Alternative Uses of Ability Grouping in Secondary Schools: Can We Bring High-Quality Instruction to Low-Ability Classes?

ADAM GAMORAN
University of Wisconsin—Madison

Despite extensive criticism, ability grouping remains a widespread practice in American secondary schools. This paper considers whether ability grouping can be implemented more effectively than is typical; in particular, it explores possible instances of high-quality instruction in low-ability classes. Data from a study of eighth- and ninth-grade English classes yield two examples of schools with apparently effective instruction in low tracks. These cases are characterized by (1) high expectations by teachers, manifested by a refusal to relinquish the academic curriculum as commonly occurs in low-track classes, (2) extra exertion by teachers to foster extensive oral discourse in class, and (3) no system of assigning weak or inexperienced teachers to lower tracks.

Tracking will remain an important part of American education. . . . Neither tracking nor heterogeneous grouping is necessarily good or bad. The effectiveness of grouping depends on the specific situation and the needs within a school. [Excerpts from the Conclusions of a Report on Academic Tracking (National Education Association 1990, pp. 27–28). Emphasis in the original.]

What shall we do about ability grouping in secondary schools? Surely there is no other current educational practice that has been challenged from as many quarters. Recent years have seen critiques founded on legal, historical, philosophical, sociological, anthropological, economic, and psychological bases. Grouping and tracking have always been debated, but the controversy appears especially heated at the present time.
Ability Grouping

Many schools and districts are reconsidering long-standing programs of grouping and tracking, searching for other means of coping with diversity in their student populations. Other communities have not attempted such sweeping changes. According to the National Education Association (1990), without support for smaller classes and intensive staff development, many educators view ability grouping as the only workable approach to instructing secondary students with different skills. Albert Shanker (1993), president of the American Federation of Teachers, has recently stated a similar position: “We should not mandate either homogeneous or heterogeneous grouping but work to develop effective models of both.”

In light of these views, it is as important to consider alternative uses of ability grouping as it is to discuss alternatives to ability grouping. Does ability grouping always work out badly? Under what circumstances, if any, can ability grouping be successful? This article has two purposes. The first is to show that ability grouping in secondary schools does not always have the same effect, and therefore it is worth seeking ways of using it more effectively than commonly occurs. A brief review of earlier studies, and a reinterpretation of the conclusions of an earlier synthesis, provide the support for this claim. The second goal is to explore instances of relatively successful uses of ability grouping, in the sense that high-quality instruction fosters significant learning among students assigned to low-ability classes. What characterizes such classes? To address this question, I draw on evidence from earlier studies by other authors, and I provide two new illustrations taken from a larger study of eighth- and ninth-grade English classes in 25 midwestern schools. Although these examples are far from conclusive, common elements emerge that, taken together, may help to characterize effective instruction in low-ability classes in secondary schools.

Varied Effects of Ability Grouping on Achievement

A common conclusion about ability grouping is that it raises achievement for students in high-level classes but depresses achievement for...
those in positions of lesser status, resulting in widening inequality between high and low achievers over time. For example, many studies of high school tracking in the United States show that, as the years pass, the gap in achievement increases between students in academic programs and those in nonacademic programs (see, e.g., Gamoran and Mare 1989). These studies document growth in inequality, but they cannot say whether widening gaps occur because tracking helps students in academic programs or harms those in other programs or some combination of the two. This issue is addressed in a British study of ability grouping. After studying a national sample of students from age 11 to age 16, Fogelman (1983) and Kerckhoff (1986) reported that, compared with students in mixed-ability classes, gains for students in high-ability classes were offset by losses among students in low-ability classes. Hence, ability grouping yielded no effect on average achievement, but it contributed to an increase in inequality of achievement.

Not all studies of ability grouping in secondary schools, however, follow this pattern. Some studies show no growth in inequality, and some even show a decline. Many report no impact on average achievement, but some indicate a rise, while others reveal an overall drop. Virtually every conceivable pattern of results—favoring high achievers, favoring low achievers, favoring those in the middle, and so on—can be observed in one study or another. As Slavin (1990) demonstrated after a thorough review, varying effects on achievement at the secondary level cannot be attributed to differences in the types of studies conducted (e.g., experimental vs. naturalistic) or to the different forms of ability grouping (e.g., grouping on a subject-by-subject basis vs. grouping for all subjects at once). Studies of curriculum tracking are more consistent in showing a growth in inequality of achievement, but even this conclusion is not universal (Jencks and Brown 1975; Alexander and Cook 1982). The diversity of findings in studies of secondary-school ability grouping, which centers around zero, led Slavin to conclude that the true effects on achievement of ability grouping in secondary schools, all else held constant, are indeed zero: “Comprehensive between-class ability grouping plans have little or no effect on the achievement of secondary students, at least as measured by standardized tests. . . . Assigning students to different levels of the same course has no consistent positive or negative effects on students of high, average, or low ability” (Slavin 1990, p. 494). According to this view, observed differences among studies in the effects of grouping are due to chance; taken together, the studies indicate that no real effects on achievement exist.
Ability Grouping

Another interpretation seems equally plausible: The inconsistent findings may have resulted from uncontrolled differences in the way ability grouping was implemented in the various school systems under investigation. Suppose some schools used ability grouping as a way of stratifying teachers as well as students (providing successful students with access to teachers with the best reputations), but other schools distributed teachers across tracks more equitably. What if, in some experiments with ability grouping, high, average, and low classes covered similar material, but in other experiments, high classes moved along as fast as teachers thought they could, while low classes were held back? Perhaps some schools used ability grouping to remediate low-achieving students, while others directed most of their resources toward their strongest students. Unfortunately, the studies provided little information on what was actually going on in the classes once students were assigned. As Slavin (1990, p. 493) acknowledged, “In none of the studies reviewed here were there systematic observations of teaching and learning.”

