth- and ninth-grade ELA CST scores.[6] This bias is likely small, as approximately 5 percent of students in the sample scored at ceiling on any one achievement test. X_i is a set of student-level characteristics, including demographics and prior achievement as measured by students' seventh-grade test scores and grade; α_t is a vector of year fixed effects; α_d are district fixed effects; u_c is a class-level random effect; u_s is a school-level random effect; and r_{icsdt} is the residual time-varying student-level error term.[7]

The coefficients of interest in this model, γ_1, represent the relationship between school-by-year measures of the dimensions of tracking and students' achievement, independent of other relevant measures of dimensions of tracking as well as district and year fixed effects and student-level controls. If the dimensions of school tracking systems are unrelated to student characteristics and other school characteristics, conditional on the demographic and lagged achievement measures we control for, these models would generate unbiased estimates of the independent effects of these dimensions of school tracking systems. However, students likely vary across schools on a wide range of unmeasured characteristics that could potentially confound the observed relation between the dimensions of tracking and student achievement.

To address this potential bias, we fit additional models in which we center each tracking measure on its school-level mean. These models, which are equivalent to estimating school fixed-effects models, thus estimate the effect of tracking exclusively for within-school variation in tracking systems. The models generate unbiased estimates of the effects of school tracking systems on students' achievement if the following two assumptions hold: (1) students do not select into school on the basis of cross-year variation in their eighth-grade math and ELA tracking systems, and (2) changes in school tracking systems are not associated with confounding changes in school organization. Although we cannot directly test either of these assumptions, our qualitative data suggest both are plausible. The vast majority of students in our sample attend their middle schools due to neighborhood zoning rather than school choice. Students likely sort into schools and their feeder neighborhoods based on broad reputations, but we believe it is unlikely they sort into schools based on contemporaneous and difficult-to-observe changes in school tracking regimes. Our interviews with school and district leaders suggest that most schools draw on teacher, student, and parent observations and input to continuously modify tracking systems, rather than altering these systems as part of broader systematic school reform efforts. We do not observe, for example, systematic links between leadership change or curriculum adoption in a school and changes in tracking systems.

Finally, to understand the extent to which tracking exacerbates achievement inequalities within schools, we add an interaction between students' seventh-grade test scores and the school-mean-centered version of the school dimension of tracking. Positive values on these interaction terms suggest tracking practices magnify the association between 7th-grade test scores and 10th-grade test scores, as one would expect if tracking increases achievement inequality. For simplicity, we interpret results in terms of the predicted associations between 7th- and 10th-grade achievement under different tracking regimes.

RESULTS

Measuring the Dimensions of Tracking

Based on a review of school course catalogues as well as conversations with educators at sample schools and districts, we categorize eighth-grade math and ELA courses into three levels: advanced, college prep, and remedial. We refer to the middle track as "college prep" because it is designed to prepare students to complete the high school course sequence required for admission to four-year colleges in the University of California and California State University systems. As Figure 1 indicates, schools tend to place relatively-high-achieving students in advanced and honors courses, students at the middle of the test score distribution in college prep courses, and low-achieving students in remedial courses. However, we also find evidence of considerable skills heterogeneity among students in each track.[8]

In this article, we move beyond the broad representation of tracking systems represented in Figure 1 and empirically measure each of the five dimensions of school tracking systems. Because we have access to testing and transcript data in mathematics and ELA for every student in our sample schools, we can identify the title and level of all courses that sample schools offer to eighth graders in these key academic areas. In addition, by identifying students who take the same class with the same teacher during the same school period, we can identify every peer in eighth graders' math and ELA classrooms. These data allow description of schools' tracking systems and students' places in these systems. Table 3 provides a descriptive overview of our measures of the five dimensions of tracking.

We measure the degree of curricular differentiation as the number of different course titles schools make available to eighth graders in any given year. As Table 3 shows, the schools in our sample offered an average of four mathematics classes during the study period. However, schools vary appreciably on this measure. Some schools offered only two distinct eighth-grade mathematics courses (algebra and prealgebra), and others offered as many as seven (including a remedial general mathematics skills course, prealgebra courses in English and Spanish, algebra courses in English and Spanish, an honors algebra course, and a doubly-advanced honors geometry course). Our sample schools offered slightly fewer ELA courses to eighth graders, but we observe similar cross-school variation in eighth-grade ELA course offerings.

We measure the degree of classroom skills homogeneity in schools' eighth-grade math and English classes by using students' eighth-grade

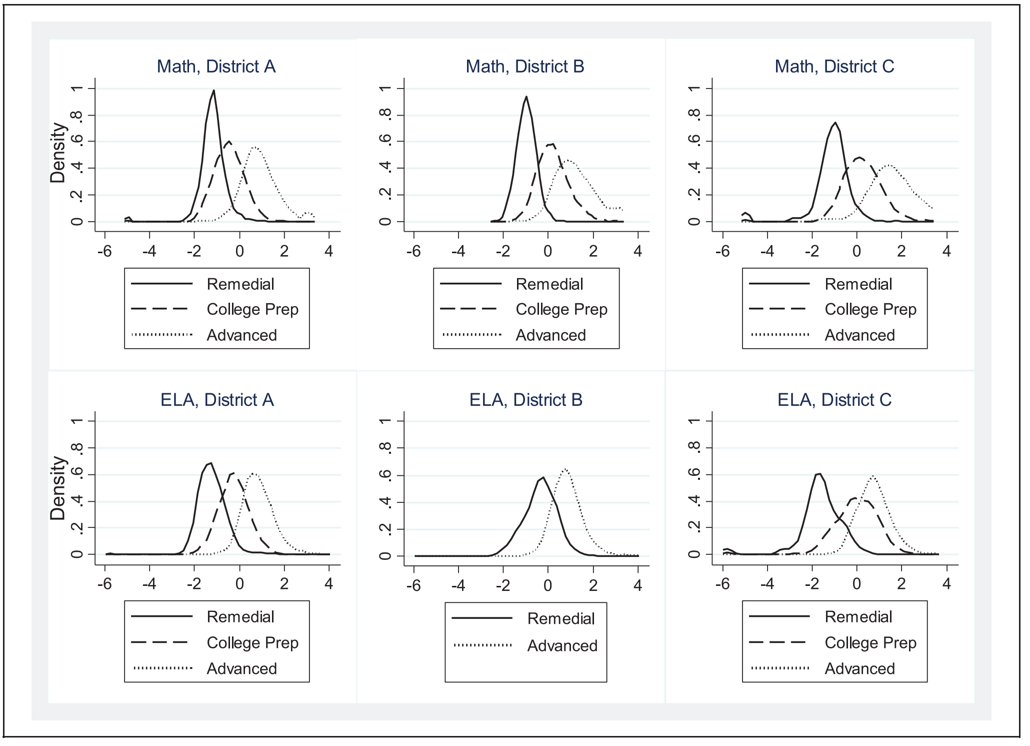

Figure 1. Distribution of standardized seventh-grade achievement scores by eighth-grade course track, math and English language arts in three California school districts, 2010 to 2012.

classroom assignments to predict their 7th-grade standardized test scores within each school and year for which we have data. The intraclass correlation (ICC) from this multilevel model captures the amount of between-class variation that exists within a given school by year based on students' prior achievement. We interpret this ICC as the degree of skills homogeneity in eighth-grade mathematics and English classrooms in a school in a given year on a 0-to-1 scale. For our sample schools, this measure has a mean of 0.52 and a standard deviation of 0.17 for mathematics, and a mean of 0.50 and a standard deviation of 0.18 in ELA.

Conversations with school and district leaders revealed substantial variation in course assignment policies across and within schools over time. Educators articulated a wide range of opinions about tracking. One teacher told us, "I like it that the advanced kids are separated, because then those kids are not in a class where it is constantly a behavior thing happening. . . . I think they've earned that." Another teacher in the same school disagreed, saying, "I don't believe in the honors system. I believe that all kids should be deserving of a high-quality curriculum." The profound variation we observe across and within schools over time reflects this debate and the very different ways schools and districts approached tracking over the study period. During the three years our data cover, District B encouraged schools to enroll students in eighth-grade math and ELA exclusively on the basis of prior test scores. Teachers reported that they occasionally overruled the district's placement formulae, but our analyses indicate that classroom assignments were relatively skills homogeneous in District B over time. However, the test score and grade thresholds District B used as benchmarks to guide course placements changed over the study period. In the study's first year, the district used fairly inclusive standards in an attempt to boost enrollments in accelerated and honors classes. Under these inclusive standards, many students had to repeat the courses they were placed in, so District B revised its placement guidelines. By contrast, Districts A and C gave

Table 3. Descriptive Statistics for Measures of Dimensions of Organizational Differentiation in Three Southern California Public School Districts, 2009–10 and 2011–12 School Years.

