



RESEARCH SYNTHESIS ON ABILITY GROUPING

The opening of the first graded school in this country in 1848 started a revolution in American education (Otto, 1950). With its multiple classrooms, each under the direction of a different teacher and all governed by a single principal, the graded school introduced economies and efficiencies into education that were hard to ignore. By 1860 nearly all the larger cities in America had graded schools, and by 1870 the age-graded school was firmly established as the dominant model for school organization.

By the turn of the century, however, critics were raising questions about the adequacy of this model. What they objected to most was classification of students exclusively by age in the new schools. At each age level, the range in student mental age was likely to be as great as the number of their grade—a two-year range in mental age in the second grade, a three-year range in the third, and so on. How can children who differ so much be taught in the same classroom, critics asked. Revisionists therefore soon began modifying schools again to accommodate the individual differences among children. Ability grouped classrooms were among the earliest and most enduring of these modifications. From the turn of the century when they were first introduced until the present, such classrooms have been a prominent feature on the landscape of American education—sometimes singled out for praise, sometimes criticized, and nearly always a focus for discussion and debate.

Supporters of ability grouping claim that homogeneous groups of students are easier to teach. According to such advocates, teachers have fewer individual differences to contend with in ability-grouped classes, and students learn more when instruction can be aimed at the right level. Opponents of grouping,

"The effects of grouping were clearer on student attitudes than on student achievement."

on the other hand, bitterly criticize the practice. They complain that grouping is popular because it appeals to a basic human need to stratify society into ingroups and outgroups. In this account, grouping does not foster better learning, but confers unnecessary distinction on those in the fast tracks while placing a stigma on those in the slow groups.

In 1916 the first serious attempt was made to study homogeneous grouping with something resembling controlled experimentation. In that year, Guy M. Whipple studied a gifted class consisting of 13 boys and 17 girls, chosen on teacher recommendation from the fifth and sixth grades of a school in Urbana, Illinois. In the years since, researchers have carried out many additional studies of grouping effects. In a typical study, a researcher divides students of a given grade into experimental and control groups. Members of the experimental

group are instructed in classes that are relatively homogeneous in ability; members of the control group are taught in conventional heterogeneous classes. At the end of the experiment, the researcher compares responses of the experimental and control groups on a common examination or on a course evaluation form.

In the years since such studies first appeared, reviewers have often tried to make sense of the accumulating findings. Their central tool in summarizing results has been the "box-score" tally of positive and negative findings, and their central message has been that nothing is established with certainty about the effects of grouping. Unlike these earlier reviewers, we employed an objective and quantitative approach to draw conclusions about effects of grouping. The method we used has been called "meta-analysis," or the analysis of analyses. The term was first used by Glass (1976) to describe the statistical analysis of a large collection of results from individual studies for the purpose of integrating findings. Meta-analysts use objective search procedures to locate as many studies of an issue as possible. They then describe the features and outcomes of the studies in quantitative or quasi-quantitative fashion. Finally, meta-analysts use statistical methods to relate study features to study outcomes.

Chen-Lin C. Kulik is Assistant Research Scientist and James A. Kulik is Research Scientist, Center for Research on Learning and Teaching, The University of Michigan, Ann Arbor.

The material in this article is based on work supported by the National Science Foundation under grant no. SED 79-20742. A more detailed presentation of the findings appears in a forthcoming issue of the *American Educational Research Journal*.



Highlights from Research on Ability Grouping

Except for high-ability students in honors classes, ability grouping has little significant effect on learning outcomes, student attitudes toward subject matter and school, and self-concept. The differences that are found in grouped classes are all positive, however slight, and there is no evidence that homogeneous grouping is harmful.

Learning Outcomes. In general, students who are grouped in classes according to academic ability outperform nongrouped students only slightly. However, students in gifted and talented programs perform better than they would in heterogeneous classes. In contrast, students in classes for the academically deficient perform neither better nor worse than they would in a mixed-ability class. The effects in multi-track, as opposed to mixed, classrooms are also negligible.

Attitudes. Students who are ability-grouped for a particular subject, such as mathematics or English, have a better attitude toward the subject. There is very little difference between grouped and ungrouped students in their attitudes toward school.

Self-Concept. The effects of grouping on self-concept are positive but minor.