This interpretation does not dispute the conclusion that, if all else were held constant, ability grouping would have no effects. It suggests, however, that ordinarily, especially in nonexperimental situations, little else is held constant. A host of observers, for example, has claimed that instructional conditions are typically better in high-group classes and inferior in classes of lower rank. If the quality of instruction were invariant, ability grouping might benefit (or harm) all students equally, but if the quality of instruction varied along with the levels of the grouping system, ability grouping would serve some students well but do ill to others. Hence, the impact of ability grouping may depend, at least in part, on how it is used to distribute instruction to students in different classes.

If inconsistencies in the effects of ability grouping are substantively meaningful, and not just statistical artifacts, then it is worth asking under what conditions the losses are lessened and the gains increased. A useful starting point is to ask whether high-quality instruction can occur in low-ability classes. Most of the harmful effects of ability grouping can be traced to low-ability classes: When inequality increases, achievement in low-ability classes is typically dropping, and when average achievement falls, the steepest decline often occurs in the lowest level of the system. Understandably, critics of ability grouping are most concerned with poor conditions in low groups (see, e.g., Oakes 1985). First, I will summarize prior research indicating that instruction is typically inferior in low-ability classes. Second, I will briefly show that new data from a study of midwestern secondary schools mainly conform to this pattern. Third, I will give four examples—two drawn
from past research, and two original cases taken from the study of midwestern secondary schools—that illustrate that high-quality instruction can occur in low-ability classes. Finally, I will consider the limitations and implications of these illustrations.

Ability Grouping and Classroom Instruction

Previous research provides details on how classroom instruction varies among ability groups. The most comprehensive examination was carried out by Oakes (1985), who, following a national study of junior and senior high schools, described differences in curriculum content, instructional activities, and classroom climate. In English, for example, students in high-track classes read "standard works of literature," while low-track students typically read "young-adult fiction" (p. 76). High-track classes required expository writing and critical thinking, but low-track classes emphasized memory and comprehension. These differences in exposure to knowledge, Oakes argued, prevent low-track students from gaining access to high-status knowledge and to knowledge that is valued at higher levels of the education system. High-track classes also devoted a greater proportion of time to instructional activities and required more time spent on homework outside of class. Students in low-track classes were more likely to be off task, and more time in low- than in high-track classes was spent managing students' behavior.

In a subsequent nationwide study, Oakes (1990) noted that low-track classes in mathematics and sciences were taught by less experienced and less well qualified teachers than classes in other tracks. This finding echoed earlier high school case studies, which reported that teachers with less experience and weaker reputations were often assigned to teach lower-track classes (see, e.g., Rosenbaum 1976; Finley 1984).

Low-Ability Classes as Caricatures

Some researchers have concentrated on describing the difficulties of low-ability classes. Page (1987) was initially struck by the surface similarity of low- and regular-track history classes in a well-off high school in a professional community. On closer inspection, however, she came to characterize the low-track classes as "caricatures" of the regular ones. Despite what appeared to be an academic curriculum, thematically similar to that of the regular track, instruction in low-track classes
Ability Grouping

deehumanized academic concerns. Teachers and students believed low-track students could not learn, and teachers were not held accountable for the learning of low-track students. In one class, the teacher communicated these low expectations by offering entertainment rather than a serious academic topic and a sense of purpose. Page commented (1987, p. 453): “Academic progress is the least important aspect of Mr. Ellison’s class. His work sheets offer ‘games,’ rather than lessons, to motivate students who [purportedly] cannot learn.” Instruction in this class was highly fragmented, as classroom events were punctuated by frequent, sometimes prolonged interruptions, and knowledge was defined by daily work sheets. In another low-track history class in this school, Page (1989) observed that ambiguity about the purpose of instruction led the teacher to forestall substantive debate, whereas academic disagreements were taken seriously and encouraged in regular-track classes.

Despite these findings, Page did not claim that all low-track classes are academically ineffective. Rather, she argued that low-track classes are “versions” of regular classes, exhibiting many similarities but differing in important, yet often subtle, ways (Page 1991). In the community she studied, the low-track version of regular classes was a caricature, but Page left open the possibility of other versions in other settings.

Ability Grouping and Instruction in a Sample of Midwestern Secondary Schools

A recent study of ability grouping in a sample of midwestern secondary schools offers data that, on the whole, conform to the patterns described by Oakes, Page, and other researchers (for reviews, see Gamoran and Berends [1987] and Murphy and Hallinger [1989]). Later I will highlight exceptions to the general pattern, but at this point I will briefly describe the main tendencies.