Variable	*M*	*SD*	% Variance between schools	% Variance within schools (over time)
Degree of curricular differentiation				
Number of distinct courses				
Math	4.06	1.06	67.7	32.3
ELA	3.26	1.31	56.6	43.4
Classroom skills homogeneity				
Eighth-grade classroom ICC, seventh-grade scores				
Math	0.52	0.17	77.3	22.7
ELA	0.50	0.18	52.0	48.0
Exclusiveness				
Percentage in lower-track courses				
Math	0.15	0.12	60.2	39.8
ELA	0.37	0.27	49.8	50.2
Stability				
Percentage in same eighth- and ninth-grade track				
Math	0.59	0.16	51.2	48.8
ELA	0.66	0.27	49.2	50.8
Scope				
Correlation: Math to ELA track	0.67	0.16	53.3	46.7

Note: ICC = intraclass correlation; ELA = English language arts.

schools relatively little guidance regarding course placements. In District A, schools typically used an informal approach to course assignments, allowing teachers and parents to place students independent of their prior test scores. In this district, placement practices varied considerably across schools and over time, as school personnel changed positions in the school hierarchy and engaged in constant negotiations with parents and students over assignments. Finally, schools in District C experimented with an array of course assignment practices over time, ranging from explicitly skills-heterogeneous course assignments to rigid test score–based assignments.

Figure 2 illustrates our measure of homogeneous classroom assignments, plotting the distribution of seventh-grade mathematics test scores by eighth-grade mathematics classroom for eighth graders in one District C school in 2010 and 2012. During this period, this school moved from an informal course placement system to a system that explicitly attempted to create skills-heterogeneous classrooms in middle-track mathematics. In the process, the school's skills-homogeneity measure decreased from 0.51 to 0.24, a change equivalent to approximately 1.5 standard deviations in the sample-wide distribution. There was considerable overlap across classrooms in the distribution of student achievement in both years. However, in 2010, the bulk of students scored within half of a population-wide standard deviation of their classroom mean. The distribution of scores within classrooms was considerably broader in 2012, especially in the nine middle-track mathematics classrooms where a large proportion of students scored more than a full standard deviation higher or lower than their classroom mean.

Track exclusiveness refers to the extent to which schools assign students to high-track courses. We measure exclusiveness as the proportion of eighth graders enrolled in remedial or college prep courses in our sample schools.[9] Table 3 shows a lower degree of track exclusiveness in mathematics than in ELA in our sample schools. This is likely largely due to California's accountability policy, which pushed schools to boost eighth-grade algebra enrollments. Although the state began to move away from this effort as it transitioned to the Common Core State Standards

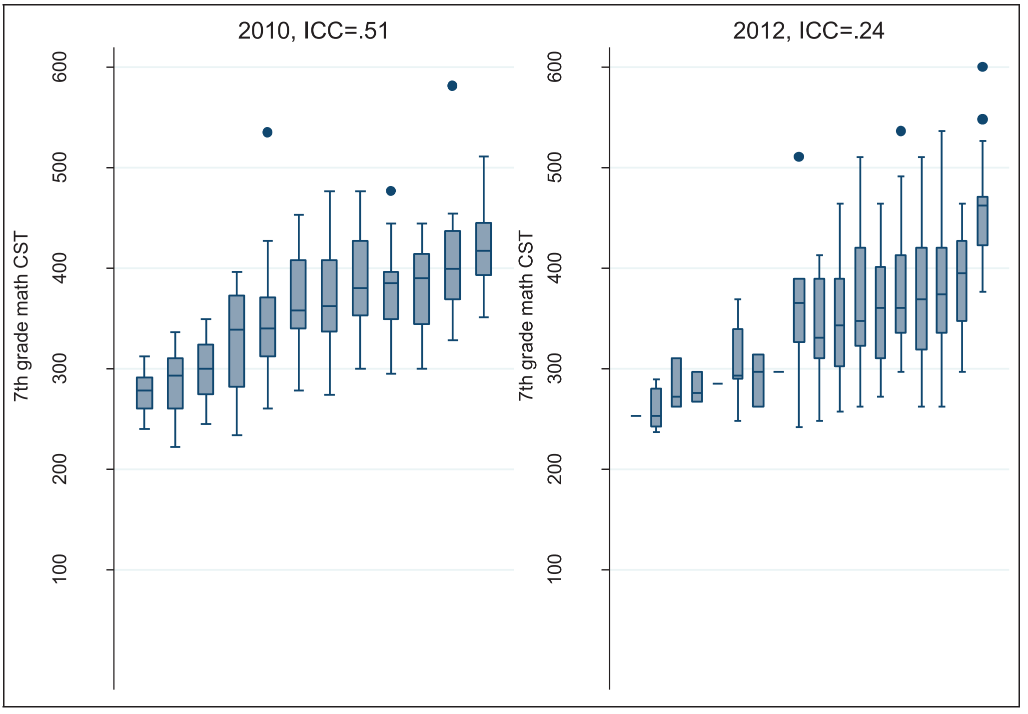

Figure 2. Illustrating classroom skills homogeneity: Distribution of seventh-grade math California Standards Test scores by eighth-grade classrooms in one school, 2010 and 2012.

in math and ELA, California schools continued to enroll students in eighth-grade algebra—a course we consider accelerated because it puts students on a track to complete calculus by the end of 12th grade—at a considerably higher rate than their peers across the United States (Domina, McEachin et al. 2015). The state's algebra-for-all effort limits the degree of variation in math track exclusiveness in our sample schools. However, we observe a large degree of between- and within-school variation in ELA track exclusiveness (the mean is 0.37 and the standard deviation is 0.27).

Track stability refers to the extent to which students remain in the same track level as they progress through school. Because differentiated courses often begin in eighth grade, we operationalize track stability over the ensuing transition to high school. This transition is partly influenced by high school policies, but it is largely a function of the organization of middle school courses. For example, the inclusivity of eighth-grade groupings is closely related to track stability between eighth and ninth grades (see below). Our sample schools enrolled a large proportion of eighth-grade students in advanced math and ELA courses, but these middle school placements do not ensure students will remain on an advanced track through high school. Consistent with Rosenbaum's (1976) observations in "Grayton High," we find that virtually no students in our sample schools moved from eighth-grade remedial classes to ninth-grade college prep classes or from eighth-grade college prep classes to ninth-grade advanced classes. However, 41 percent of students in our sample schools experienced downward mobility in mathematics between eighth and ninth grades, and 34 percent experienced downward mobility in ELA.[10]

We use the proportion of a school's eighth graders who remained in the same track during their ninth-grade year to measure track stability.[11] In interviews, teachers and district leaders reported that they prefer to place students in relatively-high-level middle school courses, so as not to foreclose students' opportunities to take advanced courses later in their educational careers. There is some evidence to suggest that state policy around eighth-grade algebra reinforced this tendency (for more detail, see Domina et al. 2016), leading schools to place students into accelerated eighth-grade algebra classes even as many students retook algebra as ninth graders. As a result, we observe a somewhat higher level of track stability in ELA than in mathematics. However, track stability varies appreciably within and between schools in both mathematics and ELA.

Finally, *track scope* refers to the relation between students' classroom assignments during one part of the school day and their assignments during the rest of the day. Following Lucas (1999), we measure scope as the correlation between a ranked measure of eighth-grade mathematics course placements and a similar ranked measure of eighth-grade ELA course placements. (In these ranked measures, remedial and low-track courses are coded as 1, midtrack courses as 2, and high-track, honors, and accelerated courses as 3.) In schools that approach 1 on this measure, students assigned to high-track mathematics courses are typically also assigned to high-track ELA courses. In schools that approach 0 on this measure, mathematics and ELA course placements are largely unrelated. On average, this measure of scope is fairly high in our sample schools, and students' math course assignments correlate with their ELA course assignments at 0.67. This correlation corresponds closely with Lucas's (1999) findings regarding track scope in a nationally representative sample of U.S. high schools. Underlying this measure, however, we find considerable variation in track scope between schools and temporally within schools. In some schools, students' math track placements rarely diverge from their ELA course placements; in other schools, it is not uncommon for students to enroll in advanced math and college preparatory ELA courses (or vice versa.)

TESTING A MULTIDIMENSIONAL CONCEPTION OF TRACKING

If tracking is a unified practice or set of practices in schools, one might expect schools that score high on one of the five dimensions of tracking to score high on the remaining four dimensions. However, as we noted earlier, these dimensions are conceptually distinct. Furthermore, prior researchers have noted that the school practices underlying tracking systems are often developed and debated within schools in an ad hoc fashion (e.g., Oakes 1985). The correlation matrix reported in Table 4 investigates the extent to which the theoretically separable dimensions of school tracking systems are separable in practice among our 69 school/year observations. We find close associations between our measures of track exclusiveness and track stability. Schools that enroll small proportions of students in advanced courses in eighth grade tend to have more students who remain in the same track location as they move into ninth grade. This correlation is particularly pronounced in ELA, at 0.95.

