DOCUMENT RESUME

ED 052 260 TM 000 770

AUTHOR Bryan, Miriam M.
TITLE Ability Grouping: Status, Impact, and Alternatives.
INSTITUTION ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO TMR-3
PUB DATE Jun 71
NOTE 15p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Ability Grouping, *Academic Achievement, Affective Behavior, Cluster Grouping, Culture Free Tests, Disadvantaged Youth, Early Childhood Education, *Educational Improvement, Educational Strategies, Environmental Influences, *Grouping Procedures, Heterogeneous Grouping, Homogeneous Grouping, Individualized Instruction, Low Achievement Factors, Minority Group Children, Predictive Validity, Self Concept, *Standardized Tests, Student Grouping, Team Teaching, Testing Problems, Tutoring

ABSTRACT

Ability grouping, defined as "the practice of organizing classroom groups in a graded school to put together children of a given age and grade who have most nearly the same standing on measures or judgments of learning achievement or capability," is considered in terms of its effect upon students' academic achievement and affective development. It is contended that grouping results in a trend toward improved achievement in superior groups and poorer achievement in the average or low groups. In addition, it tends to reinforce favorable self concepts in those assigned to high achievement groups and unfavorable self concepts in those assigned to low achievement groups. There seems to be a negative effect with regards to ethnic and socioeconomic separation. Misuses of standardized tests in grouping procedures are fully discussed. In view of the apparent inadequacies of present grouping practices, the following six alternative strategies are recommended: individualized instruction, heterogeneous grouping, stratified heterogeneous grouping, team teaching, student tutoring, and early childhood education. (Author/PR)

TM REPORTS

NUMBER 3

Ability Grouping: Status, Impact, and Alternatives

ERIC CLEARINGHOUSE ON TESTS, MEASUREMENT, & EVALUATION ■ EDUCATIONAL TESTING SERVICE, PRINCETON, NEW JERSEY 08540

Conducted by Educational Testing Service in Association with Rutgers University Graduate School of Education

The Clearinghouse operates under contract with the U. S. Department of Health, Education and Welfare, Office of Education. Contractors are encouraged to express freely their judgment in professional and technical matters. Points of view expressed within do not necessarily, therefore, represent the opinions or policy of any agency of the United States Government.

June 1971

ABILITY GROUPING:

STATUS, IMPACT, AND ALTERNATIVES

Miriam M. Bryan

ERIC Clearinghouse on Tests, Measurement, and Evaluation

ABILITY GROUPING: STATUS, IMPACT, AND ALTERNATIVES

How widespread is the use of ability grouping in the public schools of the United States? To what extent do tests represent an integral feature of ability grouping plans? What are the effects of ability grouping on the scholastic achievement and on the personal and social development of students so grouped? Is ability grouping likely to result in ethnic and socioeconomic separation within the school? What have test publishers done to determine and/or ensure the usefulness and the fairness of their tests for students who are culturally different? What have researchers reported concerning the reliability and validity of tests they have used with the disadvantaged student? What are some of the alternative strategies to ability grouping that have proved to be effective in the improvement of instruction?

These questions were among many to which answers were sought by a group of specialists in educational measurement commissioned in late 1969 by the U.S. Office of Education to study the status of ability grouping in American public schools and its impact upon the academic and affective development of school children (Findley and Bryan, 1971).

Ability grouping, as defined in that study, is "the practice of organizing classroom groups in a graded school to put together children of a given age and grade who have most nearly the same standing on measures or judgments of learning achievement or capability." Grouping and regrouping within a classroom for instruction in particular subjects is not considered to be ability grouping in the sense of this definition.

Ability grouping has been a topic of debate for more than half a century. The issue has, however, been brought into sharper focus during the last several years by three developments: (1) the launching of Sputnik and the consequent emphasis on special education for students with superior capabilities to meet the need for highly trained scientists; (2) increased attention to special education for the mentally and physically handicapped; and (3) emerging concern for equality of educational opportunity for all children, with obvious implications for the improvement and enhancement of that opportunity for those children to whom it has previously been denied.

In spite of the admission that homogeneous grouping by ability across the subjects of the school curriculum is impossible and in spite of conflicting evidence gathered over the years as to the benefits of ability grouping, such grouping is widely practiced in the nation's public schools. While grouping occurs in school districts of all sizes, it is especially characteristic of larger school systems; and while done at all grade levels, it is more common in the higher grades than in the lower grades. There is proportionately more grouping in the Northeast and Middle West than in other parts of the country.

