Coombs, P., & Hallak, J. (1987). *Cost analysis in education: A tool for policy and planning.* Baltimore: Johns Hopkins University Press.
Eicher, J. (1995). International educational expenditures. In M. Carnoy (Ed.), *International encyclopedia of economics of education* (2nd ed., pp. 443–450). Oxford, UK: Pergamon Press.
Geske, T., Davis, D., & Hingle, P. (1997). Charter schools: A viable public school choice option? *Economics of Education Review, 16*(1), 15–23.
Hanushek, E., & Jorgenson, D. (Eds.). (1996). *Improving America's schools: The role of incentives.* Washington, DC: National Academy Press.
Hartman, W. (1990). Supplemental/replacement: An alternative approach to excess costs. *Exceptional Children, 56*(5), 450–459.
Hough, J. (1994). Educational cost-benefit analysis. *Education Economics, 6*(1), 1–14.
Inter-Agency Commission for Basic Education for All. (1990). *Meeting basic human needs. A background report for the World Conference on Basic Education for All.* New York: InterAgency Commission, United Nations.
Jalil, N., & McGinn, N. (1992). Pakistan. In R. Thomas (Ed.), *Education's role in national development plans.* New York: Praeger.
Jamieson, D., Klees, S., & Wells, S. (1978). *The costs of educational media: Guidelines for planning and evaluation.* Beverly Hills, CA: Sage.
Jimenez, E., Lockheed, M., & Wattanawaha, N. (1988). The relative efficiency of public and private schools: The case of Thailand. *World Bank Economic Review, 2*(2), 139–164.
King, J. (1994). Meeting the educational needs of at-risk students: A cost analysis of three models. *Educational Evaluating and Policy Analysis, 16*(1), 1–19.
Klees, S. (1995). Economics of educational technology. In M. Carnoy (Ed.), *The international encyclopedia of economics of education* (2nd ed., pp. 398–406). Oxford, UK: Pergamon Press.
Lau, L. (1979). Educational production functions. In *Economic dimensions of education* (pp. 33–69). Washington, DC: National Academy of Education.
Levin, H. (1983). *Cost-effectiveness analysis: A primer.* Beverly Hills, CA: Sage.
Levin, H. (1995). Cost-effectiveness analysis. In M. Carnoy (Ed.), *International encyclopedia of economics of education* (2nd ed., pp. 381–386). Oxford, UK: Pergamon Press.
Levin, H. (1996). Economics of school reform for at-risk students. In E. Hanushek & D. Jorgenson (Eds.), *Improving America's schools: The role of incentives.* Washington, DC: National Academy Press.
Levin, H., Glass, G., & Meister, G. (1987). A cost-effectiveness analysis of computer-assisted instrument. *Evaluation Review, 11*(1), 50–72.
Mayoya, M. (1997). *Private costs and access to secondary education in Burundi.* Unpublished doctoral dissertation, College of Education, Michigan State University, East Lansing.
Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development, 22*(9), 1325–1343.
Tibi, C. (1986). The determinants of educational cost. *IIEP Newsletter, 4,* 1–2.
Tsang, M. (1994a). *Cost analysis of educational inclusion of marginalized populations.* Paris: International Institute for Educational Planning, UNESCO.
Tsang, M. (1994b). Costs of education in China: Issues of resource mobilization, equality, equity, and efficiency. *Education Economics, 2*(3), 287–312.
Tsang, M. (1995). Public and private costs of education in developing countries. In M. Carnoy (Ed.), *International encyclopedia of economics of education* (2nd ed., pp. 393–398). Oxford, UK: Pergamon Press.
Tsang, M. (1997). The costs of vocational training. *International Journal of Manpower, 18*(1–2), 63–89.
Tsang, M., & Taoklam, W. (1992). Comparing the costs of government and private primary education in Thailand. *International Journal of Educational Development, 12*(3), 177–190.
World Bank. (1995). *Priorities and strategies for education: A World Bank review.* Washington, DC: World Bank.

Author

MUN C. TSANG is a professor of economics of education at the College of Education, Michigan State University, East Lansing, MI 48824. His specialties are the cost and financing of education and training, and the economic effects of education.

Manuscript received December 13, 1996
Revision received July 14, 1997
Accepted July 25, 1997

Educational Evaluation and Policy Analysis
Winter 1997, Vol. 19, No. 4, pp. 325–338

Upgrading High School Mathematics Instruction: Improving Learning Opportunities for Low-Achieving, Low-Income Youth

Adam Gamoran, Andrew C. Porter, John Smithson, and **Paula A. White**
University of Wisconsin-Madison

Low-achieving, low-income students are typically tracked into dead-end math courses in high school. In this article, the authors evaluate the success of "transition" math courses in California and New York, which are designed to bridge the gap between elementary and college-preparatory mathematics and to provide access to more challenging and meaningful mathematics for students who enter high school with poor skills. The authors hypothesize that the transition courses—Math A in California and Stretch Regents and UCSMP Math in New York—allow students to keep pace with those who enter college-preparatory courses by covering rigorous mathematical content using a range of cognitive strategies. Data from 882 students in 48 math classes are analyzed using a three-level hierarchical linear model. The results show that growth in student achievement is significantly lower in general-track classes than in college-preparatory classes. Achievement in transition classes falls in between: not significantly lower than in college-preparatory classes, but not significantly greater than in general-track classes. More rigorous content coverage accounts for much of the achievement advantage of college-preparatory classes. The transition classes are judged a partial success in meeting their goal of upgrading the quality of mathematics instruction for low-achieving, low-income youth.

A serious impediment to the academic progress of low-achieving high school students is that they are typically tracked into low-level, dead-end mathematics classes (Gamoran, 1987). This problem is especially severe for students in low-income communities, where the quality of the mathematics curriculum may be weak overall and especially weak in low-level, general math classes (Oakes, 1990). How can this problem be overcome? In this study, we examine the consequences of efforts in California and New York to replace low-level general math with "transition" courses that are supposed to bridge the gap between elementary and college preparatory mathematics, leading to more challenging and meaningful mathematics instruction for students who begin high school with poor mathematics skills. Do the transition courses provide a rigorous instructional experience for low-achieving students? How much mathematics do students learn? We examine these questions with data from 882 students in 48 classes in seven high schools.

Upgrading Low-Level Math: A Response to the Tracking Problem?