More generally, however, Table 4 indicates that the correlations among the dimensions of tracking are fairly modest. For example, schools that sort students into relatively skills-homogeneous classes tend to have lower levels of enrollment in advanced math classes and higher levels of track stability, but these associations are fairly small at –0.18 and 0.17, respectively. We find positive associations between the degree of curricular differentiation in schools and the degree of within-classroom skills homogeneity, consistent with the idea that curricular differentiation facilitates the sorting of students into skills-homogeneous classrooms. In both mathematics and ELA, as the number of courses schools offer increases, so too does skills-homogeneous classroom assignments. However, these associations are modest, at 0.36 and 0.44, respectively.

The multilevel models reported in Table 5 indicate that associations between school characteristics and school tracking practices vary across the dimensions of tracking. In these models, both the dependent and independent variables are standardized, so the coefficients can be interpreted as the expected increase in the dimensions of tracking (expressed in standard deviation terms) associated with a one-standard-deviation increase in each independent variable, conditional on all other controls.

The degree of mathematics curricular differentiation appears to vary significantly across districts and over time, but none of our measured school characteristics significantly predict the number of different mathematics courses offered by schools in our sample. Similarly, we find no significant association between school characteristics and skills-homogeneity in mathematics classrooms. Indeed, the only relatively consistently significant school-level predictor of school mathematics tracking systems is schools' total enrollment. In particular, these analyses indicate

Table 4. Correlation of School-Level Measures of Dimensions of Organizational Differentiation in Three Southern California Public School Districts, 2009–10 and 2011–12 School Years.

Variable	No. courses (math)	No. courses (ELA)	Skills homogeneity (math)	Skills homogeneity (ELA)	% in lower tracks (math)	% in lower tracks (ELA)	% in same track grades 8–9 (math)	% in same track grades 8–9 (ELA)	Correlation: Math to ELA
Differentiation									
Number of courses (math)	—								
Number of courses (ELA)	.00	—							
Skills homogeneity									
Skills homogeneity (math)	.36	.08	—						
Skills homogeneity (ELA)	.20	.44	.30	—					
Exclusivity									
Percentage in lower tracks (math)	.42	–.30	.18	.23	—				
Percentage in lower tracks (ELA)	.07	–.16	–.39	.09	.23	—			
Stability									
Percentage in same track grades 8–9 (math)	.45	–.55	.17	–.02	.71	.35	—		
Percentage in same track grades 8–9 (ELA)	.04	–.31	–.40	.01	.25	.95	.46	—	
Scope									
Correlation: Math to ELA	.03	.05	–.12	–.20	–.56	.13	–.16	.17	—

Note: ELA = English language arts.

Table 5. Multilevel Models of School-Level Predictors of Dimensions of Organizational Differentiation Measures for All District A, B, and C Middle Schools, 2010 to 2012.

Variable	No. courses (math)	Skills homogeneity (math)	% in lower tracks (Math)	% in same track 8–9 (math)	No. courses (ELA)	Skills homogeneity (ELA)	% in lower tracks (ELA)	% in same track grades 8–9 (ELA)	Correlation: Math to ELA
Percentage female	–0.09	–0.02	–0.1	–0.03	–0.05	–0.05	–0.01	0.02	0.00
Disadvantage	–0.26	–0.31	0.25	0.10	0.32*	0.14	0.22*	0.42***	–0.63*
Mean CST	–0.33	–0.48	–0.27	0.67***	–0.38*	–0.61**	–0.08	0.50***	–0.23
Enrollment	0.34	0.26	–0.49**	–0.26*	0.18	0.06	0.00	–0.16**	0.69***
Standard deviation CST (math)	1.04	1.52	1.91	0.86	1.67	1.37	–0.49	–0.40	–0.38
Standard deviation CST (ELA)	0.3	1.05	0.78	0.71	–1.29	–0.99	0.01	0.37	0.73
2011	0.34	–0.03	0.26	0.11	–0.16	0.12	–0.06	–0.05	–0.27*
2012	0.71**	0.21	0.74***	0.58***	–0.32	0.23*	–0.18***	–0.14***	0.02
District B	0.81*	0.19	0.18	–0.06	–0.37	0.58	1.69***	0.97***	1.33*
District C	–0.15	0.09	0.01	0.06	–1.12***	–0.61	–0.34	–0.60***	–0.88*
Constant	–1.89	–2.47*	2.81**	–1.66*	0.17	–0.6	–0.25	–0.24	–0.73
N	69	69	69	68	69	69	69	66	69

Note: School-year-level data, with school-level random effects. Outcomes as well as percentage female, disadvantage, mean CST, and enrollment are z scored.
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

that relatively large schools tend to have less exclusive enrollments in advanced eighth-grade math courses. Furthermore, students in large schools tend to experience relatively low rates of track stability between eighth and ninth grades. Because the relatively small school-level sample size limits the power in these analyses, several nonsignificant conditional associations between the concentration of socioeconomically disadvantaged students and mathematics tracking dimensions are worth noting. These nonsignificant associations indicate that schools that educate relatively large proportions of poor, minority, and English as a Second Language students tend to offer fewer mathematics courses and place students in relatively heterogeneous mathematics classes.

The pattern of school-level predictors of the dimensions of tracking in ELA is somewhat different. Schools with relatively disadvantaged student populations tend to offer significantly *more* eighth-grade ELA courses than do more advantaged schools, net of controls. The concentration of socioeconomically disadvantaged students is also a significant predictor of track exclusiveness and track stability in ELA. Consistent with mathematics, school-mean prior achievement relates negatively with curricular differentiation and classroom skills homogeneity but correlates positively with track stability in ELA. Finally, higher school enrollment is negatively associated with track stability in ELA.

DIMENSIONS OF TRACKING AND STUDENT ACHIEVEMENT

In light of the evidence suggesting that the dimensions of tracking are empirically separable, the remaining analyses examine the links between these dimensions and student achievement. Table 6 reports results of a series of multilevel models regressing students' 10th-grade math achievement on the dimensions of mathematics tracking systems in students' 8th-grade middle schools. Table 7 reports results of parallel analyses in ELA. All dependent and independent variables are standardized in each model in both tables, such that each has a mean of 0 and a standard deviation of 1 in the student population under analysis.

The first model in Table 6 provides a fully unconditional look at these relationships; the second model investigates the relationship between each dimension and math achievement while controlling for each of the other dimensions. The third model adds student-level demographic and prior achievement controls as well as indicator variables that account for commonalities among students enrolled in the same school district (district fixed effects) and in the same grade cohort (cohort fixed effects). Finally, in the fourth model, we mean-center the time-varying school-level measures of the dimensions of tracking around schools' three-year mean scores on these measures. Doing so controls for time-invariant school characteristics that may confound the link between the dimensions of school tracking regimes and student achievement.[12]

The first model of Table 6 indicates that the unconditional relationship between the dimensions of tracking and student math achievement varies across dimensions. Mean 10th-grade math achievement is slightly higher for students from schools that offer many math 8th-grade courses than for students from schools that offer few. Similarly, math achievement is slightly higher for students from schools that maintain relatively exclusive access to high-track mathematics courses. Neither of these associations is significant after controlling for the other dimensions of tracking, and both remain nonsignificant with the addition of demographic controls and school-mean centering. The models reported in Table 6 further suggest there is no significant association between 8th-grade mathematics track stability and 8th-grade track scope and students' 10th-grade mathematics achievement.

In contrast, students in schools that have relatively skills-homogeneous 8th-grade math classroom assignments score significantly less well in 10th-grade courses than do their peers in schools where math courses are less rigidly grouped by student achievement. This association holds after controlling for other dimensions of tracking in Model 2, and it remains statistically significant in the most stringent model with school-mean centering (Model 4). Accordingly, we interpret these analyses as suggesting that homogeneous math course assignments may have small negative average effects on students' mathematics achievement.

Model 1 in Table 6 suggests there is a weak positive association between 8th-grade school math track exclusiveness and students' 10th-grade mathematics achievement. This relationship is nonsignificant when we control for the other dimensions of tracking and demographic controls

Table 6. Selected Coefficients from Multilevel Models of the Relationship between Dimensions of 8th-Grade School Tracking System and 10th-Grade Mathematics Achievement, for Students in District A, B, and C Middle Schools, 2010 to 2012.