For the study, we selected nine high schools that varied in the context of their communities, including rural, urban, and suburban schools, and public and Catholic schools. To follow students from eighth to ninth grade, we began with 16 junior high or middle schools that served as feeders for the high schools. Of the 112 English classes we studied over the two-year period, 42 in eighth grade and 50 in ninth grade were ability grouped, that is, students were assigned to them on the basis of prior performance (see further Gamoran [1992a]). The remaining classes are omitted from analyses for this article, since the present purpose is to compare alternative uses of ability grouping.
rather than alternatives to grouping. Levels of ability-grouped classes were defined by the schools (e.g., honors, regular, remedial), and there was little ambiguity about which class was which. Most schools using grouping had three levels, but some smaller schools had two, and others had as many as five. In four cases, teachers divided their time between two groups in one room.

Students in each class completed tests and questionnaires in the fall and spring. Teachers responded to a questionnaire and an interview at the end of the year and kept a weekly log of texts covered from October through April. Observers visited each class four times, focusing on classroom activities (time spent on discussion, seat work, etc.) and on classroom discourse, coding the questions asked on a variety of discourse categories (for details, see Nystrand and Gamoran [1991]).

Table 1 provides examples showing that, on average, the instructional conditions we observed were similar to what others have reported. Not only did high-track classes read more long works of fiction and low-track classes fewer, but, just as Oakes described, readings in high-track classes consisted of "standard works of literature," whereas low-track readings could often be considered "young-adult fiction."2 Similar to Page's findings, the results of our study showed that low-track classes filled in blanks more often than they wrote sustained essays, and the opposite was true in high-track classes. Low-track students reported spending less time on homework and completing fewer of their written assignments than students in other classes. We also observed more off-task behavior in low-track classes. High-track classes devoted more time to recitation and discussion and less time to seat work than other classes, but low-track and regular classes did not differ from one another in time spent in these activities. Moreover, we observed higher proportions of open-ended questions in regular classes, and other aspects of teacher questions varied inconsistently across tracks (see further Gamoran et al. 1992).3 Despite these interesting discrepancies, on the whole our findings conformed to the general picture of less serious, less demanding, and less stimulating instruction in low-track classes.

It is also interesting to note that many of the standard deviations in table 1 are large compared with their corresponding means. Although the differences between groups follow a persistent pattern, most differences are small in relation to the variability around the means. Oakes (1985) had also reported relatively small differences between high and low tracks despite a pattern that was consistent overall. Just as the achievement effects of grouping may vary, instructional conditions at a given track level appear to differ from one class to another.
TABLE 1

Means (with Standard Deviations in Parentheses) of Selected Reading, Writing, and Class-Work Activities among Honors, Regular, and Remedial Classes in Eighth- and Ninth-Grade English

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>TYPE OF CLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Honors/Advanced</td>
</tr>
<tr>
<td>Reading:</td>
<td></td>
</tr>
<tr>
<td>Standard literature (number of novels and dramas, teacher reported)</td>
<td>2.4 (1.3)</td>
</tr>
<tr>
<td>Young-adult fiction (number of novels and dramas, teacher reported)</td>
<td>.2 (.6)</td>
</tr>
<tr>
<td>Total novels and dramas</td>
<td>2.6 (1.4)</td>
</tr>
<tr>
<td>Writing:</td>
<td></td>
</tr>
<tr>
<td>Fill in blanks (times per month, teacher reported)</td>
<td>1.3 (2.2)</td>
</tr>
<tr>
<td>Write one page or more (times per month, teacher reported)</td>
<td>2.4 (3.9)</td>
</tr>
<tr>
<td>Percentage of writing completed (student reported)</td>
<td>91.5 (6.4)</td>
</tr>
<tr>
<td>Homework time (hours per week, student reported)</td>
<td>1.5 (.7)</td>
</tr>
<tr>
<td>Class work:</td>
<td></td>
</tr>
<tr>
<td>Open-ended questions (percentage observed)</td>
<td>16.6 (11.9)</td>
</tr>
<tr>
<td>Question/answer and discussion time (minutes per day observed)</td>
<td>17.9 (8.2)</td>
</tr>
<tr>
<td>Seat-work time (minutes per day observed)</td>
<td>11.4 (12.0)</td>
</tr>
<tr>
<td>Students off task (percentage observed)</td>
<td>2.0 (2.5)</td>
</tr>
<tr>
<td>Number of classes</td>
<td>24</td>
</tr>
</tbody>
</table>

Note.—Figures reflect class-level data, averaged across classes of each type. Student-level items (writing completed and homework time) were averaged by class before computation of class-type means. Because of teacher nonresponse, the number of regular classes was 43 for "write one page or more" and the number of remedial classes was 23 for "fill in blanks."
Alternatives

There is enough variability in past research to suggest that, although the pattern of lower achievement and weaker instruction in low tracks is common, it may not be inevitable. Under what circumstances, if any, does a different pattern emerge? In Slavin’s (1990) review, two of the studies carried out since 1960 reported positive effects of grouping on achievement for low-group students. Do such findings result from measurement error, or are there systematic conditions that may bring them about? Have any versions other than caricatures been observed for low-track classes?

