Synthesis of Research on Grouping in Elementary and Secondary Schools

Certain arrangements appear to be instructionally effective for students, but the psychological drawbacks may offset any advantages.

One of the most difficult decisions school administrators have to make, yet one of the most important, is how to group students for instruction. Should students be assigned to classes heterogeneously or according to ability? What kinds of grouping strategies, if any, should be used within classes? Should schools be departmentalized, self-contained, or some mix of the two? These questions are at the heart of school and classroom organization, and debates about them have gone on since the 19th century.

Decisions about grouping may be made on many grounds, from concerns about students' self-concepts and behavior to concerns about equity to concerns about efficiency or teacher morale. However, any discussion of grouping must consider the effects of alternative grouping plans on student achievement. My purpose here is to summarize what is known about the achievement effects of various forms of grouping at the elementary and secondary levels. I draw primarily on a "best-evidence synthesis" of research on ability grouping in elementary schools (Slavin 1987a,b), plus additional evidence concerning ability grouping at the junior and senior high school levels and evidence relating to alternatives to ability grouping at the elementary and secondary levels.
Ability grouping plans may stigmatize classes. Methods of dealing with students provided by heterogeneous students of the examples and stimulation to meet the needs of a more diverse student body (see Billett 1932). This too has a distinctly modern ring.

The principal rationale for various forms of grouping has been basically the same since grouping was first proposed. Ability grouping exists to deal with one central fact of mass education: students differ in knowledge, skills, developmental stage, and learning rate. If a teacher is to present a lesson to a class, then it seems intuitively obvious that the lesson should be neither too easy nor too difficult for the students. If the class is highly heterogeneous, then one lesson will of necessity be easier than would be optimal for some students and more difficult than would be optimal for others. For the sake of instructional efficiency, students should be grouped so that they will profit from one lesson.

Yet virtually every means of grouping students by ability or performance level has drawbacks that may be serious enough to offset any advantages. Ability grouping plans may stigmatize low achievers, put them into classes or groups for which teachers have low expectations, or lead to the creation of academic elites. Grouping may doom children who are not in the top tracks to second-class instruction and, ultimately, second-class futures. It may deprive students of the examples and stimulation provided by heterogeneous classes. Methods of dealing with student heterogeneity within the classroom, such as use of reading or math groups, create problems in terms of the management of multiple groups and reduction in the direct instruction received by individual students.

### Instructionally effective cooperative learning methods provide group rewards based on the individual learning of all group members.

#### Background and Rationale for Grouping

The study of alternative grouping arrangements has a long and distinguished pedigree. As early as 1929, Luther Purdom referred to the "great mass of literature" on the topic dating back to 1917. He complained that grouping decisions were too often based on personal impressions rather than hard evidence—a complaint that remains applicable 60 years later. The considerable interest in research on grouping in the 1920s was stimulated in part by a concern that since schools were serving a larger number of immigrants and children of the poor, they needed to differentiate instruction to meet the needs of a more diverse student body (see Billett 1932). This too has a distinctly modern ring.

The principal rationale for various forms of grouping has been basically the same since grouping was first proposed. Ability grouping exists to deal with one central fact of mass education: students differ in knowledge, skills, developmental stage, and learning rate. If a teacher is to present a lesson to a class, then it seems intuitively obvious that the lesson should be neither too easy nor too difficult for the students. If the class is highly heterogeneous, then one lesson will of necessity be easier than would be optimal for some students and more difficult than would be optimal for others. For the sake of instructional efficiency, students should be grouped so that they will profit from one lesson.

Yet virtually every means of grouping students by ability or performance level has drawbacks that may be serious enough to offset any advantages. Ability grouping plans may stigmatize low achievers, put them into classes or groups for which teachers have low expectations, or lead to the creation of academic elites. Grouping may doom children who are not in the top tracks to second-class instruction and, ultimately, second-class futures. It may deprive students of the examples and stimulation provided by heterogeneous classes. Methods of dealing with student heterogeneity within the classroom, such as use of reading or math groups, create problems in terms of the management of multiple groups and reduction in the direct instruction received by individual students.