Despite extensive research and criticism, tracking for mathematics remains a near-universal practice in American high schools (National Center for Education Statistics, 1994). Schools routinely assign students to classes based on tested or teacher-perceived performance levels (Oakes, Gamoran, & Page, 1992). This practice is supposed to allow teachers to gear instruction to an appropriate level and pace for each separate group of students (Barr & Dreeben, 1983). Although the system seems logical, two major problems have emerged. First, because of stratification in the wider society, dividing students by "ability" tends to separate students on other grounds as well, particularly race, ethnicity, and social class (Gamoran, Nystrand, Berends, & LePore, 1995). Second, tracking promotes inequality: The achievement gap between students in high-level classes and those in low-level classes grows over time (Oakes et al, 1992). This occurs as students in high tracks learn more and low-track students learn less than similar students in mixed-ability classes (Gamoran & Nystrand, 1994; Hoffer, 1992; Kerckhoff, 1986). A major cause for increasing inequality is that the pace, complexity, and challenge of classroom instruction are higher in high-track classes than elsewhere (Gamoran, 1993; Gamoran et al., 1995; Hoffer & Gamoran, 1993; see Oakes et al., 1992, for a review). Tracking in mathematics involves an additional problem of rigidity; high school mathematics courses are arranged in a sequence that is tightly controlled in most schools, so it is difficult to catch up if one starts out behind (Grossman & Stodolsky, 1995; Hallinan, 1996; Loveless, 1994; Rosenbaum, 1976). Overall, it is clear that tracking is not working as it was intended or, at best, it is working well only for students in the top levels of the system.

The tracking problem has two possible responses. One, the solution favored by most critics, is to eliminate tracking (e.g., Oakes, 1985; Wheelock, 1992). In this approach, all students would be placed in mixed-ability classes, where the curriculum and instruction would be, in principle, at the same level of rigor that at present one typically finds only in college-preparatory classes. Although many schools are experimenting with de-tracking, success appears hard to achieve (Gamoran & Weinstein, in press). Oakes (1994) has identified three problems that make detracking difficult: the *normative* problem of overcoming prevailing beliefs about variability in cognitive ability and the need to separate students who score differently on tests, the *political* problem of overcoming resistance of stakeholders who believe that tracking benefits their interests (e.g., parents of high-achieving students, teachers of high-ability classes), and the *technical* problem of finding effective methods for teaching students who differ in their performance levels. None of these problems has been solved. Most research concentrates on documenting the normative and political problems, and less attention has been paid to the technical complexities of teaching mixed-ability classes at the high school level.

A second response to the tracking problem is to improve instruction for low-achieving students without attempting to teach all students in the same learning context. This approach, if successful, would mitigate the harmful effects of tracking on inequality. It would not eliminate the problem of separating students from varied social backgrounds, but in inner-city schools where virtually all students are poor, that may be a lesser consideration if the overall level of learning can be improved. Because it does not mix all students together, this approach risks creating new low tracks that may deny opportunities for advanced study. However, previous research suggests this outcome is not inevitable (White, Gamoran, Smithson, & Porter, 1996). Is it feasible to enhance student learning in courses specifically designed to improve instruction for low-achieving students? In this study, we evaluate the success of efforts to improve the quality of mathematics instruction for low-achieving high school students in two states—California and New York. In light of the difficulties encountered by de-tracking reforms, this alternative approach seems worthy of consideration.

The Policy Context of Upgrading High School Math

Raising academic standards has been a recurring theme in American education reform. In the 1950s, excellence in mathematics and science was seen as a national priority to maintain international competitiveness (Conant, 1959). More recently, the 1983 report of the National Commission on Excellence in Education, *A Nation At Risk*, called for increases in academic coursework as a tonic for educational mediocrity. In response, many states mandated or encouraged stiffer academic requirements for high school graduation, and on average, students in the late 1980s were taking more academic courses than had those in the early 1980s (Clune & White, 1992; Massell & Fuhrman, 1994; Wilson & Rossman, 1993). However, part of the increase in academic course-taking occurred in low-level courses, such as general math (Wilson & Rossman, 1993).

Another landmark in this era of reform was the 1989 publication of *Curriculum and Evaluation Standards for School Mathematics* by the National Council of Teachers of Mathematics (NCTM). In calling for high standards, NCTM emphasized excellence for all students, not just those at the top of the academic hierarchy, in contrast to the reforms of the 1950s and the early 1980s (Conant, 1959; National Commission on Excellence in Education, 1983). In addition, the NCTM *Standards* emphasized the more complex cognitive tasks of understanding, application of knowledge, and reasoning rather than memorization and routine computation, and again these standards were meant for all students, not just the most academically successful. Porter, Archbald, and Tyree (1991) characterized this call for reform as "hard content for all students."

In this policy environment, many states and districts began to develop new ways of providing more challenging and meaningful mathematics instruction to students who entered high school with poor mathematics skills. Instead of tracking students into non-academic courses or dead-end academic courses, a more innovative approach was to create transition courses intended to bridge the gap between elementary arithmetic and college-preparatory math. One example with which we were familiar from previous research was Math A, a course developed by mathematics teachers in California with state support. Math A was designed to integrate college-preparatory mathematics using instructional methods as proposed in the NCTM reforms, i.e., emphasizing understanding and reasoning over memorization (Porter, Kirst, Osthoff, Smithson, & Schneider, 1993). Through contacts with state administrators, educational researchers, and the research literature, we identified a number of other states besides California in which mathematics upgrading reforms were occurring, including Connecticut, Florida, Illinois, Iowa, Kentucky, Louisiana, Montana, New York, Oregon, North Carolina, South Carolina, Texas, Vermont, Washington, and Wisconsin. Most of these states, however, were in the pilot stages of upgrading the mathematics curriculum. We decided to focus on New York, in addition to California, because upgrading was more extensive and further developed in New York than in the other states (White, 1995). Moreover, we discovered that in New York it would be possible to examine two different upgraded courses: Stretch Regents, a revision of the Regents math sequence that allows students a longer period of time to cover the college-preparatory curriculum and UCSMP Transition Math, a course in the University of Chicago School Mathematics Project that provides an alternative form of college-preparatory mathematics.

Following these decisions, we prepared to evaluate the success of three transition courses: Math A in California and Stretch Regents and UCSMP Transition Math in New York.

Math A

In California, Math A was intended to replace general math courses (California State Department of Education, 1985). Originally, Math A was designed as an alternative to the college-preparatory sequence of algebra and geometry. In practice, however, it served as a bridge to the college-preparatory courses (White, Gamoran, Smithson, & Porter, 1996). That is, students who are successful in Math A subsequently enroll in algebra. Some students, typically those who are less successful in Math A, complete a second transition course (Math B) and then move on to algebra. In previous research, we showed that students who enroll in Math A are more likely than those who enroll in general math to complete a college-preparatory sequence (White, Gamoran, Smithson, & Porter, 1996).