Studies of Catholic High Schools

Evidence for alternative uses of ability grouping comes from observational studies of Catholic high schools. In a study of three urban schools, Valli (1986, p. 29) found that “a challenging learning environment was prevalent at all track levels.” Unlike other observational studies, Valli’s study found both students and teachers speaking favorably about lower-track classes. Students believed remedial classes were beneficial for their long-term progress. In contrast to the frustration reported in other studies, teachers of low-track classes believed they could be successful. “Nowhere did we hear teachers say, ‘I don’t know how to reach this type of student’” (Valli 1986, p. 26).

On the basis of a follow-up study at one of the schools, Valli (1990) described instruction in the lower track as a “parallel curriculum,” in an explicit contrast with Page’s (1987) characterization of low-track classes as caricatures. Observers found “teachers determined to make lower-track classes as educationally rigorous as upper-track classes” (Valli 1990, p. 58). Classes were smaller, and teachers reported working harder, to provide more individualized attention to students in lower-track classes. School policy directed stronger and more experienced teachers to lower-track classes. Overall, the school climate stressed effort and caring. According to Valli, this climate supported instruction in lower-track classes by expressing high expectations and by requiring close monitoring of students’ progress.

Comparing the tracking systems in three public and four Catholic high schools, Camarena (1990) also reported more effective uses of grouping in Catholic schools. Echoing staff in Valli’s schools, teachers expressed high expectations for low-track students. As one English department chair put it, “I believe remedial kids can learn anything.
Ability Grouping

It just takes longer. I think they should use the same book. . . . The method of teaching should be different" (Camarena 1990, p. 176). Camarena found that, unlike the public schools in her study, the Catholic schools had academic standards that all students were expected to master, regardless of track level. Tracking in the Catholic schools also emphasized flexibility in assignment. Guidance counselors played a prominent role in implementing this policy, monitoring students and advising them on the steps they needed to take to progress. The comparability of instructional content across tracks also made shifting tracks feasible.

Survey research on public and Catholic schools is consistent with the claim that Catholic high schools use tracking more effectively. Students in Catholic schools take more academic courses, regardless of track, and the academic demands in non-college-track courses are greater in Catholic than in public schools (Hoffer et al. 1985; Lee and Bryk 1988). Catholic schools produce smaller achievement gaps between tracks (at least in math), and they do so in a context of higher overall achievement (Gamoran 1992b).

These studies of Catholic schools suggest that there is variability in the implementation of grouping and tracking and that such differences are tied to student outcomes. With this conclusion in mind, we searched our data for examples of effective uses of ability grouping for students in low-level classes. My purpose here is not to test for variation in the effects of ability grouping. Rather, supposing that such variation occurs, my aim is to illustrate what goes on inside the classrooms of schools that make effective use of ability grouping.

Successful Low-Ability Classes in the Sample of Midwestern Secondary Schools

We examined the 18 schools that used ability grouping for English, seeking examples of schools that had (a) high-quality instruction in low-track classes, considering both curriculum content and student-teacher interaction, and (b) higher-than-expected achievement on a year-end literature test among students in low-track classes. We addressed the first criterion by searching for cases in which works of literature were taught through extensive oral interaction and written work. We addressed the second by inspecting residuals from regressions on literature achievement predicted by student characteristics: sex, race and ethnicity, socioeconomic status, and prior reading and writing skills.5 One school met both criteria for eighth grade, and one did for ninth. Although we did not intentionally focus on the five
Catholic schools that used ability grouping, in light of previous research it is not surprising that our two exemplary cases are both Catholic schools.

St. Elias is a small K–8 parish school. Although the school is located in a suburb, its middle-class student body comes from a neighboring city as well as from the suburb itself. There were two eighth-grade classes in St. Elias, and, for English, students were divided according to their performance on standardized tests. Overall, students in St. Elias scored more than four points above their expected achievement on our spring test of literature mastery (scores ranged from two to 32). Students in the low-ability class alone averaged 3.5 points above their predicted scores. Their teacher, Mrs. Grant, had spent seven of her eight years in the profession at St. Elias.

Immaculate High School is located in an urban area, in a city different from the one next to St. Elias. Like St. Elias, Immaculate is coeducational, but it is affiliated with a religious order rather than a parish. Students at Immaculate tend to be well-off economically, though there are exceptions. At both schools, over 95 percent of the students are white. There were six ninth-grade English classes at Immaculate the year we visited. One was an honors class, which included students who scored above the eighty-third percentile on a standardized reading test. Another was a remedial class, which included students who scored below the fortieth percentile. The remaining four classes were called “regular.” Mrs. Turner, a second-year teacher, taught the honors class, the second semester of the remedial class, and several regular sections as well. During the first semester, the remedial class was taught by Mrs. Beatty, who is considered the school’s reading specialist. On our test of literature mastery, students at Immaculate averaged about three points above what would be expected on the basis of their background and fall achievement, and, relative to their starting points, students in the remedial class scored just as high as students in other classes.