#### Types of Grouping

The principal types of grouping arrangements fall into two major categories: between-class and within-class. Between-class plans are school-level arrangements by which students are assigned to classes. Several means of assigning students to classes by ability fall into this category, as does departmentalization. Within-class grouping arrangements may attempt to reduce the heterogeneity of instructional groups, as in the use of within-class ability grouping or mastery learning. Finally, cooperative learning is a within-class grouping strategy that uses heterogeneous rather than homogeneous subgroups.

The following sections define the various grouping plans and summarize the research on each.

#### Between-Class Ability Grouping

Perhaps the most controversial form of grouping is assignment of students to groups according to ability or performance. Arguments about the desirability of between-class ability grouping have raged from the 1920s (e.g., Miller and Otto 1930) to the present (e.g., Good and Marshall 1984; C.-L. Kulik and J. A. Kulik 1982, 1984). Proponents have argued that ability grouping lets high achievers move rapidly and gives low achievers attainable goals and extra help. Opponents have countered that ability grouping is unfair to low achievers, citing the problems of poor peer models, low teacher expectations, and slow instructional pace.

Yet ability grouping is not a single practice but has many fundamentally different forms that have different educational as well as psychological effects. The most important forms of between-class ability grouping are discussed in the following sections.

- **Ability-grouped class assignment.** In many elementary schools, students are assigned to self-contained classes on the basis of a general achievement or ability measure. This method might produce, for example, a high-achieving 4th grade class, an average-achieving class, and a low-achieving class, with students assigned to classes according to some combination of a composite achievement measure, IQ scores, and/or teacher judgment. Students remain with the same ability-grouped classes for all academic subjects. At the secondary level, ability-grouped class assignment usually means that students are assigned to a particular track within which they receive all, or almost all, instruction (e.g., advanced, general, and remedial or academic, general, and vocational).

The achievement effects of ability-grouped class assignment (compared to heterogeneous grouping) are essentially zero. I identified 14 methodologically adequate studies of this practice at the elementary level and found the median effect size on standardized achievement measures to be approximately .00 (Slavin 1987a). Research at the junior high school level also fails to find consistent positive effects of between-class ability grouping (Borg 1965, Stoakes 1964, Peterson 1966). High school research is relatively rare, but even at this level controlled studies find few benefits of tracking (Borg and Perpich 1966, Billett 1932, Purdom 1929). There is some evidence that high achievers...
may gain from ability grouping at the expense of low achievers (e.g., Borg 1965, Flair 1965, Tobin 1966), but most studies find no such trend (see, e.g., Morgenstern 1963, Peterson 1966, Bremer 1958, Hartill 1936). Overall, the effects of ability grouping cluster closely around zero for students of all achievement levels.

One probable reason that ability-grouped class assignment has little effect on student achievement is that this plan typically has only a limited impact on the heterogeneity of the class. For example, Goodlad (1960) estimated that dividing a group of elementary students into two ability groups on the basis of IQ reduced total variability in each class by only 7 percent. With three groups, heterogeneity was reduced by 17 percent, still not likely to be enough to have a measurable impact. Even though a student's performance in any one subject is correlated with performance in other subjects, this correlation is far from perfect. This means that grouping students on any one criterion is sure to leave substantial heterogeneity in any specific skill domain. On the other hand, assigning students to "high" and "low" ability classes may have a stigmatizing effect on low achievers and may evoke low expectations for student achievement and behavior even if the grouping has a minimal impact on class heterogeneity. Thus, ability-grouped class assignment may be enough to produce psychological drawbacks but does not do enough to reap the potential educational benefits of reducing student heterogeneity in any particular skill. Nonetheless, it is interesting that ability grouping is ineffective even at the secondary level, where student heterogeneity may be extreme.

Regrouping for reading and/or mathematics Another commonly used ability grouping arrangement has students remain in heterogeneous classes most of the day but regroup for selected subjects. For example, three 4th grade classes in an elementary school might have reading scheduled at the same time. At reading time, students might leave their heterogeneous homes rooms and go to a class organized according to reading levels. At the secondary level, students might be grouped for some subjects, but not for others.