Math A involves an integrated curriculum of arithmetic, algebra, and geometry (Porter et al., 1993). It emphasizes problem-solving, real-world applications, empirical reasoning, and manipulatives. Implementation of Math A has varied across districts and schools (White, 1995). Some educators view the course as an innovative approach for readying students for collegepreparatory mathematics. Others see Math A as important for all students to improve their mathematical understanding and skills.

Stretch Regents

Regents Math I-II-III is a three-year sequence of college-preparatory mathematics in which algebra, geometry, and higher mathematics are integrated. To allow students with lower-level mathematics skills to enter the college-preparatory sequence, New York educators designed Stretch Regents as a course that covers the first two years of the Regents curriculum, but at a slower pace. Some schools implemented Stretch Regents by covering half of each Regents textbook each year. Other schools covered a whole textbook in the first year, but only the "easy" parts. They covered the "hard" parts of the same textbook in the second year (White, 1995).

Like Math A, the Stretch Regents curriculum integrated algebra and geometry content. However, Stretch Regents has not involved a revision of pedagogy. Traditional emphases on computation and routines, in addition to problem-solving, remain in Stretch Regents, the same as in the Regents curriculum.

UCSMP Math

In Buffalo, New York, educators responded to high failure rates in Regents courses by adopting the UCSMP curriculum (White, 1995). This six-year sequence is designed to begin in grade seven, but in practice many low-achieving students enroll in the first course in grade nine. Hence, first-year UCSMP essentially serves as a bridge to college-preparatory mathematics, like Math A. Another similarity to Math A is that UCSMP Math emphasizes problem-solving and real-world applications. Thus, like Math A and in contrast to Stretch Regents, UCSMP Math involves not only more challenging content but pedagogical reforms such as working in groups and using manipulatives, along the lines promoted by the NCTM *Standards*.

Hypotheses

In earlier work, we found that the transition courses were partially successful in leading students to more meaningful mathematics (White, Gamoran, Smithson, & Porter, 1996). Students who entered ninth grade in transition courses were more likely to complete two years of college-preparatory math than those who began in the general track. Moreover, students in transition courses did not suffer any loss of the mathematics credits needed for high school graduation. However, students in transition courses were less likely to complete a college-preparatory sequence than those placed in college-preparatory mathematics from the beginning, even when prior grades or test scores were statistically controlled.

To further evaluate the viability of transition courses as a solution to the problem of low-level, dead-end high school math, we need to examine student achievement. How much mathematics do students learn in transition courses? Do they keep pace with those in college-preparatory classes? If so, then the transition courses may be judged a success. For this study, we consider two hypotheses:

- Learning gains of students in transition courses are similar to those of students in college-preparatory courses, while those in traditional general-track courses lag behind.

- The reason students in transition courses keep pace in learning with those in college-preparatory courses is that, like the college-preparatory classes, transition courses cover a rigorous mathematics curriculum and emphasize complex cognitive demands.

Addressing these hypotheses requires data about course types, instructional content and methods, and student achievement.

Sample and Data

We obtained data for our study from two school districts in each state. To select the districts, we used the following criteria: (a) urban districts with large percentages of low-achieving students (because the transition math courses were designed to assist low-achieving students), and (b) districts that had initiated the new mathematics initiatives several years previously. The district selection process involved interviews with state department of education officials, district level mathematics coordinators, school administrators, and mathematics teachers. This process led us to select San Francisco and San Diego in California and Rochester and Buffalo in New York.

We then selected two high schools in each district (an exception was the Buffalo district where only one high school met our criteria). Because the transition courses were designed to assist low-achieving students, we selected schools with high concentrations of low-income and lower-achieving students. The following criteria were used to select the schools:

- Comprehensive grade 9–12 school that maintained the same grade-level organization before and after the implementation of the new mathematics initiatives.

- Average student achievement in the lowest quartile of schools within the district and state (based on SAT, CAP, and Regents test scores).

- High percentage of low-income students (based on the number of students receiving free and reduced lunches).

- Schools that offered at least one of the following arrangements: an innovative course used as a transition to college-preparatory mathematics, a stretch version of college-preparatory mathematics, or an alternative approach to college-preparatory mathematics.

- A math transition program that had been in place for at least two years.

- At least three mathematics teachers in the school teaching the transition math courses.

Within each school, we selected transition courses, along with at least one traditional, low-level course (e.g., general math or pre-algebra) and at least one college-preparatory course (e.g., Regents I or algebra). San Francisco was an exception to this design because all the lower-level mathematics courses had been eliminated by the time we began our data collection. Initially, our data collection included 56 classes in the seven schools (8 per school) and over 1,600 students.

During the 1992–1993 school year, we administered tests and questionnaires to students and interviewed teachers and administrators. Findings from interviews were summarized in an article about the implementation of the transition courses (White, 1995). In addition, we collected transcript data for earlier cohorts of students, and these results have also been reported (White, Gamoran, Smithson, & Porter, 1996). In this article, we focus on growth in student achievement for students in the different types of mathematics courses.

Mathematics Achievement Test

Our main concern in creating a mathematics test was for the validity of a single exam to test achievement across a wide variety of mathematics courses of varying levels of difficulty and course content. We created a test from NAEP 1990 public-release items that was oriented toward higher-order thinking and problem-solving skills rather than computational skills and designed for administration to a general population of students. Seventy-five percent of the items were multiple-choice, and 25% were completing short answers. There were no extended-response items.

We used a taxonomy of mathematics content areas and tasks to identify topics and cognitive demands (Porter et al., 1993). This taxonomy consisted of 10 general areas of mathematics, with each area divided into 7 to 10 specific topics, for a total of 93 topics. Cognitive demand was defined according to six levels: (1) memorize facts, (2) understand concepts, (3) perform procedures/solve equations, (4) collect/interpret data, (5) solve word problems, and (6) solve novel problems. Crossing topics with cognitive demand yielded 558 specific types of content that might have been taught and/or tested. Using this approach, examples of mathematics content are "solving problems using measures of central tendency" and "performing procedures involving units of time." Each teacher reported the amount of instructional time for each topic and each type of cognitive demand associated with that topic. Separately, we also determined the percent of test items for each combination of topic and cognitive demand.

The test consisted of 26 problems in the following content areas: 15% arithmetic, 20% measurement, 15% algebra, 20% geometry, 20% probability, and 10% numbers and sets. Types of problems included were 30% concepts, 15% procedures, 20% data interpretation, 27% routine word problems, and 8% novel word problems. These content areas and problem types are consistent with the NCTM call for rigorous mathematics content that promotes understanding and emphasizes problem-solving.