We observed a number of instructional similarities in the way ability grouping was used in the two schools. First, in both schools the same teacher taught both high- and low-track classes. Remedial classes were not used as “dumping grounds” for teachers or for students. Unlike other cases reported in the literature, these schools did not have a system of assigning less experienced or less successful teachers to the lower-level classes.

Second, the teachers implemented similar literature curricula across tracks in both schools. At St. Elias, Mrs. Grant used the same set of readings with both classes; these included Animal Farm and The Glass Menagerie, as well as condensed versions of Hamlet and Macbeth. At Immaculate, all classes read Romeo and Juliet and To Kill a Mockingbird,
as well as other texts. Although the remedial class at Immaculate read fewer novels than the regular and honors classes, readings at all levels would be considered standard works of literature rather than young-adult fiction, just as at St. Elias. This pattern contrasts sharply with Oakes's finding that "low-track English classes rarely, if ever, encountered these kinds of knowledge. . . . Not only did they not read works of great literature, but we found no evidence of good literature being read to them or even shown to them in the form of films" (1985, p. 76).

Third, observations in both schools revealed a preference for spending class time on oral work, in all sections. In St. Elias, both classes averaged close to 17 minutes per day in recitation and discussion. In Immaculate, the low-track class averaged over 23 minutes in recitation and discussion, while the honors and regular classes averaged 19 and 17 minutes, respectively, in oral activities.

Observations and interviews indicated that Mrs. Grant, of St. Elias, and Mrs. Turner, of Immaculate, shared characteristics that Valli (1990) and Camarena (1990) linked to successful uses of ability grouping. Both teachers expressed high expectations for students in all sections, including the lower-level classes, and both showed that success with low-track classes requires special effort on the teacher's part. These views differ greatly from the opinions of teachers in Page's (1987) study. They also contrast with Rosenbaum's (1976) finding that teachers typically prepare less and put forth less effort for low-track classes.

According to Mrs. Grant, "I know I have to put out more for (the low-track class), and that's often very difficult for me to do. . . . I really feel tired when I'm through. . . . I also have to give them more encouragement as to, 'The only thing that's a stupid answer is the answer that's left unsaid.' You know, [to] get them to open up and say something." In the low-track class, she finds it necessary to present a more structured lesson and to monitor the students more carefully to keep them on track. Regarding group work, she says, "I have to float constantly among that lower group, where[as] I can let the other group go a little bit . . . and know that they're still on the right track." Mrs. Grant uses the blackboard more with the lower-track class, writing down what she and the students have said and rephrasing their ideas at times. Compared with the higher class, in the low-track class she spends less time in small groups, and, during whole-class instruction, she presents more examples and tries to draw on students' personal experiences more often. In addition, she takes time outside of class as well as during class to meet individually with students. "I have to schedule the lower group at least once a week . . . individually to talk to . . . each student." Grading standards were the same in the two classes.

12 American Journal of Education
Mrs. Turner’s exertion with the low-track class was evident in one lesson, a relentless attempt, despite some resistance from students, to provide them with a background for reading *To Kill a Mockingbird*. After explaining conditions of racism in the South during the 1930s, she asked:

*Mrs. Turner.* Now, let me just ask you, ’cause this word’s gonna come up a lot in *To Kill a Mockingbird*, how do you guys feel when you hear the word “nigger”?

*Student.* It’s stupid.

*Mrs. Turner.* Um-hm, it’s stupid. [Pause] No other response?

*Student.* It’s like a childish thing, to say that.

*Mrs. Turner.* It’s childish, all right.

Her tone was not condescending, but collegial; she frequently used phrases such as “you guys” and “folks” in an attempt to reach out to students. Asking students how they felt reflected her attempt to personalize the story for them and suggested that she had something to learn from them. Not all students took the question seriously, however.

*Student.* I think of a booger. [Laughter]

*Mrs. Turner.* Well, all right, that’s

*Student.* I mean, I don’t think of a black person . . . I just think of “nigger,” “booger,” [it] rhymes, you know.

*Student.* [Sarcastically] Good analysis.

*Mrs. Turner.* Well, ok, let me just, let me just say this.

*Student.* [To Mrs. Turner] You wouldn’t say that [nigger] if there was a black person in here.

*Mrs. Turner.* Would I say that?

*Student.* They call themselves

*Mrs. Turner.* I’m quoting! Look, Bill and Corey, I’m asking you how do you feel when you hear the word? How do you think he [a black person] feels when he hears the word?

*Student.* Doesn’t—they call themselves that sometimes.

Mrs. Turner encourages the students to express themselves, but she does not let matters lie. Instead, she confronts the students:

*Mrs. Turner.* Oh, boy, OK, here we go, here we go. This is the thing. I had to teach myself to be able to say that word in front of a class. . . . And because that word is part of society in . . . Maycomb in 1933, nobody really saw much wrong with using the word “nigger” to apply to a black person [then]. It’s in the book, guys, and you’re going to have to deal with it. . . . So Corey, you’re saying that I wouldn’t say that word if a black person were in
Ability Grouping

here, and that’s not true. Because I’m saying it ‘cause it’s in the book. But I want to know, I mean is that, is it a wrong word to you? Would you call somebody that?