Regrouping for selected subjects has three important advantages over ability-grouped class assignment. First, students remain in a heterogeneous setting most of the day, so they are likely to identify with that group, thereby reducing the labeling effect of all-day grouping. Second, students are grouped solely on the basis of their achievement in the particular subject, not by general achievement or ability level, so a meaningful reduction in heterogeneity in the skill being taught is possible. Third, elementary regrouping plans tend to be more flexible in operation than ability grouped class assignment, because changing students between reading or mathematics classes is less disruptive than changing basic class assignments. For this reason, any errors in assignment can be easily remedied, and any changes in student performance level can be accommodated with a change in grouping.

The limited research on regrouping plans suggests that regrouping can be instructionally effective if: (1) instructional level and pace are completely adapted to student performance level, and (2) the regrouping is done for only one or two subjects so that students stay in heterogeneous placements most of the day (Slavin 1987b). Studies at the elementary level that met these conditions have generally found positive effects on student achievement in reading (Berkun et al. 1966), in mathematics (Provis 1960), and in reading and mathematics taken together (Balow and Ruddell 1963, Morris 1969). On the other hand, when regrouping has been done in elementary schools without adapting the pace or level of instruction (Moses 1966, Davis and Tracy 1963) or in more than two different subjects (Koontz 1961), no benefits have been found. At the junior high school level, grouping in one subject has been no more successful than ability-grouped class assignment (Bick 1962, Fick 1963), and similar results have been
obtained in high school studies

**Joplin Plan.** One interesting form of regrouping applicable primarily to elementary schools is the Joplin Plan (Floyd 1954), in which students are regrouped for reading without regard for grade levels. That is, a reading class at 4th grade–first semester level (4-1) might contain some 3rd, some 4th, and some 5th graders. One important consequence of this grouping plan is that it allows for the reduction or elimination of within-class grouping for reading, as students in each reading class may all be at the same reading level. Thus, teachers can spend more of the reading class time doing direct instruction, reducing the time during which students must do unsupervised follow-up seatwork.

Effects of the Joplin Plan and closely related forms of nongraded plans have been quite positive overall. In a recent review, I (Slavin 1987a) estimated the median effect size for achievement on reading to be +.44 for reading achievement, and one study (Hart 1962) found similar effects for mathematics. Two of the studies I reviewed (Morgan and Stucker 1960, Hillson et al. 1964) used random assignment of students and teacher to treatments, and most of the others were good-quality matched equivalent studies.

**Nongraded plans.** Nongraded plans (Goodlad and Anderson 1963) are elementary grouping arrangements in which formal grade levels are abolished in favor of flexible cross-age groupings for different subjects. Many different forms of nongrading have been evaluated. In some cases (e.g., Hillson et al. 1964, Ingram 1960) where nongrading is done in reading or mathematics only, it is essentially identical to the Joplin Plan. At the other extreme, nongraded plans may involve many subjects, constant and flexible regrouping within and between classes, extensive use of individualized instruction, team teaching, and other features more in line with forms of the open classroom, than with the still quite traditional Joplin Plan (see, e.g., Bowman 1971, Ross 1967).

I found achievement effects of comprehensive nongraded plans to be inconsistent but generally positive (Slavin 1987a, b). However, most of the studies that failed to find positive effects of nongraded plans either took place in laboratory schools (e.g., Otto 1963, Ross 1967) or found few implementation differences between nongraded and traditional programs (e.g., Carbone 1961, Hopkins et al. 1965). Studies in regular classrooms in which the nongraded program was used conscientiously usually found positive effects on student achievement (e.g., Bowman 1971, Machiele 1965).

**Gifted programs.** Gifted programs may be offered in only one subject (often mathematics), or separate classes for high achieving or otherwise exceptional students may be provided. Most research on special programs for the gifted contains a serious systematic bias. Much of this research (e.g., Simpson and Martson 1961) compares students in gifted programs to students in the same schools who did not participate in the gifted program, matching on IQ or other measures. However, if
two students have IQs of 130 and one is accepted for a gifted program while the other is rejected, then we can be sure that there were other factors, such as motivation, achievement, and so on, that also differentiate these students, all of which are likely to work to the advantage of the students accepted for the special program. Nonrandomed comparisons across schools with and without gifted programs (e.g., West and Sievers 1960, Baldauf 1959) are somewhat better, but also suffer from this systematic selection bias. Studies of acceleration (see J. A. Kulik and C. L. Kulik 1984), exposing gifted students to higher levels of information in some subjects or skipping them to higher grades, share the same problems of selection bias, with the additional problem that the accelerated students are exposed to material not seen by regular students, making comparisons difficult.