ASCD's Research Information Service will help ASCD members locate sources of information on topics related to curriculum, supervision, and instruction. Send your specific question in writing to Research Information Service, ASCD, 225 N. Washington St., Alexandria, VA 22314.

It was impossible for us to cover all aspects of ability grouping in our review, however. Like other reviewers, therefore, we adopted a specific focus. Our review was limited to comparative studies at the secondary school level. Among the major questions that we asked were: What are the effects of ability grouping in the typical study? Does grouping have different effects on students of different ability levels? Does grouping have different effects in different areas of student development—for example, in academic, attitudinal, and personal areas? Finally, do the effects of grouping vary with study settings, with methodological features of the studies, and with types of grouping practices?

Method

The data for the meta-analysis came from 52 objective, comparative studies of grouping that we located through computer searches of the educational literature. The studies differed somewhat in their basic approach to grouping. Some grouping programs covered the full range of student ability ("multi-track" programs); some provided special classes only for gifted or talented

students (honors programs); and some provided special classes only for low-achieving or low-aptitude students (remedial programs). The studies also differed in experimental design, course setting, and publication history. We created categorical variables, similar to those used in our previous meta-analyses (Kulik, Kulik, and Cohen, 1980) to classify the 52 studies according to such features.

The 52 studies also described educational outcomes in three different areas: learning, attitudes, and self-concept. To quantify effects of grouping in each of these areas, we used Glass's (1976) index of effect size (*ES*). This index gives the number of standard deviations that separate the group averages that are being compared. The index is defined as the difference between the means of two groups divided by the standard deviation of the control group.

Learning Outcomes

Fifty-one of the 52 studies described effects of grouping on examination scores of students. In more than 70 percent of these studies, students from the grouped classes outperformed students

from the ungrouped classes by at least a small amount. The average *ES* in the 51 studies, however, was only .10. This means that in a typical study, grouping raised the performance of students by .10 standard deviations, or from the 50th to the 54th percentile. Cohen (1977) has referred to effects of this magnitude as small in size.

Further examination of the data showed that studies with certain features consistently produced strong effects. An especially important feature was the basic type of grouping program. Effects were largest (approximately .40) in the 14 programs designed for talented and gifted students; these students accomplished more in special "honors" programs than they would have in mixed ability classrooms. Effects were near zero in the four programs designed especially for academically deficient students; such students learned as much in mixed-ability classrooms as they would have in homogeneous classrooms. Finally, effects were also near zero in the 33 studies that compared effects of multi-track vs. mixed-ability classrooms on unrestricted populations.

We carried out a further analysis of the 33 studies of unrestricted populations. Each of these studies included students of high, middle, and low ability, and in each study students were assigned to both homogeneous and heterogeneous classes. In 19 of the 33 studies, the investigators reported results in enough detail so that we could calculate separate measures of effect size for students at each ability level, and thus determine whether effects of multitrack programs were strongest on high-, middle-, or low-ability students. The effects turned out to be similar at each ability level. The average effect size was .00 for students in the high-ability classes; $-.06$ for students in the middle-ability classes; and .002 for students in low-ability classes.

Attitudinal Outcomes

Eight studies reported results on student attitudes toward the subject matter taught in grouped and ungrouped classrooms. In each of these eight studies grouping was used only for teaching a specific subject matter—for example, mathematics or English composition—and not for an entire school program. In seven of these studies, student attitudes were more positive in the grouped class. The average *ES* was .37. Even though the number of studies available was small, results were consistent enough

for us to conclude with statistical confidence that grouping had a positive effect on student attitudes toward the subject being taught.

Another 11 studies reported results on attitudes of students toward the school they were attending. In eight of the studies the students from grouped classes expressed more favorable attitudes toward their schools; in the other three studies the attitudes of the students from ungrouped classes were more favorable. The average *ES* in the 11 studies was .09. This effect is a very small one at best, and was not large enough to be considered statistically reliable.

Effects on Self-Concept

Fifteen studies reported results on student self-concept. In seven studies self-concept was higher for students in grouped classrooms; in six studies self-concept was higher for students from ungrouped classrooms; and in two studies self-esteem was equal for the two groups. The average *ES* in these studies was .01, a trivial value.