Variable	Model 1 (fully unconditional)	Model 2 (dimensions only)	Model 3 (controls)	Model 4 (controls, school-mean centered)	Model 5 (controls, school-mean centered, interaction Homogeneity × Seventh-Grade Scores)
Differentiation (number of courses)	.03***	.01	–.01	–.01	–.00
	(.01)	(.01)	(.01)	(.01)	(.01)
Classroom skills homogeneity	–.03***	–.05***	–.03	–.02*	–.02**
	(.01)	(.01)	(.01)	(.01)	(.01)
Exclusivity (percentage in lower tracks)	.04***	.04	.00	.11*	.06**
	(.01)	(.02)	(.01)	(.05)	(.02)
Stability (percentage in same track grades 8–9)	.04	.00	.02	.04	.00
	(.01)	(.03)	(.02)	(.02)	(.02)
Scope (correlation: math to ELA)	.02	.05	.02	–.01	–.00
	(.01)	(.02)	(.01)	(.01)	(.01)
Interaction: Classroom Skill Homogeneity × Seventh-Grade Scores	—	—	—	—	.02***
	—	—	—	—	(.01)
Demographic controls	No	No	Yes	Yes	Yes
Prior achievement	No	No	Yes	Yes	Yes
School-mean centered	No	No	No	Yes	Yes
District FE	No	No	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes	Yes
School RE	Yes	Yes	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes	Yes	Yes
N	20,545	20,545	20,545	20,545	20,545

Note: All independent variables and the dependent variable in these models are *z*-score standardized. Demographic controls include indicators for students' gender, race-ethnicity, free or reduced-price lunch status, English-language-learner status, and special education enrollment. Prior achievement controls are continuous measures of students' seventh-grade math and ELA test scores.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table 7. Selected Coefficients from Multilevel Models of the Relationship between Dimensions of 8th-Grade School Tracking System and 10th-Grade ELA achievement, for Students in District A, B, and C Middle Schools, 2010 to 2012.

Variable	Model 1 (fully unconditional)	Model 2 (dimensions only)	Model 3 (controls)	Model 4 (controls, school-mean centered)	Model 5 (controls, school-mean centered, interaction Homogeneity × Seventh-Grade Scores)
Differentiation (number of courses)	–.01	–.02	.00	.02	.01
	(.01)	(.01)	(.01)	(.01)	(.01)
Classroom skills homogeneity	.00	.00	.02	.00	.01
	(.02)	(.02)	(.01)	(.01)	(.01)
Exclusivity (percentage in lower tracks)	.09***	–.03	–.10***	.04	.05
	(.02)	(.06)	(.02)	(.03)	(.03)
Stability (percentage in same track grades 8–9)	.12***	.15**	.11**	.00	.02
	(.03)	(.06)	(.03)	(.05)	(.05)
Scope (correlation: math to ELA)	.03	.02	.01	.02	.02
	(.02)	(.02)	(.01)	(.01)	(.01)
Interaction: Classroom Skill Homogeneity × Seventh-Grade Scores	—	—	—	—	.06***
	—	—	—	—	(.01)
Demographic controls	No	No	Yes	Yes	Yes
Prior achievement	No	No	Yes	Yes	Yes
School-mean centered	No	No	No	Yes	Yes
District FE	No	No	Yes	Yes	Yes
Cohort FE	No	No	Yes	Yes	Yes
School RE	Yes	Yes	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes	Yes	Yes
N	20,545	20,545	20,545	20,545	

Note: All independent variables and the dependent variable in these models are z-score standardized. Demographic controls include indicators for students' gender, race-ethnicity, free or reduced-price lunch status, English-language-learner status, and special education enrollment. Prior achievement controls are continuous measures of students' seventh-grade math and ELA test scores.

$^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

in Models 2 and 3. After controlling for time-invariant school characteristics in Model 4, we find evidence to suggest that attending middle schools with highly exclusive 8th-grade math tracking systems significantly *increases* student achievement by the 10th grade. This estimate suggests that, all else equal, a student who moves from a school that enrolls 27 percent of students in advanced math classes to a school that enrolls 15 percent in advanced math classes will experience a test score increase of approximately one-tenth of a standard deviation. Supplementary analyses, available by request, confirm that these results hold when investigating changes in track exclusiveness without controlling for other dimensions of school tracking systems. Although somewhat counterintuitive, these findings are consistent with evidence elsewhere in the research literature suggesting that efforts to intensify middle school mathematics curricula may have unintended negative consequences for students' achievement (Clotfelter, Ladd, and Vigdor 2014; Domina et al. 2015; Penner et al. 2015).

Table 7 reports parallel models exploring the relation between dimensions of school ELA tracking systems and students' 10th-grade ELA achievement. In general, the results indicate that ELA test scores are less sensitive than math scores to the dimensions of tracking. We find that 8th-grade ELA exclusiveness and stability are both associated with 10th-grade ELA scores. Model 3 indicates that students in schools with more exclusive ELA track placement systems demonstrate significantly lower 10th-grade ELA scores than do demographically similar students in schools that allow more access to honors and other high-track English classes. At the same time, students in schools with a high degree of track stability between eighth and ninth grades score significantly higher than students who face greater prospects of downward track mobility. However, Model 4 indicates that neither of these associations is significant after controlling for student characteristics and time-invariant school characteristics, suggesting that these relationships are largely a function of unmeasured school effects.

Taken together, the results in Tables 6 and 7 suggest the constellation of practices researchers often refer to as "tracking" has mixed and modest average effects on student achievement. We find that placing students into skills-homogeneous 8th-grade mathematics classrooms has a small negative effect on students' 10th-grade mathematics achievement. However, our findings regarding the average effects of track exclusiveness suggest that efforts to detrack mathematics instruction by enrolling all students in accelerated courses may have unintended negative consequences. We find no evidence to suggest that any dimensions of 8th-grade ELA tracking systems influence student achievement on the 10th-grade ELA California high school exit exam.

SENSITIVITY ANALYSES

One possible explanation for the weak link between school tracking practices and mean student achievement is that the dependent variable in these analyses—10th-grade exit exam scores—is measured two years after students experience their 8th-grade learning environments.[13] Although the structure of state testing in California schools precludes intermediary analyses in mathematics, we can analyze the relation between the dimensions of school tracking systems in eighth grade and achievement on standardized tests in ELA administered in the spring of students' eighth-grade year. The results of these analyses, reported in Appendix Table A1, provide limited evidence of a short-term effect for the dimensions of tracking. In particular, our school-mean-centered analysis reported in Model 4 suggests that attending a school in which students are assigned to skills-homogeneous eighth-grade ELA classrooms may slightly but significantly decrease eighth-grade achievement. However, consistent with the 10th-grade ELA findings reported in Table 6, we find no robust evidence of a link between the other dimensions of tracking and students' eighth-grade ELA scores.[14]

A related empirical concern is that students' disparate high school experiences could attenuate the observed relation between middle school tracking systems and 10th-grade achievement. Each middle school for which we have data enrolls seventh and eighth graders exclusively. All but one of these schools explicitly serve as a feeder for a single high school. (The exception is a middle school that feeds into two neighboring high schools.) In most cases, our focal middle schools occupy a separate building on a shared or a closely neighboring space with their linked high school. As a result, over 80 percent of students enroll in the same high school as the bulk of their middle school classmates. Appendix Table A2 reports

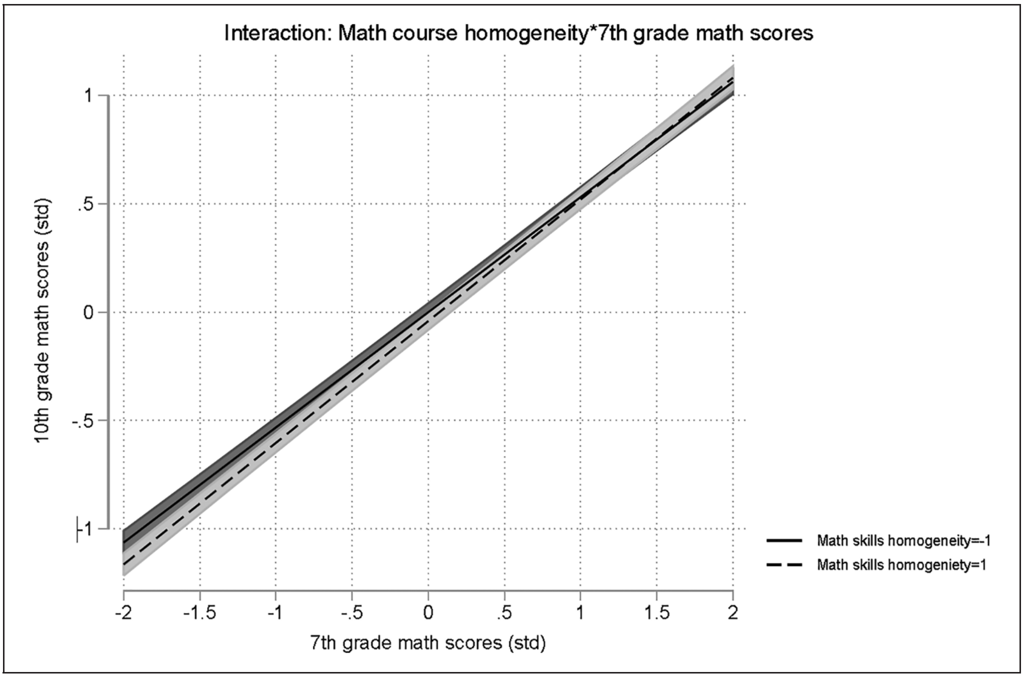

Figure 3. Predicted 10th-grade mathematics achievement scores for students in schools with high and low levels of skills-homogeneous assignment in 8th-grade mathematics classrooms.

a series of supplementary analyses in which we estimate the link between the dimensions of 8th-grade tracking and 10th-grade achievement separately for students who follow their middle school's modal high school path and for students who move out of the modal path. The results of these analyses are strikingly similar to one another and to the results reported in Tables 6 and 7.