The literature on gifted programs at the elementary level is small, inconclusive, and methodologically inadequate for the reasons described above and others. A few studies have reported achievement benefits of special programs for the gifted (e.g., West and Sievers 1960, Bell 1957, Atkinson and O’Connor 1956), while others found no significant advantages (e.g., Baldauf 1959, Cluff 1964, Becker 1963). Generally, acceleration programs (e.g., providing Algebra I to 6th or 7th graders) produce better achievement outcomes than enrichment programs (see Fox 1979, Justman 1954, Pasow et al. 1961).

Special education. Assignment to full- or part-time special education programs for learning problems can also be seen as a form of between-class ability grouping. Unfortunately, most research on the efficacy of special education has the same type of methodological flaws characteristic of studies of gifted programs. Matched comparisons between students assigned to special education and those remaining in regular classes are biased toward regular class assignment because of the likelihood of systematic differences between students of, for example, the same IQ assigned to regular or special classes (Madden and Slavin 1983).

However, there are a few randomized studies at the elementary level of special education vs. mainstreaming (regular class placement) for students with moderate academic handicaps, such as learning disabilities and mild retardation. For example, Goldstein and colleagues (1966) found that students randomly assigned to a high-quality special education program learned no more than did students who remained in regular classes. Calhoun and Elliott (1977) found that mildly retarded and emotionally disturbed students randomly assigned to regular classes that used individualized instruction gained substantially more in achievement than did students who experienced the same individualized program with the same teacher in self-contained special education classes. Overall, evidence from studies of special education and mainstreaming supports accommodating student differences within the regular classrooms rather than separating students into special classes (Madden and Slavin 1983).

Summary and conclusions: Between-class ability grouping. Evidence from studies of various forms of between-class ability grouping in elementary schools indicates that achievement effects depend on the types of programs evaluated. In general, ability grouping plans are beneficial for student achievement when they incorporate the following features (adapted from Slavin 1987b):

1. Students remain in heterogeneous classes most of the day and are regrouped by performance level only in such subjects as reading and mathematics in which reducing heterogeneity is particularly important.
2. The grouping plan reduces heterogeneity in the specific skill being taught.
3. Group assignments are flexible and are frequently reassessed.
4. Teachers adapt their level and pace of instruction in regrouped classes to accommodate students’ levels of readiness and learning rates.

The between-class grouping plan that most completely incorporates these four principles is the Joplin Plan. Evidence on the Joplin Plan strongly supports the effectiveness of this arrangement and of within-grade regrouping plans and nongraded plans that most resemble it.

In contrast, ability-grouped class assignment, special programs for the gifted, and special education for students with learning problems do not generally meet the four criteria. Typically, they segregate students all or most of the day, are based on general ability or achievement rather than skill in a specific subject, and tend to be highly inflexible. Teachers may or may not adjust their level and pace of instruction to adapt to students’ needs in these plans. Evidence at all levels shows no benefits for ability-grouped class assignment or special education assignment and only inconsistent and flawed evidence in favor of special programs for the gifted.

Departmentalization
One between-class grouping plan that is nearly universal in secondary schools and increasingly seen in elementary schools is some form of departmentalization, in which teachers teach one or a few (but not all) subjects to multiple class groups. At the
secondary level, departmentalization usually implies that students change classes for all subjects. Alternatively, in elementary or middle schools, one teacher might see classes for more than one, but not all subjects, for example, teaching combined reading and language or mathematics and science blocks. Semi-departmentalized plans may have all teachers teach a set of subjects to students in their home-room groups in the morning, while other subjects are departmentalized in the afternoon.

The main advantage of departmentalization is that it allows teachers to specialize, teaching subjects they are most able and willing to teach. However, departmentalization risks diffusing responsibility for individual children, making it difficult for a student to identify with a single caring adult. In elementary schools, it may represent an unhealthy intrusion of the "subject-centered" secondary organization into the more typically "child-centered" elementary program.