Discussion

The practice of ability grouping has had a long and a sometimes stormy history. When first introduced as a way of individualizing instruction in large schools at the turn of the century, ability grouping excited the imagination of teachers, administrators, and educational researchers. In the years since then, teachers have continued to support the practice, but the reputation of ability grouping has waxed and waned among educational researchers and administrators. Although reviewers have never lacked an adequate pool of research studies with which to answer the basic questions about ability grouping, they have lacked a way of reading the message that the research studies hold.

Meta-analysis showed that only one type of grouping has clear effects on student achievement. This is the type in which students of high ability are put into a special honors class for enriched instruction in their secondary school subjects. Studies of this type usually report significant results, and they usually report effects on achievement that are medium in size. High ability students apparently benefit from the special curricula that grouping made possible.

Other types of studies reported only trivial effects on student achievement. Most studies that compared achievement in multi-track and single-track

schools, for example, reported no significant effects of grouping. Students learned as much in the heterogeneous classrooms as in the homogeneous ones. And this was as true of the students assigned to the high aptitude classes in a multi-track school as it was of the students who ended up in the low aptitude classes. Studies that dealt only with low-ability students—assigned either to special remedial classes or to heterogeneous ones—also reported no examination differences associated with grouping.

“the effect of grouping was near zero on the achievement of average and below-average students....”

Our conclusions about student achievement were generally consistent with the conclusions drawn by narrative reviewers. Others have also noted the effectiveness of honors classes in reviews of grouping research (Ekstrom, 1961), and they have also noted that ability grouping has only trivial effects on the achievement of average and below-average students (Findlay and Bryan, 1971). Our results support these observations, but they do not support the view that grouping has unfavorable effects on the achievement of low aptitude students. In our analysis, the effect of grouping was near zero on the achievement of average and below-average students; it was not negative.

The effects of grouping were clearer on student attitudes than on student achievement. Students assigned to grouped classes for work in certain subject areas, such as mathematics or English composition, responded more favorably to these subjects than did similar students assigned to heterogeneous classes. Effects were positive in nearly all the studies of attitudes toward subject matter, and in the typical study these attitudinal effects were medium in size. Effects of grouping on attitudes toward school and on self-concept were

also positive, but these effects were smaller and less consistent.

Our conclusions about attitudinal effects introduce a new note into reviews of grouping research. Many narrative reviewers did not consider attitudinal outcomes at all in their reviews, and those who did tended to emphasize the negative effects of grouping on the attitudes and self-concepts of low-ability students. Such conclusions, however, were based primarily on anecdotal and uncontrolled studies. The controlled studies that we examined gave a very different picture of the effects of grouping on student attitudes. Students seemed to like their school subjects more when they study them with peers of similar ability, and some students in grouped classes even develop more positive attitudes about themselves and about school.

This meta-analysis thus confirmed some common beliefs about the effects of grouping, and it showed that other common beliefs are not supported by the facts. More important, however, this meta-analysis provided precise, quantitative estimates of the size of grouping effects, based on a large number of diverse studies. We believe these estimates give researchers and policy makers alike a new starting point for planning future policy and research in this area. ■

References

- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Rev. ed. New York: Academic Press, 1977.
- Ekstrom, R. B. "Experimental Studies of Homogeneous Grouping: A Critical Review." *The School Review* 69 (1961): 216-226.
- Findlay, W. G., and Bryan, M. M. *Ability Grouping: 1970, Status, Impact and Alternatives*. Athens, Ga.: University of Georgia, Center for Educational Improvement, 1971.
- Glass, G. V. "Primary, Secondary, and Meta-Analysis of Research." *Educational Researcher* 5 (1976): 3-8.
- Kulik, J. A.; Kulik, C.-L.C.; and Cohen, P. A. "Effectiveness of Computer-Based College Teaching: A Meta-Analysis of Findings." *Review of Educational Research* 50 (1980): 525-544.
- Otto, H. J. "Elementary Education—III. Organization and Administration." In *Encyclopedia of Educational Research*. Edited by W. S. Monroe. New York: The Macmillan Co., 1950.
- Whipple, G. M., ed. *The Grouping of Pupils*. Thirty-fifth Yearbook, Part 1. National Society for the Study of Education. Chicago: University of Chicago Press, 1936.

Copyright © 1982 by the Association for Supervision and Curriculum Development. All rights reserved.