DO HIGH- AND LOW-ACHIEVING STUDENTS EXPERIENCE TRACKING DIFFERENTLY?

The analyses reported in the first four models in Tables 6 and 7 focus on the mean effects of school-level tracking systems, so they neglect crucial questions regarding the relationship between tracking systems and achievement inequality. Model 5 in Tables 6 and 7 addresses the equity effects of tracking by taking a closer look at one key dimension of school tracking systems—the degree to which schools group students into classrooms based on their prior test scores. In these models, we investigate the extent to which the effects of school-level ability grouping vary with students' seventh-grade test scores. To ease interpretation, we also report these interactions in Figures 3 and 4.

The *y*-axis in Figure 3 represents students' *z*-scored predicted 10th-grade mathematics achievement scores, and the *x*-axis represents students' *z*-scored 7th-grade mathematics scores. The dashed line represents the predicted relation between 7th-grade achievement and 10th-grade achievement in mathematics for students in schools that implemented ability grouping to an above-average degree in 8th-grade mathematics classrooms. The solid line represents the same relation in schools that implemented a below-average degree of ability grouping in eighth-grade mathematics classrooms. The shaded areas around both lines represent 95 percent confidence intervals. Consistent with the results indicating a negative average effect of classroom skills homogeneity in eighth-grade mathematics (see Table 6), the dashed line is lower than the solid line across the seventh-grade math test score distribution in Figure 3. Notably, however, the disadvantage associated with attending a school in which eighth-grade mathematics courses are largely skills homogeneous is particularly pronounced for students at

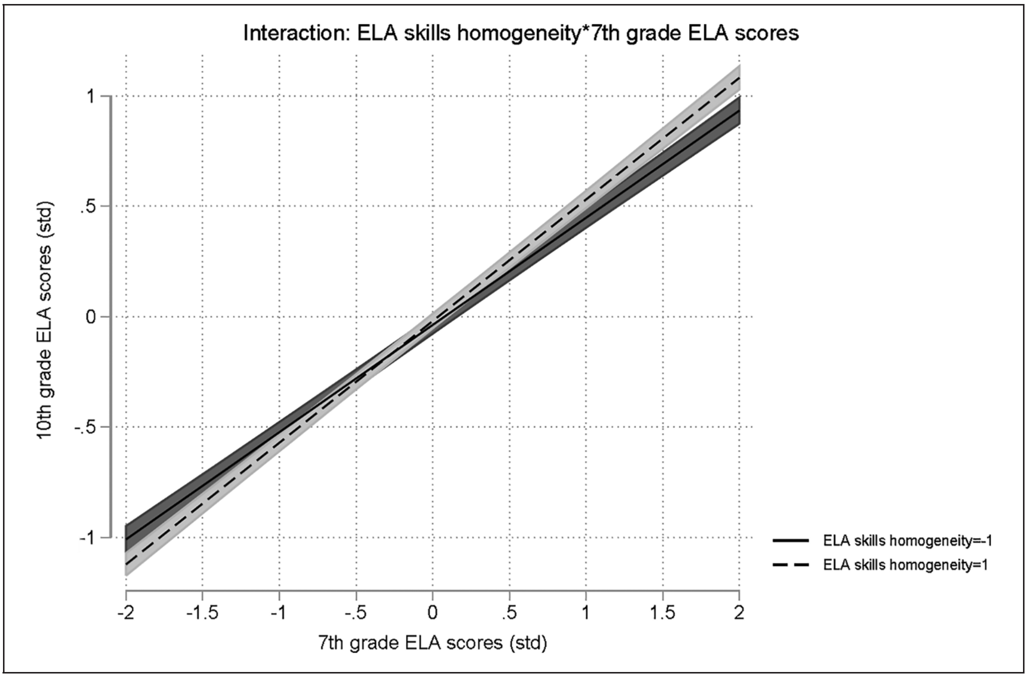

Figure 4. Predicted 10th-grade English language arts (ELA) achievement scores for students in schools with high and low levels of skills-homogeneous assignment in eighth-grade ELA classrooms.

the bottom of the seventh-grade mathematics test score distribution. Put differently, this figure suggests that low-achieving students disproportionately bear the achievement costs associated with ability grouping in middle school mathematics. Model 5 in Table 6 indicates that this interaction term is statistically significant, if small in magnitude. All else equal, this model suggests that enrolling in a school with a high degree of ability grouping will increase the gap between students who come into eighth grade 1 standard deviation above and below the math test score average by approximately 0.08 standard deviations. The models reported in Figure 4 and Model 5 in Table 7 indicate that these inequitable effects are more pronounced in ELA.[15]

DISCUSSION

This study measures multiple dimensions of tracking and identifies their effects on student achievement. Building on the work of Sørensen (1970) and others (Becker 1987; Gamoran 1992; Kelly 2007; Lucas 1999), we identify five theoretically distinct dimensions of school math and ELA tracking systems: (1) the degree of curricular differentiation, (2) the degree of skills homogeneity in classrooms, (3) track exclusiveness, (4) track stability, and (5) track scope. We take advantage of a unique set of student-level administrative data from three public school districts, as well as qualitative data gathered from administrators and educators in those districts, to measure the dimensions of tracking systems in 23 schools and the ways those systems changed over the course of three years.

In these 23 schools, tracking was a multidimensional phenomenon. We found considerable variation on each dimension between our sample of 23 middle schools and within these schools over time. In addition, the dimensions of school tracking systems did not correlate highly with one another. Consistent with a multidimensional conception of school tracking practices, our analyses indicate that the predictors of school tracking systems vary across the dimensions of tracking. These findings resonate with a long tradition of case study research on school tracking systems (Dreeban and Barr 1988; Metz 2003; Oakes 1985; Rickles 2011; Rosenbaum 1999; Watanabe 2006; Wells and Oakes 1996), but they also suggest that many recent quantitative studies considering individual track locations likely obscure important organizational variation in track practices and their consequences.

Our investigation of the dimensions of tracking reveals new insights into how school tracking systems influence student achievement, at least in the 23 schools for which we have data. In ELA, our findings are largely consistent with earlier sociological research on the effects of school tracking. Although we find little evidence to suggest that the five dimensions of school tracking systems have an effect on student achievement, this null effect conceals statistically significant if substantively modest inequality-producing consequences of tracking systems. In particular, we find that when schools group students into ELA classes based on prior achievement, high-achieving students tend to experience more rapid test score growth in ELA than do low-achieving students (see Argys et al. 1996; Gamoran and Mare 1989).

Furthermore, the dimensions of tracking have potentially crosscutting effects on students' mathematics achievement growth. Our results suggest that students experience lower levels of mathematical achievement growth in schools that place students into skills-homogeneous mathematics classrooms, and this negative effect is particularly pronounced for low-achieving students. This finding suggests that students may benefit from placement in relatively skills-heterogeneous classrooms for secondary mathematics instruction. Based on this finding, one might be tempted to endorse recent policy efforts that attempt to expose all students to high-quality instruction and high-achieving peers by universalizing accelerated course placements. However, our analyses also reveal positive effects of math track exclusiveness and stability on students' mathematics achievement. Consistent with several recent policy analyses (Allensworth et al. 2009; Clotfelter et al. 2015; Domina et al. 2015; Penner et al. 2015; Stein et al. 2011), these findings suggest that efforts to detrack instruction by enrolling more students in accelerated courses can have negative effects. From a practitioner's perspective, our findings point to a tension between the benefits of skills-heterogeneous learning environments and the shortcomings of instruction that is insensitive to students' skills. Curricular reform efforts that simultaneously provide disadvantaged students with access to higher-achieving peers and sufficient skill-building opportunities are one promising strategy for resolving this tension (Nomi and Allensworth 2012; Nomi and Raudenbush 2016).