Because the test was used to determine the effects on student achievement gains of three quite different sets of courses—general math, first-year college preparatory math, and transition math—there is a concern for content validity overall and differential content validity across course types. In comparison to teacher descriptions of what they taught, the test underemphasized number and number relations (0.038 on the test versus 0.24 in instruction) and arithmetic (0.15 versus 0.25). The test overemphasized geometry (0.19 versus 0.06) and probability (0.19 versus 0.04). Test and content of instruction matches were quite good on measurement, algebra, trigonometry, statistics, advanced algebra, and finite/discrete mathematics. Across all 10 categories of topics, the mean proportion of classroom instruction on tested content was quite similar for the different sets of courses. The largest difference was between Math A/B (0.30) and algebra (0.26). Thus, with regard to general content areas of mathematics, the test was not seriously biased for or against any of the types of courses being compared.

At each of the seven high schools, the test was administered three times during the 1992–1993 school year: to 1,068 students at the beginning of the school year (in September 1992), to 1,228 students at the beginning of second semester (in January 1993), and to 1,013 students at the end of the school year (in May 1993). The total number of individual students tested was 1,678, and the total number of tests administered was 3,399. The test was administered by the teacher; students had a class period to complete the test.

Response rates of students enrolled in the courses were 80% in the fall, 83% in the winter, and 75% in the spring. (Some new course sections were tested in the spring so that students who had moved to different sections of the same course could be retained in the sample.) Of the 1,068 students who responded to the fall test, 833 (78%) were retested in the winter or the spring, allowing us to measure their achievement growth. (Most of these students, 620, were retested in both the winter and the spring.) An additional 287 students missed the fall test but responded in the winter and spring, allowing us to assess their achievement growth also.

Means and standard deviations on the achievement test and other variables are reported in Table 1. The average score on the pretest was about 10 out of a total of 26 possible points.[1] On average, students gained about 1.7 points over the course of the school year. While some portion of these significant gains may have been due to repeated use of the same test, much of the gains can be attributed to instruction. First, the gains were significantly different across different types of math courses taken. Second, the gains were predicted by teachers' descriptions (through questionnaires) of what was taught. In fact, the ability to predict student achievement gains through knowledge of what students had studied is an important side result of the work, about which more will be said later.

Content Coverage

To assess instruction in each class, we used our test instrument as a measuring rod, examining the extent to which mathematical content and cognitive demands in our test were included in the sample classes. Teacher questionnaires provided information on the extent to which the topics on our tests were covered in the sample classes and whether the cognitive demands made on our test also occurred in mathematics instruction.

Our indicator of content coverage reflects both the proportion of instructional time that was spent covering tested content (level of coverage) and the match of relative emphases of types of content between instruction and the test (configuration of coverage). These are the classic level and configuration properties that are used in profile comparisons. Specifically, the level of coverage was defined as the proportion of instructional time that addressed one or more types of content tested (mean of 0.07, standard deviation of 0.02). Nineteen types of content were tested out of the total possible of 558.

The configuration of coverage was calculated in three steps. First, we calculated the proportion of instructional time for each of the 19 unique content areas tested relative to the total amount of instructional time spent on tested content. The result was 19 separate proportions representing instruction for the tested content. Second, we computed the proportion of the 19 content areas represented on the test. Because the test consisted of 26 items, each item represented 0.038 of the entire test ($1/26$ = 0.038). However, some content areas were represented by two or three questions, thus four content areas represented 0.076 of the test (= 2 x 0.038), and one content area represented 0.114 (= 3 x 0.038) of the test. In the third and final step for calculating the configuration of coverage, the absolute value of the differences between the proportions for each content area represented on the test and the proportion of classroom instruction spent on that content was summed and the sum re-scaled by dividing by 2.0 and subtracting the result from 1.0 to create an index that ranges from 0.0 to 1.0, with 1.0 representing a perfect match in configuration between instruction and tested content (mean of 0.58, standard deviation of 0.08).

Our final indicator of content coverage, which we use in our analyses, is the product of level and configuration (mean of 0.04, standard deviation of 0.01, range from 0.02 to 0.07). We used this compound indicator because we expected test scores to be higher not just when appropriate topics are covered, but when they are covered in depth. We examined several alternatives for combining level and configuration, including the sum of level and configuration and level and configuration as separate indicators alone and in combination with their product, and found that our models were most stable when only the product was included. The product of level and configuration correlated 0.451 with class gains and 0.259 with student gains. Content was defined as the intersection of topics (93) and cognitive demand (6) because it was at that level of detail that correlations with student achievement gains were highest. Using topics only, the correlations with student achievement gains were −0.205 at the class level and 0.103 at the student level. For cognitive demand only, the correlations were 0.112 at the class level and 0.069 at the student level. Clearly, to predict student achievement gains from knowledge of the content of instruction, a micro-level description of content that looks at cognitive demand by topic is the most useful approach considered to date. The strong correlation of instructional alignment to tested content further establishes the validity of our test for the purposes used here Thus, our model assumes that level and configuration are ineffective alone and matter only in com bination. This assumption seems reasonable: Grea range with shallow depth and great depth in a nar row range of coverage both seem unlikely to resul in substantial achievement.

Table 2 shows that coverage varies among th types of classes in a predictable way. Coverage highest in Regents and algebra classes, followe closely by UCSMP Math, then Math A and Stretc Regents. The lowest coverage occurred in gener and pre-algebra classes.

Control Variables

In a study of achievement growth, it is importa to take account of pre-existing conditions that m be related to both course type and achievement, ov and above the pretests. From student questionnair we obtained information about students' gend race, ethnicity, and their final grades in mathem ics in the previous year. As seen in Table 1, c sample consists of 41% Black, 24% Hispanic, a 18% Asian students. (The Asian students in c sample were predominantly poor, with origins Southeast Asia, i.e., Vietnam, Laos, and Camb dia.)

We did not obtain data on socioeconomic bac grounds from students, but we asked teachers

TABLE 1
Means and Standard Deviations of Variables

Variable	*M*	*SD*
Within students (N = 2,262 test scores)		
Time (0 = fall, 1 = winter, 2 = spring)	1.04	0.79
Test score[a]	10.99	4.71
Student level (N = 882 students)		
Male	0.56	0.50
Black	0.41	0.49
Hispanic	0.24	0.43
Asian	0.18	0.38
Previous grade (4 = A, 3 = B, etc.)	2.22	1.22
Class level (N = 48 classes)[b]		
General math/pre-algebra	0.17	0.38
Stretch Regents	0.17	0.38
Math A/B/UCSMP	0.38	0.49
Algebra	0.17	0.38
Regents	0.13	0.33
Content coverage	0.04	0.01
Class socioeconomic status	1.81	0.55

[a]Average of scores across three time points.
[b]Class types do not sum to 1.00 because of rounding.