*Students.* No.

*Mrs. Turner.* No, you shouldn’t. Of course not.

Although Mrs. Turner drives home her point emphatically, she is still communicating with the students, showing deep concern for their views. Even when they do not take her seriously, she is serious about them:

*[Following further statements by students that border on racist generalizations]*

*Mrs. Turner.* Be quiet! Don’t turn this in to a joke. . . . All right, I, I just can’t believe some of the things I just heard, I really can’t, I don’t, I don’t think it’s funny. We’ve got a problem. Um, if you’re going to take this book seriously, then you’ve got to, you’ve got to think about some of the things that you’re saying. All right. It’s time to relearn some things for some of you because I, I don’t, I hope you don’t grow up like this. All right, I hope some of you are joking, I really do. I’m getting frightened, that, that you, some of you may have these attitudes. Um, go ahead Helen.

*Student.* Well we’re talking about how white people, um, are racist to black people, but blacks do that to us too.

*Mrs. Turner.* Um-hm, sure.

*Student.* I mean just last night, we were on the bus, and, we weren’t even doing anything, all I was like looking out the window, and, and these, these five black guys sitting right next to me . . . and then they start saying like they were going to start a fight. And we didn’t even do anything. . . .

*Mrs. Turner.* Helen, I go along with you . . . the thing is, though, think about, growing up, any of you . . . different from other people. . . . And, if you got people all of your life telling you that you’re, you’re nasty, and you’re a drug dealer, and you can never amount to anything. You can’t go to college, you can’t do this and you can’t do that. How do think you’re going to start to act? *[Students murmur]* You’re going to act just like how people tell you to act, OK, that’s called the self-fulfilling prophesy. And so, you’re right, Helen, I know things like that happen, but some people just sort of live up to what they’re, they’re asked to be. All right. And that comes from us, that doesn’t come, that doesn’t come from anybody’s skin color. That comes from hatred. That’s what hatred causes, and it’s, it’s a disease in our society.

Here again, Mrs. Turner gives the student’s opinion a hearing and then expresses her own view. This is not to say the teacher-student
relationship is symmetrical; on the contrary, the teacher draws on her authority and has the last word. But that is a far cry from the low-track teachers observed by Page (1987, 1989), who attempted to forestall debate and tended to ignore students when they expressed opposing views.

From this point, Mrs. Turner elicited a discussion of peer pressure in closed communities such as Immaculate High School, which she finally brought back to the text:

Mrs. Turner. In this book, . . . Scout Finch and Jem Finch and their friend Dill are going to be three kids, who don't understand all the hatred and racism in their community. All right, and they're going to try, in their own small way, they're going to try and fight it, and they're going to get in trouble for it. And people are going to beat them up, and people are going to cast them out, and people are going to threaten their lives. But they're still going to stick to their guns, about what they believe.

To an observer, it was clear that this lesson required a great deal of energy on the teacher's part. Mrs. Turner was shaken at one point, as she told the students: "I'm getting frightened . . . that some of you hold these attitudes." At another point, she expressed to the students both her frustration and her determination: "I know it's not easy, you guys, I know it's not easy, but we're not going to read a Weekly Reader in this class. All right. You deserve to have this information. So stick with it." Like Mrs. Grant, Mrs. Turner found that teaching the low-track class was demanding work.

It is important to recognize that teaching the low-track class effectively did not mean, for either teacher, teaching it in the same way as other classes. Mrs. Turner's class again provides an instructive example. On the same day as the lesson described above, Mrs. Turner also engaged in a discussion of racism with the honors class, relating it to the book they were reading, *The Adventures of Huckleberry Finn*. Compared with the lesson in the honors class, the remedial class was much more structured. Whereas the honors class began with a joke to warm up the students for a discussion, the remedial class began with a call to attention. Mrs. Turner soon reminded students to take notes and occasionally gave detailed instructions for where students should be in their note-taking outlines (e.g., "Part II, letter A"). In the remedial class Mrs. Turner wrestled with students for their attention ("Alan, ya ready?" "Stop it!" "Pay attention"), but this did not occur at all in the honors class. In contrast with the difficulty Mrs. Turner faced in holding a discussion on racism in the remedial class, discussion in the honors class seemed almost effortless:

November 1993   15
Mrs. Turner. [Can you] recall things from Huck Finn that, um, seemed racist to you?

Student. . . . Miss Watson's, that guy she's always calling "Miss Watson's nigger."


[Mrs. Turner called on other students and interspersed their responses with her comments.]

As in the remedial class, Mrs. Turner treated students' views seriously but gave her own perspective:

Student. Isn't [Twain] being historically accurate when he says "those niggers"?

Mrs. Turner. Oh, yes, absolutely.

Student. So why is it racist?

Mrs. Turner. Well this, that's kind of what I was trying to bring out on the first day, is that Twain is really just trying to mirror the society, and especially the society of . . . Missouri . . . at the time . . . but Twain is using the word rather sarcastically. I mean, you're right, he's being historically accurate, but he's also trying to make a point, um, about the different people who are saying things like [that].