In interpreting these findings, it is important to note that our data can speak directly only to the experiences of approximately 20,000 students enrolled in 23 middle schools in three Southern California public school districts during the 2010–11 to 2012–13 school years. Although we have no reason to suspect these middle schools are idiosyncratic, we make no strong claims regarding our findings' generalizability. Future research should investigate the dimensions of tracking in other settings. Doing so is data intensive, because many of our measures of the dimensions of school tracking systems require at least two years of longitudinal data on all students enrolled in a school as well as indicators for students' classroom placements and the title or content of the classes in which students enroll. However, the growing availability of student-level administrative data in U.S. public K–12 schooling, combined with new approaches to machine-learning and text-based data analysis, provide new opportunities for investigating the dimensions of tracking on a larger scale. In particular, future research should explore and refine track stability in secondary school. Our analyses indicate that track stability matters, but because we do not observe students after 10th grade, we provide an admittedly limited view of stability. In addition to attempting to replicate this study's findings about the dimensions of tracking and their links to student achievement, larger-scale research should investigate the link between school resources and organizational characteristics and the development of school tracking systems.

Future researchers should also examine other dimensions of school tracking systems. Our measures capture just a few of the important ways tracking systems may vary. For example, our analyses do not consider the extent to which student ascriptive characteristics, such as race, class, and gender, align with course placements and the consequences of such alignment for identity formation and ascriptive inequalities in school (Frank et al. 2008; Legewie and DiPrete 2012; Lewis and Diamond 2015; Tyson 2011). Our data also provide limited information on the extent to which schools allow students to choose their own track locations (i.e., track electivity) (Lucas 1999; Powell et al. 1985; Sørensen 1970). Relatedly, due to the age of our sample, we were not able to separate horizontal from vertical differentiation, as the former does not occur widely in middle schools. Future work may be able to test these factors in other K–12 contexts, like high school science. Similarly, our analyses provide no information about the sorts of between-school tracking that occur when school admissions are selective or school enrollments are otherwise unequal (cf. Hanushek and Wößmann 2006; Van de Werfhorst and Mijs 2010).

Despite these limitations, this article suggests a set of strategies for incorporating into the quantitative study of educational inequality a key insight that emerges from the long history of qualitative and case-study research on school tracking systems. Much quantitative research on tracking—indeed, much research in the sociology of education and inequality—takes an individualistic approach, focusing on the consequences of students' track locations. By contrast, our analyses challenge sociologists and educators to consider a broad array of school organizational practices and the ways they cohere into a wide variety of educational tracking systems. Careful study of this organizational variation can help elucidate and evaluate the complex and interacting mechanisms through which schools produce, reproduce, and even ameliorate social inequality. Ultimately, such an approach may point to promising strategies for building more effective and equitable organizations.

Table A1. Selected Coefficients from Multilevel Models of the Relationship between Dimensions of Eighth-Grade School Tracking System and Eighth-Grade ELA Achievement, for Students in District A, B, and C Middle Schools, 2010 to 2012.

Variable	Model 1 (fully unconditional)	Model 2 (dimensions only)	Model 3 (controls)	Model 4 (controls, school-mean centered)
Number of courses	–.05***	–.05***	–.01	0
	(.01)	(.01)	(.01)	(.01)
Classroom skills homogeneity	–.02	.00	–.04***	–.06***
	(.02)	(.02)	(.01)	(.01)
Percentage in lower tracks	–.06*	–.23***	–.07**	.03
	(.02)	(.05)	(.02)	(.03)
Percentage in same track grades 8–9	.14***	.38***	.08*	–.06
	(.02)	(.05)	(.03)	(.04)
Correlation: Math to ELA	.00	–.03	–.03**	–.01
	(.01)	(.01)	(.01)	(.01)
Demographic controls	No	No	Yes	Yes
Prior achievement	No	No	Yes	Yes
School-mean centered	No	No	No	Yes
District FE	No	No	Yes	Yes
Cohort FE	No	No	Yes	Yes
School RE	Yes	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes	Yes
N	20,771	20,771	20,771	20,771

Note: All independent variables and the dependent variable in these models are z-score standardized. Demographic controls include indicators for students' gender, race-ethnicity, free or reduced-price lunch status, English-language-learner status, and special education enrollment. Prior achievement controls are continuous measures of students' seventh-grade math and ELA test scores.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Table A2. Selected Coefficients from School-Mean Centered Multilevel Models of the Relationship between Dimensions of Eighth-Grade School Tracking System and Eighth-Grade Math and ELA Achievement, for Students in District A, B, and C Middle Schools Who Moved to Modal and Nonmodal High Schools, 2010 to 2012.

	Outcome: 10th-grade math score		Outcome: 10th-grade ELA score	
Variable	Students who enrolled in modal high school	Students who enrolled in nonmodal high school	Students who enrolled in modal high school	Students who enrolled in nonmodal high school
Number of courses	–.01	.01	.01	.05
	(.01)	(.03)	(.01)	(.03)
Classroom skills homogeneity	–.03**	.00	–.01	–.02
	(.01)	(.03)	(.01)	(.04)
Percentage in lower tracks	.10***	.07	.08*	.03
	(.03)	(.07)	(.04)	(.10)
Percentage in same track grades	.04*	–.03	.04	.08
8–9	(.02)	(.05)	(.05)	(.14)
Correlation: Math to ELA	.00	–.01	.02	.03
	(.01)	(.03)	(.01)	(.03)
Demographic controls	Yes	Yes	Yes	Yes
Prior achievement	Yes	Yes	Yes	Yes
School-mean centered	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
School RE	Yes	Yes	Yes	Yes
Classroom RE	Yes	Yes	Yes	Yes
N	16,914	3,210	16,914	3,210

Note: All independent variables and the dependent variable in these models are z-score standardized. Demographic controls include indicators for students' gender, race/ethnicity, free or reduced-price lunch status, English-language learner status, and special education enrollment. Prior achievement controls are continuous measures of students' 7th-grade math and ELA test scores. ELA = English language arts; FE = fixed effects; RE = random effects.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

ACKNOWLEDGMENTS

Research reported in this publication was supported by Spencer Foundation (Grant No. 201400180). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Spencer Foundation.

DECLARATION OF CONFLICTING INTEREST

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Andrew McEachin prepared the work as an employee of RAND Corporation.

NOTES

1. Indeed, our conception of skills homogeneity is nearly identical to Sørensen's (1970:363) notion of selectivity, which he defines as "the amount of homogeneity that educational authorities intend to produce by the assignment, in terms of the index of learning used."
2. We conducted 15 formal/informal interviews and two focus groups with district administrators, math coaches, and teachers during the 2014–15 school year in the three districts that are part of this study. During the summer of 2014, the team observed six professional development sessions for the three districts and a pilot-testing effort at one school district with the director of assessments. Interviews

consisted of open-ended questions on the district policies and challenges pertaining to student testing, student course placement, curriculum changes, and the implementation of new state standards. Observations were done naturalistically and recorded using field notes. Some focus groups and interviews were audio recorded and transcribed; others were recorded manually.

3. See http://www.cde.ca.gov/.
4. Standardized percentage black or Latinx correlates with standardized percentage free or reduced lunch at 0.86 and standardized percentage English-language learner at 0.74. Standardized percentage free or reduced lunch correlates with standardized percentage English-language learner at 0.86. School-mean test-score measures correlate at 0.96. Because school-level standard deviations in math and ELA test scores correlate less closely (0.61), we enter these variables separately into the models.
5. California has since reversed course on the requirement that students pass the California High School Exit Exam (CAHSEE).
6. We are unable to estimate similar models using mathematics California Standards Test scores, because beginning with algebra, California middle and high school students take different mathematics tests based on their math course enrollment.
7. In Equations 1 and 2, the fixed-effects terms represented by α_t and α_d absorb all time-invariant district-level variation in the outcomes. By contrast, the school-level random-effects term u_s (as well as the course-level random-effects term u_c in Equation 2) simply accounts for the nonindependence of repeated observations of schools across time.
8. We compared observed track placements to hypothetical course assignments based solely on prior content-relevant achievement, holding course enrollments constant. Observed placements correlated with these hypothetical assignments at approximately 0.7 in eighth-grade mathematics and English language arts (ELA).
9. Our data include text variables describing the titles of each mathematics and ELA course students enrolled in during their eighth-grade year. We sort these courses into three levels: remedial, college prep, and advanced. In most cases, these categories were self-evident based on course titles. Courses labeled "remedial" or "developmental" clearly belonged in the lowest level, and courses labeled "advanced" or "honors" clearly belonged in the highest level. However, in all cases, we triangulated our course labels in conversation with district-level mathematics and ELA curricular specialists and school-level department chairs.
10. Many educators we interviewed expressed frustration at the lack of upward track mobility for their students. Curricular planners in Districts A and C attempted to facilitate upward mobility by creating multiple "course acceleration" opportunities, including double-dose and summer courses. However, these efforts were not in place in sample schools during the study period.
11. Most downward mobility in mathematics occurred when students took algebra in the eighth grade and retook it in the ninth grade. In ELA, the most common example of downward mobility was from an eighth-grade honors course to a ninth-grade college prep course.
12. This is a correlated random-effects model. All models include school- and classroom-level random-effects terms to adjust standard error estimates for the clustering of students in schools and classrooms.
13. However, as noted earlier, the CAHSEE, especially in math, is a function of sixth- to 8-eighth-grade standards. Therefore, the relationship between 8th-grade tracking practices and 10th-grade achievement is relevant, because these practices occurred at the same time students were largely exposed to the material tested on the CAHSEE.
14. Our operationalization of stability (organizational patterns in the transition to high school) requires special attention for the interpretation of this coefficient in models with eighth-grade achievement as an outcome. From the perspective of eighth-grade experiences, this stability reflects the structural prospects afforded by curricular differentiation practices in middle school. Results from initial models suggest that arrangements with greater prospects of (downward) mobility are negative, but this association does not hold when accounting for time-invariant school characteristics.
15. Because these analyses are available only for students who remained in school through the 10th grade, they understate the degree to which students' prior achievement moderates the effects of school tracking regimes. Approximately 20 percent of students for whom we have tracking measures in the 8th grade leave our sample before taking the 10th-grade exit exam. Although our data do not allow us to differentiate between students who move to other districts or private schools and those who drop out of school, we suspect high school dropout is an important source of attrition. Consistent with that explanation, students in the bottom quartile on seventh-grade achievement are approximately twice as likely to leave the sample as their peers in the top quartile.