TABLE 2
Content Coverage in Different Types of Classes

	Class type						
	General	Pre-algebra	Math A/B	UCSMP	Stretch Regents	Algebra	Rege
Level of coverage	0.046 (0.003)	0.053 (0.012)	0.063 (0.016)	0.059 (0.004)	0.068 (0.021)	0.082 (0.028)	0.0 (0.0
Configuration of coverage	0.566 (0.027)	0.538 (0.042)	0.598 (0.091)	0.715 (0.131)	0.555 (0.081)	0.580 (0.066)	0.5 (0.0
Content coverage (level × configuration)	0.026 (0.0003)	0.028 (0.006)	0.038 (0.012)	0.042 (0.010)	0.037 (0.012)	0.047 (0.012)	0.0 (0.0
Number of classes	2	6	16	2	8	8	6

Note: Figures are means, with standard deviations in parentheses.

estimate the proportion of students whose family incomes were low, lower-middle, middle, middle-upper, and high. These estimates were scaled 1–5 with 5 as high and are used as class-level indicators of student socioeconomic background. As Table 1 shows, the students in our sample were economically disadvantaged; the average of the class estimates was less than 2, indicating low to lower-middle family income for the typical student in our sample.

Methods

To examine achievement growth, we used a three-level hierarchical linear regression model (Bryk & Raudenbush, 1992). The first level measures individual achievement growth over time for each student. This level, known as the "within-student" level, allows achievement to vary for each student across the three time points. Because we had only three data points for achievement, our model assumes that achievement growth was linear over the course of the school year.

The second level of the model measures differences between students within classes. Here, each student's initial achievement and achievement growth are predicted as a function of student-level background characteristics, including gender, race, ethnicity, and previous mathematics grade.

The third level of the model estimates differences between classes. At this level, we predict differences among classes in average achievement growth. We present three versions of this model, adding more class-level variables to each succeeding version. First, we estimate a class-level baseline model in which class socioeconomic status is the only class-level variable. Second, we estimate a model of class-type effects, in which achievement growth varies among college-preparatory, transition, and general-track classes. Third, we add content coverage to this model to see whether differences in achievement growth across classes are explained by differences in content coverage.

All three levels of the model are estimated simultaneously, using the HLM/3L program (Bryk, Raudenbush, & Congdon, 1995). This procedure estimates separate error variances for each level, ensuring that parameters at the class level are not distorted because of similarities among students within classes (Bryk & Raudenbush, 1992). For ease of computation and interpretation, the student-level controls and class socioeconomic status are centered at their grand means (i.e., they are deviated from their total-sample means). Effects of student-level variables are treated as fixed across classes; only initial achievement and achievement growth are allowed to vary, both within and among classes. (Bryk & Raudenbush, 1992, provide further details on centering variables and fixing effects.)

Of the original 56 classes, we dropped from the sample 6 bilingual classes and 2 classes that used the Interactive Math Program because initial examination of results showed them to be outliers in instruction and achievement. Thus, 48 classes were included in our analyses. Because of the small number of classes, we collapsed the general and pre-algebra classes into a single low-track category, and we combined Math A, Math B, and UCSMP Math classes into a single category. We placed UCSMP Math with Math A because both emphasized revisions in pedagogy as well as an integrated curriculum.

After reducing the sample to 48 classes and dropping a few students with missing background data, our sample contained 498 students who were tested three times and an additional 384 who were tested twice. These 882 students constitute our sample for the analysis. At least three time points are needed to estimate the growth curve, but the estimation procedure accommodates cases for which data are available at two of the three time points (Bryk & Raudenbush, 1992). Initially we tried to include all students who were tested at least once, by imputing fall test scores when necessary, but students with imputed scores appeared substantially different from students with at least two scores present, so that approach was untenable, and we limited the analyses to students with at least two test scores.

Because students were not randomly assigned to classes, we are relying on our statistical model to purge observed achievement differences from pre-existing differences among students.[2] Two aspects of our model address this concern: First, we examine achievement growth, measured at three time points, rather than achievement differences at one or even two points in time. More time points improve the reliability of estimating achievement differences between class types. Variation in achievement resulting from unobserved differences among students should be reflected in the intercept (i.e., initial achievement in the fall) so that differential growth over time reflects differences among the classes, not pre-existing differences among students. Second, we controlled for background differences among students, including race, ethnicity, previous math grade, and (at the class level) socioeconomic status.[3] This procedure adjusts differences in average fall achievement and average achievement growth for demographic differences among students. Still, one cannot be certain that selection differences have been eliminated completely, so caution is warranted in interpreting the results.

Results

Table 3 presents the results of our three-level models. At the top of the first column of results ("Between-class baseline"), one can see that the average initial score was 9.716 and average achievement growth was 0.858 points per time period.[4] The between-student model shows no significant sex difference in initial score and no effects of previous grade. However, Blacks, Hispanics, and Asians all scored significantly lower than Whites (the omitted category) on the initial test. In the prediction of achievement growth, by contrast, one can see that the coefficients for the different race/ethnic categories are relatively small, and none are statistically significant. Either achievement inequality in mathematics had ceased to expand by this time or the time period of one year was too short to reliably measure changes in the achievement gap between Whites and other students. Similarly, in the between-class portion of the baseline analysis, socioeconomic status predicts initial score but not achievement growth.

In the between-class portion of the second column of results ("Course-type effects"), one can see differences among types of classes in average achievement growth. In comparison to students in Regents classes (the omitted category), students in general math and pre-algebra learned significantly less over the course of the year. Whereas achievement growth averaged 1.127 points per time period in Regents classes (see the top of the second column), growth was only 0.657 points (1.127 – 0.470 = 0.657) per time period in general math/pre-algebra. Growth for algebra, Math A/B/UCSMP, and Stretch Regents lies between Regents and general math—not significantly less than Regents, but, according to a post-hoc test, not significantly greater than general math, either.

Finally, the between-class portion of the third column of results shows effects of content coverage. First, one can see that achievement growth was greater in classes with more content coverage. Second, the gaps between Regents classes and other classes appear greatly diminished. In particular, the coefficient for general math/pre-algebra is reduced by almost half and is no longer statistically significant, suggesting that greater content coverage accounts for much of the advantage of Regents over general math/pre-algebra. The coefficients for Stretch Regents and Math A/B/UCSMP are also substantially reduced, suggesting that if coverage were more like that of Regents Math, achievement would be closer as well. This point is graphically illustrated in Figure 1. The first panel, based on the second column of Table 3, shows achievement growth for a hypothetical average student in five types of classes. The second panel, based on the third column of Table 3, simulates what achievement growth would be like for the hypothetical student in the different types of classes if content coverage were the same in all classes as we observed it to be in the Regents classes. As one can see, achievement growth would be more similar if content coverage were standardized, according to our model.