In contrast to the remedial class, where Mrs. Turner had to drive home her points most emphatically, in the honors class the major point was first made by a student. Mrs. Turner asked, "How does that [racism in the book] make you guys feel?" and a discussion ensued:

Student. Everyone claims it's so historical, you can find that anywhere . . . "nigger," you know, you just hear that . . . and people always think . . . it's so historical.

Mrs. Turner. Like, oh, we wouldn't do that anymore.

Student. Yeah, like oh, we're not primitive, you know, and it's not, I mean everybody does that, all the time. Well, not everybody, but people, people do that. . . . People can't get in(to) apartment buildings because they're black.

Mrs. Turner. Um-hm.

Student. They can't go to certain stores because they're black, or they're arrested because they're black . . . you know, it's just, I mean, everybody is always saying how historical it is, and it's right here, and it's right now.

This view, which developed out of a conversation between Mrs. Turner and a student, served as a foundation for Mrs. Turner's concluding comment on the issue:
Gamoran

Mrs. Turner. I like that comment, because do you remember . . . when I gave you that whole list of things that Twain is making fun of in the story? Well all of those things still exist, all right? Gullibility, religious convention, um, all kinds of things. Did this book stop being an accurate mirror of society? At any point?

[Students shake their heads.]

Mrs. Turner. I don’t think so. Um, our society is more technologically advanced . . . but this is still true, I mean these are still problems in our society, they’re the dirt in our society. And we can kind of look into this book as if we were looking into a mirror and saying, oh, OK, these are the things that are wrong with me.

In the low-track class, we observed more structure, more emphasis on order, and more effort on the teacher’s part. At the same time, both lessons exhibited a serious academic purpose and high expectations. In neither class was there a “treaty” or “bargain” allowing students to slide by if they behaved themselves (see Powell et al. [1985] for a contrast). The low-track class session was not the same as the honors class, but it was not a caricature either.

Conclusions

This article has three main points: (1) It is very likely that there are differences among schools in the implementation and effects of tracking. (2) There appear to be instances of successful low-track classes, at least in Catholic secondary schools. (3) The characteristics of these classes include high expectations, an academic curriculum, oral interaction between teachers and students, great effort on the part of teachers, and the absence of a system of assigning weak or less experienced teachers to the lower track.

Despite these conclusions, the article is severely limited by the narrowness of the evidence for success with low-track students. The two new examples came from Catholic schools whose students mainly have economically advantaged backgrounds. Thus, the article adds to the small amount of literature on the effective use of tracking in Catholic secondary schools but provides little basis for knowing whether its findings may generalize to other settings.

Why does tracking appear more effective in Catholic secondary schools? Our study did not address this issue, but it is consistent with Valli’s (1990) and Camarena’s (1990) conclusion that a Catholic school
Ethos of caring and effort is the source of success. According to these authors, the emphasis on caring leads to close monitoring of students, providing them with feedback and taking corrective steps (e.g., requiring summer school) if students falter. Emphasis on effort leads to academic rigor in all types of classes and to a seriousness of purpose among students and teachers in low as well as high tracks. If this explanation is correct, it may reflect in part the ability of Catholic schools to select and reject students. Unlike most public schools, Catholic schools do not have to accept all applicants, and this may allow them to avoid precisely those students who present the greatest challenge in low-track classes: those who are extremely disruptive, and those with exceptionally low achievement levels. Unfortunately, the present study does not shed light on what other school-level conditions might encourage a positive academic climate in low tracks.

More because of their locations than as a result of selection policies, both St. Elias and Immaculate contained mainly economically advantaged, college-bound students, even in the low-track classes. Unlike the schools I described, though, Valli's (1990) school had a diverse student body, and many were not college bound, yet she still found "a curriculum of effort" in the lower track. Moreover, lower-track students at Immaculate were recalcitrant in class just like low-track students elsewhere, as evidenced by observed off-task behavior and flippant responses to teacher questions. As we saw, Mrs. Turner's effectiveness came in the face of such resistance. Other schools in our study, public and Catholic, which had student bodies similar to St. Elias and Immaculate, did not show similar success in low tracks. Thus, student-body composition is not the whole explanation for the presence of high-quality instruction in low-track classes in some Catholic schools.

Yet another limitation of the study, also a form of narrowness, is that it relied on higher-than-expected achievement as a sign of effectiveness without considering other sorts of outcomes. Critics of ability grouping, however, maintain that low-track assignment is stigmatizing, producing harmful social outcomes apart from effects on achievement (see, e.g., Schwartz 1981; Oakes 1985). Cases studied by Valli (1986) and by Camarena (1990) seemed to open the possibility of counteracting this problem. However, in this study, Mrs. Turner commented in the year-end interview that assigning students to a remedial class stigmatizes them and depresses their motivation, and she views this as reason to avoid assigning them to a separate class. In fact, both Mrs. Turner and Mrs. Grant told us that, although they see the ability-grouping question as complex and multisided, on balance they both prefer mixing low-track students with other students. Thus, our examples of teachers who succeeded with low tracks—at least with respect
to instruction and achievement—would actually prefer to end that arrangement. Perhaps, then, these are simply examples of good teachers, who would be effective regardless of how students were assigned. In any case, given the likelihood that ability grouping will continue to be used, we need to know much more about how to use it well.