RESEARCH ETHICS

All research on human subjects has been IRB approved has therefore been performed in a way that is consistent with the ethical standards articulated in the 1964 Declaration of Helsinki and its subsequent amendments and Section 12 ("Informed Consent") of the ASA's Code of Ethics. All subjects that participated in the qualitative portion of the study gave their informed consent prior to their participation in the research. Since the administrative data we draw upon were collected in the process of routine educational practice and analyzed to support educational delivery in accordance with FERPA, students who contributed administrative data did not need to provide consent to participate.

REFERENCES

Allensworth, Elaine, Takako Nomi, Nicholas Montgomery, and Valerie E. Lee. 2009. "College Preparatory Curriculum for All: Academic Consequences of Requiring Algebra and English I for Ninth Graders in Chicago" *Educational Evaluation and Policy Analysis* 31(4):367–91.

Argys, Laura M., Daniel I. Rees, and Dominic J. Brewer. 1996. "Detracking America's Schools: Equity at Zero Cost?" *Journal of Policy analysis and Management* 15(4):623–45.

Attewell, P., & Domina, T. (2008). "Raising the Bar: Curricular Intensity and Academic Performance." *Educational Evaluation and Policy Analysis*, 30(1): 51–71.

Becker, Henry Jay. 1987. "Addressing the Needs of Different Groups of Early Adolescents: Effects of Varying School and Classroom Organizational Practices on Students from Different Social Backgrounds and Abilities." Report No. 16. Baltimore: Johns Hopkins University Center for Research on Elementary and Middle Schools.

Betts, Julian R., and Jamie L. Shkolnik. 2000. "The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools." *Economics of Education Review* 19(1):1–15.

Buchmann, Claudia, and Hyunjoon Park. 2009. "Stratification and the Formation of Expectations in Highly Differentiated Educational Systems." *Research in Social Stratification and Mobility* 27(4):245–67.

Carbonaro, William. 2005. "Tracking, Students' Effort, and Academic Achievement." *Sociology of Education* 78(1):27–49.

Carbonaro, William J., and Adam Gamoran. 2002. "The Production of Achievement Inequality in High School English." *American Educational Research Journal* 39(4):801–27.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2015. "The Aftermath of Accelerating Algebra Evidence from District Policy Initiatives." *Journal of Human Resources* 50(1):159–88.

Cremin, Lawrence A. 1965. *The Genius of American Education*. Pittsburgh: University of Pittsburgh Press.

Cronbach, L. J., and Snow R. E. 1977. *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. Oxford, UK: Irvington.

Domina, T., Hanselman, P., Hwang, N., & McEachin, A. (2016). "Detracking and Tracking Up: Mathematics Course Placements in California Middle Schools, 2003–2013." *American Educational Research Journal*, 53(4):1229–66.

Domina, T., McEachin, A., Penner, A., & Penner, E. (2015). "Aiming High and Falling Short: California's Eighth-Grade Algebra-for-all Effort." *Educational Evaluation and Policy Analysis*, 37(3): 275–95.

Domina, T., Penner, A., & Penner, E. (2017). "Categorical Inequality: Schools as Sorting Machines." *Annual Review of Sociology*, 43:311–30.

Domina, T., & Saldana, J. (2012). "Does Raising the Bar Level the Playing Field? Mathematics Curricular Intensification and Inequality in American High Schools, 1982–2004." *American Educational Research Journal*, 49(4):685–708.

Dougherty, Shaun M., Joshua S. Goodman, Darryl V. Hill, Erica G. Litke, and Lindsay C. Page. 2015. "Middle School Math Acceleration and Equitable Access to Eighth-Grade Algebra Evidence from the Wake County Public School System." *Educational Evaluation and Policy Analysis* 37(1 Suppl.):80S–101S.

Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review* 69(5): 613–35.

Dreeben, Robert, and Rebecca Barr. 1988. "Classroom Composition and the Design of Instruction." *Sociology of Education* 61(3):129–42.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2008. *Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya* (No. w14475). Cambridge, MA: National Bureau of Economic Research.

Eccles, Jacquelynne, and Robert W. Roeser. 2011. "Schools as Developmental Contexts During Adolescence." *Journal of Research on Adolescence* 21(1): 225–41.

Epple, Dennis, Elizabeth Newlon, and Richard Romano. 2002. "Ability Tracking, School Competition, and the Distribution of Educational Benefits." *Journal of Public Economics* 83(January):1–48.

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

Frank, Kenneth A., Chandra Muller, Kathryn S. Schiller, Catherine Riegle-Crumb, Anna Strassmann Mueller, Robert Crosnoe, and Jennifer Pearson. 2008. "The Social Dynamics of Mathematics Coursetaking in High School." *American Journal of Sociology* 113(6):1645–96.

Gamoran, Adam. 1992. "The Variable Effects of High School Tracking." *American Sociological Review*: 812–28.

Gamoran, Adam. 2010. "Tracking and Inequality." Pp. 213–28 in *The Routledge International Handbook* *of the Sociology of Education*, edited by M. W. Apple, S. J. Ball, and L. A. Gandin. New York: Routledge.

Gamoran, Adam, and Robert D. Mare. 1989. "Secondary School Tracking and Educational Inequality: Compensation, Reinforcement, or Neutrality?" *American Journal of Sociology* 94(5):1146–83.

Gamoran, Adam, and Martin Nystrand. 1994. "Tracking, Instruction and Achievement." *International Journal of Educational Research* 21(2):217–31.

Gamoran, Adam, Martin Nystrand, Mark Berends, and Paul C. LePore. 1995. "An Organizational Analysis of the Effects of Ability Grouping." *American Educational Research Journal* 32(4):687–715.

Hallinan, Maureen T. 1992. "The Organization of Students for Instruction in the Middle School." *Sociology of Education* 65(2):114–27.

Hallinan, Maureen T. 1994a. "School Differences in Tracking Effects on Achievement." *Social Forces* 72(3):799–820.

Hallinan, Maureen T. 1994b. "Tracking: From Theory to Practice." *Sociology of Education* 67(2):79–84.

Hallinan, Maureen T. 1996. "Track Mobility in Secondary School." *Social Forces* 74(3):983–1002.

Hallinan, Maureen T., and Aage B. Sørensen. 1985. "Ability Grouping and Student Friendships." *American Educational Research Journal* 22(4):485–99.

Hanushek, Eric A., and Ludger Wößmann. 2006. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." *Economic Journal* 116(510): C63–C76.

Hoffer, Tomas B. 1992. "Middle School Ability Grouping and Student Achievement in Science and Mathematics." *Educational Evaluation and Policy Analysis* 14(3):205–27.

Horvath, Hedvig. 2015. "Classroom Assignment Policies and Implications for Teacher Value-Added Estimation." Unpublished manuscript.

Katz, Michael B. 1971. *Class, Bureaucracy and Schools: The American Illusion of Educational Change*. New York: Praeger.

Kelly, Sean. 2007. "The Contours of Tracking in North Carolina." *High School Journal* 90(4):15–31.

Kelly, Sean. 2009. "The Black-White Gap in Mathematics Course Taking." *Sociology of Education* 82(1): 47–69.

Kelly, Sean, and William Carbonaro. 2012. "Curriculum Tracking and Teacher Expectations: Evidence from Discrepant Course Taking Models." *Social Psychology of Education* 15(3):271–94.

Kelly, Sean, and Heather Price. 2011. "The Correlates of Tracking Policy Opportunity Hoarding, Status Competition, or a Technical-Functional Explanation?" *American Educational Research Journal* 48(3):560–85.

Kerckhoff, Alan C. 1986. "Effects of Ability Grouping in British Secondary Schools." *American Sociological Review* 51(6):842–58.