Discussion and Conclusions

This article shows that more rigorous content coverage distinguishes college-preparatory math classes from general-track math classes, and it also shows that, consistent with previous research, students learn more in the college-preparatory classes. Transition courses appear to be located in between, both in coverage and in achievement. On this basis, one may judge the transition courses a partial success in upgrading the quality of mathematics instruction for low-achieving high school students.

One conclusion seems clear: General-track math classes should be eliminated. Instruction is weak, achievement growth is shallow, and general math is a dead end for students' mathematics careers. Certainty about this conclusion must be tempered by the possibility that low achievement in general tracks may reflect unexamined characteristics of students rather than the course designation or curriculum, but the results strongly suggest that the course and curriculum are implicated in students' lower achievement growth compared to other classes.

Our study is less conclusive on the best alternative to the general track. On the one hand, transition courses appear to improve students' chances of learning and attainment. Content coverage is more rigorous, and our best estimate suggests that learning is greater, compared to the general track. These benefits are compounded because unlike the

TABLE 3
Class-Type Differences in Student Learning Over the Course of a Year[a]

	Between-class baseline	Course-type effects	Instructional effects
Within-student model			
Intercept (initial score)	9.716***	9.718***	9.722***
	(0.312)	(0.311)	(0.310)
Achievement growth	0.858***	1.127***	0.596*
	(0.065)	(0.177)	(0.311)
Between-student model			
Effects on initial score			
Male	0.298	0.285	0.286
	(0.269)	(0.270)	(0.270)
Black	−2.300***	−2.318***	−2.315***
	(0.388)	(0.390)	(0.390)
Hispanic	−2.092***	−2.097***	−2.100***
	(0.441)	(0.444)	(0.444)
Asian	−2.895***	−2.887***	−2.885***
	(0.476)	(0.479)	(0.479)
Previous grade	0.017	0.021	0.016
	(0.113)	(0.114)	(0.114)
Effects on achievement growth			
Male	0.102	0.110	0.111
	(0.128)	(0.129)	(0.129)[a]
Black	−0.112	−0.108	−0.108
	(0.180)	(0.183)	(0.183)
Hispanic	−0.025	−0.035	−0.026
	(0.200)	(0.204)	(0.203)
Asian	0.166	0.127	0.127
	(0.214)	(0.217)	(0.216)
Previous grade	0.085	0.082	0.087
	(0.054)	(0.055)	(0.055)
Between-class model			
Effects on average initial score			
Class socioeconomic status	1.064*	1.070*	1.076*
	(0.574)	(0.572)	(0.571)
Effects on average achievement growth			
Class socioeconomic status	0.153	0.081	0.051
	(0.119)	(0.124)	(0.123)
Algebra		−0.167	−0.178
		(0.214)	(0.212)
Stretch Regents		−0.264	−0.174
		(0.260)	(0.261)
Math A/B/UCSMP		−0.320	−0.242
		(0.212)	(0.212)
General math/pre-algebra		−0.470*	−0.251
		(0.252)	(0.268)
Content coverage			11.615*
			(5.795)
Residual variance in class average achievement growth	0.0123	0.0076	0.0037
Net between-class variance explained in average achievement growth[b]		0.381	0.703

[a]Standard deviations are in parentheses. $N = 2{,}262$ time points for 882 students in 48 classes.
[b]Proportion of variance explained net of between-class baseline model.
*$p < .10$. **$p < .05$. ***$p < .01$.

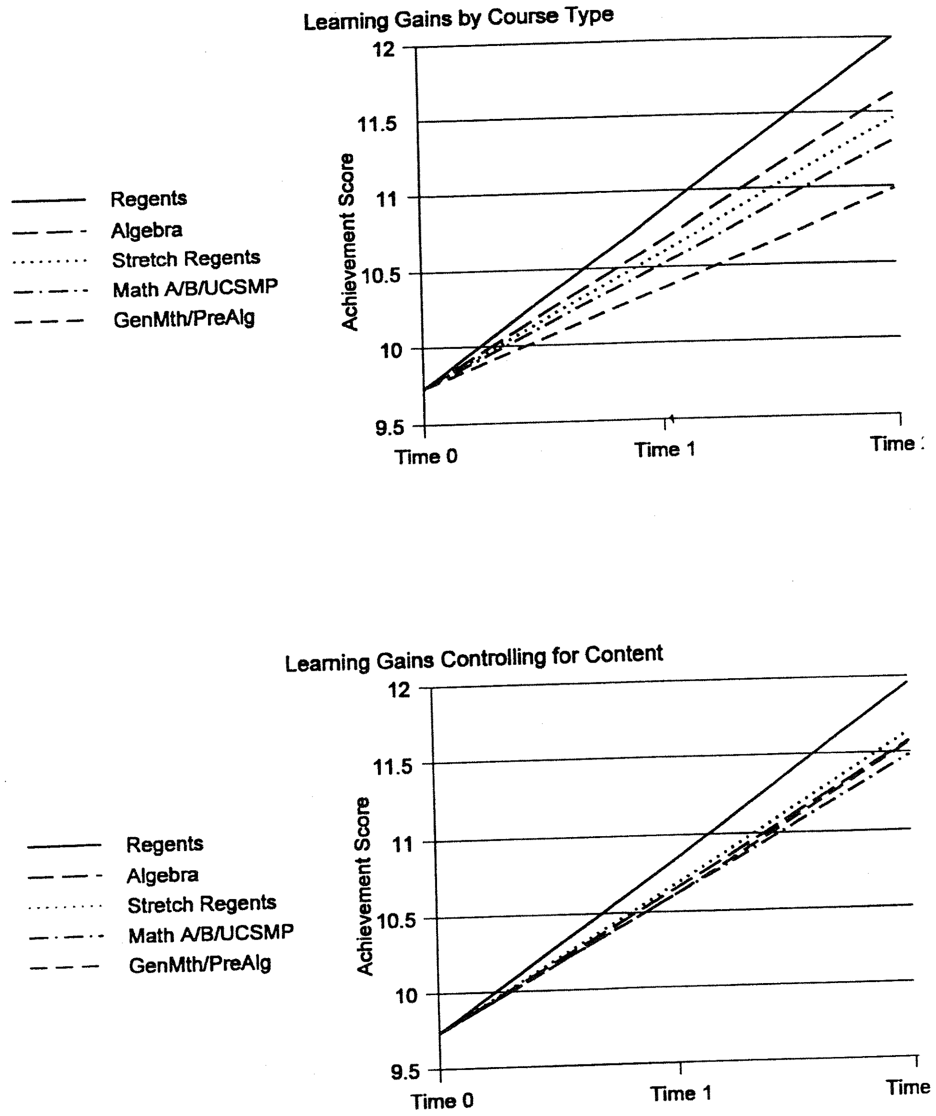

FIGURE 1. *Estimated learning gains.*

general track, transition courses are not a dead end. Math A, in particular, tends to lead students toward college-preparatory classes. In our study of high school transcripts, we found that almost half the students who entered Math A in ninth grade completed at least two years of college-preparatory math (typically algebra and geometry) by the end of four years of high school (White, Gamoran, Smithson, & Porter, 1996). (By contrast, less than a quarter of students who began in general math complete years of college-preparatory math.) Thus, stu who begin in Math A often move into alg where the learning trajectory may be even hi