While a relatively small proportion of schools rely on test scores alone for ability grouping, virtually all ability grouping plans depend on tests of aptitude and/or achievement as an integral feature. Findley and Bryan (1971 found that test scores alone constituted the basis for grouping in 13 per cent of the school districts reporting, but were among the multiple criteria reported by 82 per cent. Other criteria included teacher, counselor, and/or principal judgment, school grades, and student and/or parent interest, or a combination of these.

Although ability grouping is widely approved by school administrators and school teachers, opinion polls show that an overwhelming number of teachers express preference for average, mixed, or superior classroom groups over classes of low ability, in which emotional disturbance and rebellious behavior, as well as poor achievement, are likely to abound. Research on "streaming" (ability grouping) in England's schools indicates that the most detrimental effects occur in "non-streamed" classes taught by "pro-streaming" teachers. This generalization could apply equally well to American schools.

Early research studies on ability grouping were almost entirely concerned with the effect of grouping on academic achievement. While the evidence, then as now, was conflicting, the earlier studies more often than not reported gains by low groups and losses by high groups when compared with similar students taught in heterogeneous classes. More recent studies tend to show that separation into ability groups, when all children involved are considered, has no clear-cut positive or negative effect on average academic achievement, and the slight trend toward improving achievement in superior groups is counterbalanced by poorer achievement in the average and low groups, particularly the latter. One possible explanation for this difference is that in the earlier period the prevailing emphasis in instruction was on drill, with strong academic motivation accepted as a favorable but not a necessary characteristic, while today both strong academic motivation and academic achievement are emphasized; another is that low-achieving groups contain far more children of minority and low socioeconomic groups today than they did earlier, when the comparisons were between groups within a narrower range of ethnic and socioeconomic variation.

Research evidence regarding the effect of ability grouping on the affective development of students has, until recently, been very thin, perhaps because emotional and social growth is more difficult to assess than intellectual growth. As with the studies of impact on achievement, there has been little uniformity among the findings reported for the research studies that have been made. However, much of the evidence, especially the more recent evidence with ethnic and socioeconomic overtones, supports the generalization that the effect of ability grouping on the affective development of students is to reinforce favorable self-concepts in those assigned to high achievement groups and to reinforce unfavorable self-concepts in those assigned to low achievement groups. Low self-concept operates against motivation for academic achievement in all students, but especially among those from minority groups and lower socioeconomic backgrounds.

Most recently, researchers have become concerned with the effect of ability grouping on ethnic and socioeconomic separation. Here the evidence has been more conclusive. Students from minority groups and from unfavorable socioeconomic backgrounds tend to score lower on tests and to be judged less accomplished by teachers than students from middle-class homes. To the extent that these students are over-represented in low ability groups, then, they are being made to suffer the unfavorable results of ability grouping. A grouping plan which creates classes where disadvantaged students are in the majority deprives them of the stimulation of middle-class children as learning models and helpers, and commonly produces poorer achievement on their part. The greatest positive impact on the school learning of disadvantaged children occurs when the proportion of middle-class children in a group is highest.

Children of many minority groups come disproportionately from lower socioeconomic backgrounds. The disadvantages of their backgrounds are further compounded by language disabilities. For some of them, English, in which teaching and testing are generally done, is a "second language"; for others, the language patterns differ markedly from "standard American English." Language disabilities not only have the direct effect of making learning more difficult; they have the indirect effect of lowering self-concept because of frequent correction.

There have been no studies to date of the reliability and validity of tests administered to culturally limited populations for the specific purpose of ability grouping. As a matter of fact, until recently few publishers have studied the general usefulness of their tests with disadvantaged students. Now systematic efforts are being made by test publishers and research agencies to review present test offerings and to introduce new emphases to meet the particular problem of assessing the capabilities and achievement of the disadvantaged group.

The research that has been done to date shows that standardized aptitude tests, as they are currently constructed, are no less reliable for disadvantaged students than they are for others. They do, however, tend to overpredict for the disadvantaged group; that is, the disadvantaged student may not perform subsequently as successfully as his test scores indicate that he should. The same findings, in a slightly more limited way, apply also to standardized achievement tests. This is not to say that certain items in a standardized test may be more easily answered by students of one culture than by those of another, but, rather, that minority students who select the intended responses do not

always perform up to expectations. The evidence that tests standardized on other populations tend to overpredict the subsequent performance of disadvantaged students and, hence, are not unfair to them, is less than comforting. The challenge is to develop ways of describing learning progress directly rather than to settle for measures that are "fair" only in the sense that they reflect "fairly" the results of educational disadvantages.