Kerckhoff, Alan C. 1995. "Institutional Arrangements and Stratification Processes in Industrial Societies." *Annual Review of Sociology* 21:323–47.

Kliebard, Herbert M. 1975. "The Rise of Scientific Curriculum Making and its Aftermath." *Curriculum Theory Network* 5(1):27–38.

Legewie, Joscha, and Thomas A. DiPrete. 2012. "School Context and the Gender Gap in Educational Achievement." *American Sociological Review* 77(3):463–85.

Lewis, Amanda E., and John B. Diamond. 2015. *Despite the Best Intentions: How Racial Inequality Thrives in Good Schools*. Oxford: Oxford University Press.

Long, Mark C., Dylan Conger, and Patrice Iatarola. 2012. "Effects of High School Course-Taking on Secondary and Postsecondary Success." *American Educational Research Journal* 49(2):285–322.

Loveless, Tom. (2011). *The Tracking Wars: State Reform Meets School Policy*. Washington, DC: Brookings Institution Press.

Lucas, Samuel R. 1999. *Tracking Inequality: Stratification and Mobility in American High Schools*. Sociology of Education series. New York: Teachers College Press.

Lucas, Samuel R., and Mark Berends. 2002. "Sociodemographic Diversity, Correlated Achievement, and De Facto Tracking." *Sociology of Education* 75(4): 328–48.

Lucas, Samuel R., and Aaron D. Good. 2001. "Race, Class, and Tournament Track Mobility." *Sociology of Education* 74(2):139–56.

Mann, Horace. [1848] 1957. "Twelfth Annual Report to the Massachusetts Board of Education." P. 79 in *The Republic and the School: Horace Mann on the Education of Free Men*, edited by L. A. Cremin. New York: Teachers College Pres.

Marsh, Herbert W. 1987. "The Big-Fish-Little-Pond Effect on Academic Self-Concept." *Journal of Educational Psychology* 79(3):280.

McFarland, Daniel. A. 2006. "Curricular Flows: Trajectories, Turning Points, and Assignment Criteria in High School Math Careers." *Sociology of Education* 79(3):177–205.

Metz, Mary Haywood.1978. *Classrooms and Corridors: The Crisis of Authority in Desegregated Secondary Schools*. Berkeley: University of California Press.

Metz, Mary Haywood. 2003. *Different by Design: The Context and Character of Three Magnet Schools*. New York: Teachers College Press.

Mickelson, Roslyn Arlin. 2001. "Subverting Swann: First- and Second-Generation Segregation in the Charlotte-Mecklenburg Schools." *American Educational Research Journal* 38(2):215–52.

Nomi, Takako, and Elaine M. Allensworth. 2012. "Sorting and Supporting Why Double-Dose Algebra Led to Better Test Scores but More Course Failures." *American Educational Research Journal* 50(4):756–88.

Nomi, Takako, and Stephen W. Raudenbush. 2016. "Making a Success of 'Algebra for All': The Impact

of Extended Instructional Time and Classroom Peer Skill in Chicago." *Educational Evaluation and Policy Analysis* 38(2):431–51.
Nystrand, Martin, and Adam Gamoran. 1997. "The Big Picture: Language and Learning in Hundreds of English Lessons." Pp. 30–74 in *Opening Dialogue*, by M. Nystrand. New York: Teachers College Press.
Oakes, Jeannie. 1985. *Keeping Track: How Schools Structure Inequalities*. New Haven, CT: Yale University Press.
Oakes, Jeannie, and Gretchen Guiton. 1995. "Matchmaking: The Dynamics of high School Tracking Decisions." *American Educational Research Journal* 32(1):3–33.
Oakes, Jeannie, Amy Stuart Wells, Makeba Jones, and Amanda Datnow. 1997. "Detracking: The Social Construction of Ability, Cultural Politics, and Resistance to Reform." *Teachers College Record* 98: 482–510.
Page, Reba Neukom. 1991. *Lower-Track Classrooms: A Curricular and Cultural Perspective*. New York: Teachers College Press.
Penner, A. M., Domina, T., Penner, E. K., & Conley, A. (2015). "Curricular Policy as a Collective Effects Problem: A Distributional Approach." *Social Science Research*, 52:627–l41.
Powell, Arthur G., Eleanor Farrar, and David K. Cohen. 1985. *The Shopping Mall High School*. Boston: Houghton Mifflin.
Raudenbush, Stephen W., and Robert D. Eschmann. 2015. "Does Schooling Increase or Reduce Social Inequality?" *Annual Review of Sociology* 41:443–70.
Rickles, Jordan H. 2011. "Using Interviews to Understand the Assignment Mechanism in a Nonexperimental Study: The Case of Eighth Grade Algebra." *Evaluation Review* 35(5):490–522.
Rosenbaum, James E. 1976. *Making Inequality: The Hidden Curriculum of High School Tracking*. New York: Wiley.
Rosenbaum, James E. 1999. "If Tracking Is Bad, Is Detracking Better?" *American Educator* 23(4):24–29.
Shavit, Yossi, and Hans-Peter Blossfeld. 1993. *Persistent Inequality: Changing Educational Attainment in Thirteen Countries*. Social Inequality series. Boulder, CO: Westview Press.
Slavin, Robert E. 1988. "Synthesis of Research on Grouping in Elementary and Secondary Schools." *Educational Leadership* 46(1):67–77.
Snow, Richard. 1989. "Aptitude-Treatment Interaction as a Framework for Research on Individual Differences in Learning." Pp. 13–59 in *Learning and Individual Differences*, edited by P. Ackerman, R. J. Sternberg, and R. Glaser. New York: Freeman.
Sørensen, Aage Bøttger. 1970. "Organizational Differentiation of Students and Educational Opportunity." *Sociology of Education* 43(4):355–76.
Stainback, Kevin, Donald Tomaskovic-Devey, and Sheryl Skaggs. 2010. "Organizational Approaches to Inequality: Inertia, Relative Power, and Environments." *Annual Review of Sociology* 36(1):225–47.
Stein, Mary Kay, Julia Heath Kaufman, Milan Sherman, and Amy F. Hillen. 2011. "Algebra: A Challenge at the Crossroads of Policy and Practice." *Review of Educational Research* 81:453–92.
Tyson, Karolyn. 2011. *Integration Interrupted: Tracking, Black Students, and Acting White After Brown*. New York: Oxford University Press.
Van de Werfhorst, Herman G., and Jonathan J. B. Mijs. 2010. "Achievement Inequality and the Institutional Structure of Educational Systems: A Comparative Perspective." *Annual Review of Sociology* 36:407–28.
Van Houtte, Mieke. 2004. "Tracking Effects on School Achievement: A Quantitative Explanation in Terms of the Academic Culture of School Staff." *American Journal of Education* 110(4):354–88.
von Hippel, Paul T., Joseph Workman, and Douglas B. Downey. 2018. "Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, of 'Are Schools the Great Equalizer?'" *Sociology of Education* 91(4): 323–57.
Watanabe, Maike. 2006. "Some People Think This School Is Tracked and Some People Don't": Using Inquiry Groups to Unpack Teachers' Perspectives on Detracking. *Theory into Practice* 45(1):24–31.
Wells, Amy Stuart, and Jeannie Oakes. 1996. "Potential Pitfalls of Systemic Reform: Early Lessons from Research on Detracking." *Sociology of Education* 69(Extra Issue):135–43.
Zimmer, Ron. 2003. "A New Twist in the Educational Tracking Debate." *Economics of Education Review* 22(3):307–15.

Author Biographies

Thurston Domina is associate professor of educational policy and organizational leadership at the University of North Carolina-Chapel Hill. He works in partnership with educational practitioners to better understand the relationship between education and social inequality in the contemporary U.S.

Andrew McEachin is a policy researcher in the Economics, Statistics, and Sociology Department at the RAND Corporation and a professor at the Pardee RAND Graduate School. His research focuses on the determinants of persistent achievement gaps, as well as evaluating the effect of popular responses by policymakers and educators to reduce these gaps.

Paul Hanselman is assistant professor of sociology at the University of California-Irvine. His research focuses on how schools influence social inequalities and the impacts of school-based educational interventions on these processes.

Priyanka Agarwal is a PhD candidate at School of Education, University of California-Irvine. Her research is focused on designing and examining inclusive learning environments in the context of K-12 mathematics classrooms through design collaborations with teachers and youth.

NaYoung Hwang is a postdoctoral associate researcher at the Center for Research on Educational Opportunity (CREO) at the University of Notre Dame. NaYoung is interested in the extent to which educational policies and practices can mitigate and aggravate educational inequality.

Ryan W. Lewis is a research associate at WestEd within the organization's Comprehensive School Assistance Program. He is interested in the intersection of education policy, implementation, and targeted programs, specifically in settings that confront educational and social inequalities.