On the other hand, our model cannot rule o claim that a given student learns the most wh or she is placed in a college-preparatory clas gardless of initial mathematics skills. This ment is often used as the basis for advocating tracking policy in which all students are placed in college-preparatory classes from the beginning of high school (e.g., Oakes, 1985; Rosenbaum, 1976). Although the teachers in our study criticized such proposals (White, Gamoran, Smithson, & Porter, 1996), it may be that they did not know how to provide a rigorous curriculum to an academically diverse group of learners. Clearly, this type of proposal needs to be attempted and evaluated, as is currently occurring in New York City and elsewhere. Observations reported by Porter et al. (1993) indicate that curriculum content does not necessarily deteriorate when academically diverse students are included in college-preparatory classes. A survey analysis by Gamoran and Hannigan (1997) further showed that all students gain from high school algebra, even those who enter high school with low test scores. Notwithstanding the difficulties, therefore, de-tracking warrants further implementation efforts.

This study and others have shown that low-achieving high school students are capable of learning much more than is typically demanded of them. The key is to provide a serious, meaningful curriculum: "hard content for all students" (Porter, Archbald, & Tyree, 1991). The normative, political, and technical problems of de-tracking have not been fully solved, but neither have we demonstrated that transition courses are as effective as college-preparatory classes for students with weak math skills. Hence, both de-tracking and transition courses must be considered viable policy options, worthy of continued implementation and evaluation. Either way, dead-end math classes should be replaced with one of these alternatives.

Notes

Research for this article was supported by the Consortium for Policy Research in Education, which is supported by the U.S. Office of Educational Research and Improvement (Grant No. OERI-R117-G10007) and by the Wisconsin Center for Education Research, University of Wisconsin-Madison. Portions of the work were presented in seminars at the University of Wisconsin, the University of Chicago, Tel Aviv University, and RAND and at the 1996 meeting of the American Educational Research Association. The authors are grateful for reactions received in those forums and for comments from Steve Raudenbush and from the *EEPA* reviewers on earlier versions of the article. Opinions and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the supporting agencies.

[1]As expected, the achievement level of students in our sample was low compared with national norms. Whereas students in our sample, mainly 9th- and 10th-graders, averaged about 10 on the pretest, the national average among 8th-graders for these items was 11.8 (National Center for Education Statistics, 1992).

[2]Although there was considerable overlap in student achievement among students assigned to the different types of classes, average fall score differed by class type, ranging from 7.53 in general/pre-algebra classes to 8.78 in Math A/B/UCSMP classes, 10.79 in Stretch Regents, 11.73 in algebra classes, and 13.17 in Regents classes. The difference between Regents and general/pre-algebra classes is about one and one quarter times the standard deviation on the fall test—not a trivial amount but smaller than the two-standard-deviation difference commonly found between the highest and lowest class types in studies of tracking (Slavin, 1990). Because class assignment is, in part, a consequence of fall achievement but not a cause, we include class type as a predictor of achievement growth but not as a predictor of fall achievement.

[3]In preliminary work summarized by White, Porter, Gamoran, and Smithson (1996), we also included initial achievement as a predictor of growth to control for regression to the mean in achievement. However, this procedure placed initial status on both the left side (as the first point on our growth line) and the right side (as a predictor) of our multi-level equation. Hence, we have omitted initial achievement as a predictor from the current analyses. Results for effects of class type and curriculum coverage were the same in either case.

[4]Because the student background variables are centered, these figures are averages for all students.

References

Barr, R., & Dreeben, R. (1983). *How schools work.* Chicago: University of Chicago Press.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1995). *Hierarchical linear modeling with the HLM/2L and HLM/3L programs.* Chicago: Scientific Software International.

California State Department of Education. (1985). *Mathematics framework for California public schools: Kindergarten through grade twelve.* Sacramento, CA: Author.

Clune, W. H., & White, P. A. (1992). Education reform in the trenches: Increased academic course taking in high schools with lower achieving students in states with higher graduation requirements. *Educational Evaluation and Policy Analysis, 14,* 2–20.

Conant, J. B. (1959). *The American high school today.* New York: McGraw-Hill.

Gamoran, A. (1987). The stratification of high school learning opportunities. *Sociology of Education, 60,* 135–155.

Gamoran, A. (1993). Alternative uses of ability grouping in secondary schools: Can we bring high-quality instruction to low-ability classes? *American Journal of Education, 101,* 1-22.

Gamoran, A., & Hannigan, E. (1997, March). *Algebra for everyone? Benefits of college-preparatory mathematics for students with diverse abilities in early secondary school.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Gamoran, A., & Nystrand, M. (1994). Tracking, instruction, and achievement. *International Journal of Educational Research, 21,* 217–231.

Gamoran, A., Nystrand, M., Berends, M., & LePore, P. C. (1995). An organizational analysis of the effects of ability grouping. *American Educational Research Journal, 32,* 687–715.

Gamoran, A., & Weinstein, M. (in press). Differentiation and opportunity in restructured schools. *American Journal of Education.*

Grossman, P., & Stodolsky, S. S. (1995). Content as context: The role of school subjects in secondary school teaching. *Educational Researcher, 24*(8), 5–11, 23.

Hallinan, M. T. (1996). Track mobility in secondary school. *Social Forces, 74,* 983–1002.

Hoffer, T. (1992). Middle school ability grouping and student achievement in science and mathematics. *Educational Evaluation and Policy Analysis, 14,* 205–228.

Hoffer, T., & Gamoran, A. (1993, August). *Effects of instructional differences among ability groups on student achievement in middle-school science and mathematics.* Paper presented at the annual meeting of the American Sociological Association, Miami, FL.

Kerckhoff, A. C. (1986). Effects of ability grouping in British secondary schools. *American Sociological Review, 51,* 842–858.

Loveless, T. (1994). The influence of subject areas on middle school tracking policies. *Research in Sociology of Education and Socialization, 10,* 147–175.