Unfortunately, there is very little research on departmentalization. However, that which does exist fails to support the practice. Ward (1970) found that students in grades 4-6 learned reading and science significantly better in self-contained classes than in departmentalized classes, and there were no differences in mathematics and social studies. Hosley (1954) also found that 6th graders' achievement was higher in self-contained K-6 schools than in semi-departmentalized junior high schools. Jackson (1953) found no achievement differences between departmentalized and nondepartmentalized plans at the elementary or junior high school levels, and Spivak (1956) found higher achievement for 7th and 8th graders in self-contained than in departmental

Ability grouping is ineffective even at the secondary level, where student heterogeneity may be extreme.
Methodologically adequate research on within-class ability grouping has unfortunately been limited to the study of mathematics grouping, perhaps because few teachers would be willing to participate in an experiment in which they had to teach heterogeneous classes without breaking students into reading groups. However, the research on within-class grouping in mathematics clearly supports this practice. Every one of the eight studies of within-class ability grouping in mathematics that I identified (Slavin 1987), including five randomized studies, favored the grouped treatment (median effect size = +.34). Effects of within-class grouping were somewhat higher for low achievers (median effect size = +.65) than for average and high achievers. There was some trend for effects to be more positive when the number of ability groups was two or three rather than four.

Effects of within-class grouping on mathematics achievement cannot be assumed to hold for reading. In mathematics, students need to work problems independently, so there is an appropriate place for independent seatwork. The corresponding need for independent seatwork in reading is less compelling. The universality of within-class grouping in reading in North America provides at least some indication that this form of ability grouping is also instructionally necessary, although it is interesting to note that European schools do not generally use this type of grouping in reading.

Mastery learning. Mastery learning is a form of within-class grouping in which grouping decisions are constantly changed to reflect student performance. In group-based mastery learning (Bloom 1976, Block and Burns 1976), students receive a lesson as a whole class and then take a formative test. Those whose test scores exceed a present mastery criterion (e.g., 80 percent) then do enrichment activities; those who do not achieve this criterion receive corrective instruction. Group-based forms of mastery learning are by far the most commonly used in elementary and secondary schools.

I recently reviewed the research on the achievement effects of group-based mastery learning (Slavin 1987c). In methodologically adequate studies of at least four weeks’ duration, mastery learning had no significant effect on standardized achievement measures and only moderate effects on experimental-made measures. However, there are examples of successful continuous-progress forms of mastery learning, in which students proceed through a hierarchy of skills at their own rate and are frequently assessed and regrouped (see Slavin 1987d).

Cooperative learning. Cooperative learning (Sharan 1980, Slavin 1983, in press) refers to various instructional methods in which students work in small heterogeneous learning groups toward some sort of group goal. Cooperative learning differs from within-class ability grouping not only in that cooperative learning groups are small and heterogeneous, but also in that these groups are expected to engage in a great deal of task-focused interaction, such as studying together or completing group assignments. Cooper-
tive learning methods vary considerably in their basic structures. Some, such as Jigsaw Teaching (Aronson et al. 1978) and Group Investigation (Sharan and Sharan 1976) assign students specific tasks with a larger group task. In others (e.g., Johnson and Johnson 1975) students work together to complete a common group worksheet or other group product. A third category consists of methods in which students study together and are rewarded on the basis of the achievement of all group members (e.g., DeVries and Slavin 1978, Slavin 1983). For example, in Student Teams-Achievement Divisions, or STAD, students are assigned to four-member heterogeneous teams. The teacher presents a lesson, and then students study worksheets together in their teams, attempting to ensure that all team members master the material. Finally, the students are individually quizzed, and teams are rewarded with certificates or other recognition based on the average of their members' quiz scores.

The idea behind cooperative learning is that if students are rewarded on the basis of group or team performance, they will be motivated to help and encourage one another to achieve (Slavin 1983, in press). In a heterogeneous learning group, students among themselves are expected to be able to solve problems or organize material presented by the teacher and to transmit the group's understanding to each individual.