Notes

This article draws on a larger project carried out collaboratively with Martin Nystrand. Research assistance for this article was provided by Mark Berends and Paul LePore. The author is grateful for helpful comments from Reba Page, Albert Shanker, two referees, and teachers who participated in the study. An earlier version of this article was presented at the 1991 meetings of the American Educational Research Association. The study was conducted at the Center on Organization and Restructuring of Schools, Wisconsin Center for Education Research, University of Wisconsin—Madison, which is supported by a grant from the Office of Educational Research and Improvement (OERI-R117-Q00005). Additional support came from the Institute for Research on Poverty at the University of Wisconsin—Madison and from the Graduate School of the University of Wisconsin—Madison. Any opinions, findings, or conclusions expressed in this article are those of the authors and do not necessarily reflect the views of these agencies or the U.S. Department of Education.

1. The goal of the larger study was to measure variation in instructional quality among different types of classes. While drawing on such quantitative data, the present article mainly provides a qualitative account of the uses of ability grouping in different settings. Rather than testing hypotheses about grouping and instruction, this article illustrates alternatives in the ways grouping and instruction may be linked.

2. I distinguished between the two types of literature using Oakes's description of young-adult fiction as "short novels with themes designed to appeal to teenagers (love, growing pains, gang activity) and written at a low level of difficulty" (1985, p. 76). Thus, I counted novels such as Great Expectations and drama such as Romeo and Juliet as standard literature and works such as S. E. Hinton's The Outsiders as young-adult fiction. To make sure I did not underestimate the use of standard literature in low tracks, I counted a few classic works directed at youthful readers—for example, Johnny Tremain—as standard literature, but the pattern of differences would be the same if this decision were reversed.

3. Open-ended questions, or "authentic" questions, were defined as questions for which the teacher had no prespecified answer. These typically included opinion questions and questions about facts that the teacher did not know. Further analyses indicated that, although high- and low-track teachers asked similar proportions of authentic questions, the questions concerned different topics, with those in high-track classes far more often related to literature students were reading (Gamoran and Nystrand 1992).

4. The two studies, both unpublished doctoral dissertations, were conducted by Chiotti (1961) and Platz (1965). Nine other studies published since 1960 showed zero or negative effects for low-group students. Among secondary-school studies prior to 1960, only four provided separate data on achieve-
Ability Grouping

ment of high- and low-group students, of which three showed positive effects for low-group students (Billett 1928; Martin 1927; Purdom 1929).

5. The year-end test required students to answer a series of questions about readings they had read for class. Questions addressed basic recall as well as in-depth understanding. Tests of prior reading and writing were administered in the fall, as were questionnaires that provided information on students’ sex, race and ethnicity, and family socioeconomic status (see Gamoran et al. [1992] for further details). Residuals were computed at the student level and then averaged to the class level. Each class was considered separately, regardless of the number of low-ability classes in the school, but the two schools I selected each had only one low-ability class. This method of choosing cases would not be appropriate for a test of the hypothesis that instructional differences produce track differences in achievement or that high-quality instruction in low-track classes produces high achievement. That would require quantitative assessment of the impact of instruction on achievement in the different types of classes (see Gamoran et al. 1992). For this article I did not automatically select all classes with positive average residuals, although as it happens these two classes had the highest positive residuals and only one other low-ability class had an average residual greater than one. Rather, I am focusing on the low-ability classes that most clearly combined high-quality instruction with higher-than-expected achievement. I adopted this approach because I am illustrating a possibility rather than testing hypotheses.

6. All names are pseudonyms.

7. Other long works of fiction used in ninth-grade English classes at Immaculate included The Miracle Worker in the remedial class, All Quiet on the Western Front and Animal Farm in the regular classes, and The Odyssey, The Red Badge of Courage, The Adventures of Huckleberry Finn, and Animal Farm in the honors class.

8. The two classes were also similar in that low-track students reported a lot of homework (averages of 1.5 and 1.8 hours per week in St. Elias and Immaculate, respectively) and high proportions of written work completed (averages of 96.2 percent and 88.6 percent, respectively, in the low-track classes). There were also some differences. Low-track students in St. Elias wrote more than a page more than twice per month while filling in blanks less than once per month, and we observed no students off task, but only 4.3 percent of teacher questions were open-ended, and students spent 13.8 minutes per day in seat work. In Immaculate, 25.9 percent of teacher questions were open-ended in the low-track class, and students did very little seat work (2.7 minutes per day), but off-task rates averaged 7.2 percent of students, and filling in blanks was more common (about three times per month) than writing more than a page (less than once per month).

9. Our observations confirmed that Mrs. Grant used less small-group time with the lower-track class. However, questionnaire data failed to show a significant difference between the two classes in students’ perceptions of how often readings were related to their own experiences.

References


Ability Grouping