Massell, D., & Fuhrman, S. (with Kirst, M., Odden, A., Wohlstetter, P., Carver, C., & Yee, G.). (1994). *Ten years of state education reform, 1983–1993: Overview with four case studies.* New Brunswick, NJ: Consortium for Policy Research in Education.

National Center for Education Statistics. (1992). *National assessment of educational progress 1990 assessment public release items.* Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (1994). *Curricular differentiation in public high schools* (Document No. NCES 95-360). Washington, DC: U.S. Department of Education.

National Commission on Excellence in Education. (1983). *A nation at risk.* Washington, DC: U.S. Gov ernment Printing Office.

National Council of Teachers of Mathematics. (1989 *Curriculum and evaluation standards for school mat ematics.* Reston, VA: Author.

Oakes, J. (1985). *Keeping track: How schools structu inequality.* New Haven, CT: Yale University Press.

Oakes, J. (1990). *Multiplying inequalities.* Santa Monic CA: RAND.

Oakes, J. (1994). More than misapplied technology: normative and political response to Hallinan on trac ing. *Sociology of Education, 67,* 84–89.

Oakes, J., Gamoran, A., & Page, R. N. (1992). Curricu lum differentiation: Opportunities, outcomes, an meanings. In P. W. Jackson (Ed.), *Handbook of r search on curriculum* (pp. 570–608). New Yor Macmillan.

Porter, A. C., Archbald, D., & Tyree, A. (1991). Reform ing the curriculum: Will empowerment policies r place control? In S. H. Fuhrman & B. Malen (Eds. *The politics of curriculum and testing. 1990 yearbo of the Politics of Education Association* (pp. 11–36 London: Taylor and Francis.

Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, L., & Schneider, S. A. (1993). *Reform up close: classroom analysis.* Madison, WI: Wisconsin Cent for Education Research.

Rosenbaum, J. E. (1976). *Making inequality: The hi den curriculum of high school tracking.* New York Wiley.

Slavin, R. (1990). Achievement effects of ability group ing in secondary schools: A best-evidence synthesi *Review of Educational Research, 60,* 471–499.

Wheelock, A. (1992). *Crossing the tracks: Ho untracking can save America's schools.* New York The New Press.

White, P. A. (1995). *Math innovations and classroon practice: Upgrading of the math curriculum at th high school level.* Madison, WI: Consortium for Polic Research in Education.

White, P. A., Gamoran, A., Smithson, J., & Porter, A. C. (1996). Upgrading the high school mathematics curriculum: Math course-taking patterns in seven high schools in California and New York. *Educational Evaluation and Policy Analysis, 18,* 285–307.

White, P. A., Porter, A. C., Gamoran, A., & Smithson, J. (1996). *CPRE policy brief: Upgrading high school math: A look at three transition courses.* New Brunswick, NJ: Consortium for Policy Research in Education.

Wilson, B. L., & Rossman, G. B. (1993). *Mandating academic excellence.* New York: Teachers College Press.

Authors

ADAM GAMORAN is a professor of sociology and educational policy studies at the Department of Sociology, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706. His specialties are the sociology of education, school organization, and inequality in school systems.

ANDREW C. PORTER is a professor and director of the Wisconsin Center for Education Research at the University of Wisconsin-Madison, Madison, WI 53706. He specializes in research on teaching, educational policy analysis, student and teacher assessment, and psychometrics (especially the problem of measuring change).

JOHN SMITHSON is a project assistant at the Wisconsin Center for Education Research, University of Wisconsin-Madison, Madison, WI 53706. His specialties are measures of the enacted curriculum and collaborative research methodologies.

PAULA A. WHITE is the project manager at the National Institute for Science Education, University of Wisconsin-Madison, 1025 W. Johnson Street, Madison, WI 53706. She specializes in educational policy, curriculum reform, and effective teaching practices.

Manuscript received December 3, 1996
Revision received July 15, 1997
Accepted August 1, 1997

Educational Evaluation and Policy Analysis
Winter 1997, Vol. 19, No. 4, pp. 339–353

Students, Schools, and Enrollment in Science and Humanity Courses in Israeli Secondary Education

Hanna Ayalon and Abraham Yogev
Tel Aviv University

This article examines the deteriorating status of the humanities and social sciences versus mathematics and the sciences in the curriculum of Israeli high schools. We examine this tendency by conducting a multi-level analysis of the effect of school and individual characteristics on inequality in curriculum specialization on a sample of academic-track 12th-graders in 1989. The main findings are (a) more able students, males, and members of the privileged Jewish ethnic group in Israel tend to specialize in mathematics and the sciences, and (b) students' characteristics are the major determinant of course-taking in mathematics and the sciences, whereas school policy is central regarding the humanities and social sciences. The article discusses social implications of the findings.

The centrality of mathematics and the sciences constitutes a major feature of the curriculum of secondary education in many western industrial societies (Apple, 1990; Kamens & Benavot, 1992). The research treats this centrality mainly by studying the time allocated in the curriculum to these school subjects compared to the humanities and social sciences. The relative share of mathematics and the sciences in the curriculum increases with time, while that of the humanities decreases; this is interpreted as an indication of changes in the status of these areas of study in the hierarchy of school subjects: The prestige and social esteem of mathematics and the sciences are improving while those of the humanities are deteriorating (see Kamens & Benavot, 1992, for a global comparison; Goodson, 1983, for Britain; Kliebard, 1992, for the U.S.; Mitter, 1995, for Germany; Morris, 1995, for Hong Kong).

Although we find expressions of intellectual concern with the deteriorating prestige of the humanities (Bloom, 1987; Hirsch, 1988; Iram, 1995; Kliebard, 1992), the empirical research on their status in the school curriculum is still limited. Indicators of the status of the different subject areas, beyond differential allocation of time, have hardly been studied in primary and secondary education. One exception is the study of Morris (1995), who reported that in Hong Kong mathematics, the sciences, and languages are given special weight in calculating the grades of students in secondary education. More evidence exists on the differential status of these areas of study in higher education. A recent OECD comparative project indicates that in Japan only a minority of the students of the humanities and social sciences attend prestigious state universities and that in France, Germany, and the Netherlands, the humanities belong to the open, non-selective sector of the universities, which is associated with low employment prospects (Raivola, 1995).

In this article, we wish to contribute to the research of the status of the two subject areas by referring to the social profile of the students who study them in Israeli secondary education. We focus on this aspect because the social profile of the students who study the various school subjects is considered an inherent part of the stratification of these subjects. Higher-status knowledge is offered to and preferred by higher-status students, who usually do better in school. The ability and social identity of the students, in turn, reflect on the subjects' status (Goodson, 1983).

Israel is a particularly appropriate arena for studying the matching between students and the two subject areas because of a special characteristic of the