Research on cooperative learning in elementary schools has found that its effectiveness depends on how it is organized. Instructionally effective cooperative learning methods provide group rewards based on the individual learning of all group members. Stu-

---

**SCHOOL IMPROVEMENT**

**with LARRY LEZOTTE**

Larry Lezotte is America's foremost authority on effective schools and school improvement. His message to Educational Policy Makers and Administrators goes far beyond the identification of effective school correlates. He provides vivid and inspiring descriptions of the purposes of schooling, various approaches to school improvement and the need for equity in instructional programming. This 2 day institute is designed for: District Wide Supervisors • Building and District Administrators • Board Members and Policy Makers

**DATES:** December - 7, 8, 9, 1988 (Begins 8:00 p.m. December 7th)

**LOCATION:** The Conference is located in beautiful San Antonio, Texas. Information about the training site and overnight lodging will be sent to you upon receipt of your registration.

**REGISTRATION FEE:** $325.00 includes training, materials, and books.

**OBJECTIVES:** • Identify the essential components of effective schools. • Identify the steps to School Improvement. • Analyze various effective school improvement programs.

**MAIL REGISTRATIONS TO:** Teacher Effectiveness Associates
4993 April Day Garth
Columbia, Maryland 21044

**FOR INFORMATION CALL:** 301-621-5151

---

**REGISTRATION FORM**

Name ____________________________
Name You Prefer To Be Called __________
Organization/System ____________________
Address ____________________________ Zip __________
Phone ____________________________ Registration fee of $325 or purchase order must accompany registration form.
A Matter of Balance

While there is much we can learn from the research on grouping, there is still much to understand. For example, basic questions about the effects of within-class ability grouping in reading, of mastery learning, and of various forms of departmentalization remain unanswered. The mechanisms by which different grouping plans may have their effects are not at all clear.

Yet we do know enough to dispel the notion that simply grouping students by ability will in itself accelerate their achievement. Certain forms of grouping may be effective as part of a coherent plan for adapting instruction to meet individual needs, but the costs and benefits of each form of grouping must be carefully weighed.

References


Cliff, J. E. "The Effect of Experimentation and Class Reorganization on the Scholastic Achievement of Selected Gifted Sixth Grade Pupils in Wichita, Kansas." Dissertation Abstracts 24 (1964) (University Microfilms No. 64–10, 959).


Fox, L. H. "Progress for the Gifted and Talented: An Overview." In The Gifted


Author's note: This article is adapted from a review by Slavin (1987a). Preparation of this article was supported by a grant from the Office of Educational Research and Improvement, U.S. Department of Education (No. OERI-G-86-0006). However, the opinions expressed are those of the author and do not represent OERI positions or policy.

Robert E. Slavin is Director, Elementary School Program, Center for Research on Elementary and Middle Schools, The Johns Hopkins University, 3505 N. Charles St., Baltimore, MD 21218.

NATIONAL SYMPOSIUM ON STAFF & CURRICULUM DEVELOPMENT FOR COOPERATIVE LEARNING JANUARY 26-27, 1989 SAN FRANCISCO, CA

Rather than cooperative learning methodologies, this symposium will focus on implementation issues such as:

STAFF DEVELOPMENT
Inservice sequences to achieve whole-school change
Designing effective cooperative learning inservice programs

COACHING
Designing coaching programs
Presentation of existing coaching models

CURRICULUM DEVELOPMENT
Developing/choosing cooperative learning curriculum
Presentation of existing cooperative materials

ADMINISTRATION
Role in creating or supporting whole-school change
Role in establishing a collaborative school

HIGHER EDUCATION
Pre-service cooperative learning education
University-school partnership for whole-school change

PRESENTERS
Robert Slavin, Pat Wilson O'Leary, Spencer Kagan
Laurel Robertson, Carole Cooper

Sponsored by The Development Studies Center

REGISTRATION

| If you wish to present or want additional program information, contact Susan Unquhart-Brown, 415-838-7633. Information regarding hotel accommodations will be mailed with registration confirmation. Purchase order or check payable to Developmental Studies Center, 111 Deerwood Place, Suite 165, San Ramon, CA 94583. Registration Fee: $235 Deadline: Oct. 24, 1988 |

| NAME |
| POSITION |
| DISTRICT/ORGANIZATION |
| BUSINESS ADDRESS |
| PHONE |
| HOME ADDRESS |
| PHONE |

SEPTEMBER 1988