Generally speaking, researchers are not studying or trying out and evaluating tests. They are studying other matters and, with few exceptions, accept uncritically the standardized test and/or use it as the best available instrument at hand. In the search of the literature concerning the use of tests in ability grouping and, especially, with the use of tests with the culturally deprived, several misuses of tests were noted. Among these, the following should be mentioned: (1) assuming that a test designed for students of a given age or of an estimated ability level can be used indiscriminately with students of different ages and/or experiences; (2) modifying the test in some material respect, but still applying the regular norms (for example, changing items or answers because of local circumstances; or translating the entire test into another language); (3) testing so early in preschool programs that culturally deprived children are not even ready to manipulate the test materials; (4) testing so early in preschool programs that there is no opportunity for children with limited backgrounds to "learn" the abilities tested; (5) using tests written in standard American English, with heavy emphasis on vacabulary, for students for whom this is a second language or who speak in a particular dialectic style; (6) testing very small numbers of students over a very short period of time; (7) failing

to follow through for two, three, or four years or more; (8) interpreting scores of individual students on short subtests when reliability estimates make it impossible to trust such interpretations; (9) treating different measures of learning ability as though the results on them were comparable; and (10) attaching the same importance to predictive validity without intervention (in the form of compensatory training) as with it.

The research concerned with ability grouping and with the procedures for the use of tests in grouping students for learning has provided only limited information. The design for the research procedures, the selection of tests, and the interpretation of test results have frequently been questionable. Most important, the research has produced inconclusive and conflicting results. This applies equally to the research findings concerning the advantages and disadvantages of ability grouping and to those regarding either the validity of currently available tests for use with culturally limited students or the validity of the interpretations of the test results for such students.

If, then, present ability grouping practices seem inadequate, what alternative strategies are there? The six suggestions which follow are not exhaustive of all possible alternatives, but they are judged to be the most promising for the promotion of learning:

1. Individualized instruction. There are almost as many definitions of individualized instruction as there are "authorities" defining the term. It is thought of here as instruction of the individual student, once his characteristics have been defined, by the pre-

scription of sequences of learning experiences leading to the mastery of basic skills and structural knowledge.

2. Heterogeneous grouping. This involves the putting together, in unselective fashion, of students who may vary extensively in age, experience, and knowledge and may, therefore, have opportunities to learn from one another that are not always provided by homogeneous grouping. Heterogeneous grouping of this kind is practiced in the nongraded school.

3. Stratified heterogeneous grouping. Grouping of this kind -- notably the Baltimore plan of stratified heterogeneous grouping by tens -- takes into account the concern for curtailing extreme heterogeneity, while allowing for enough diversity to give leadership opportunities in each class and avoiding the concentration of defeated and stigmatized students in a low group almost impossible to inspire or teach. In the Baltimore plan, ninety students ranked in order of excellence on some composite -- a standardized test battery, for example -- are then subdivided into nine groups of ten each. Teacher A is given a class consisting of the highest or first ten, the fourth ten, and the seventh ten; Teacher B has the second, fifth, and eighth tens; and Teacher C has the third, the sixth, and the ninth tens. In this kind of grouping there is no top or bottom section; each class has a narrower range than

the full ninety students have; teachers can give attention where it is needed without feeling that there are extremes whose needs are not being met; no teacher has to teach a class of disruptive children who lack both motivation and capability.

4. Team teaching. Several different models for team teaching have been developed. Each model embraces the concepts of individualized instruction, mastery, and differentiated staff working under the leadership of coordinating master teachers. Students who need to learn the same tasks may work in groups assigned to a designated teacher for the purpose of learning the special tasks. The grouping is informal, ad hoc, and of short duration. Such grouping promotes the effective utilization of personnel and resources, and increased learning by the individual student, without the detrimental effects of homogeneous grouping.

5. Student tutoring. In student tutoring plans, top students within a class may help those having difficulty with various subjects; or older children may be "imported," and perhaps paid, to tutor younger children who are having difficulty in learning the basic skills. Such tutoring works to the advantage of both groups of students. In fact, tutors who were themselves academically retarded have been found to gain even more than the tutored.

6. Early childhood education. Such education applies to the provision of opportunities for all children, especially those in need of compensatory education, to enjoy intellectual stimulation in a supportive emotional climate, at least from Kindergarten at age five and somewhat earlier when possible. Competence generated by the nature of early stimulation should increase the readiness of the children to participate in the conventional schooling of the primary grades.

Taken together, these alternative strategies constitute a constructive challenge to the uncertain advantages and the harmful effects of ability grouping on academic achievement, affective development, and the ethnic and socioeconomic separation of children. In each of them, tests and other evaluative measures may be used constructively if they are used with care and caution.

In conclusion, the following recommendations are offered: (1) Ability grouping, as defined, should not be used; however, flexible grouping within classes may be used to advantage when the information gained by testing and/or observation is the first step in a program of diagnosis and individualized instruction. (2) In any grouping plan, provision should be made, as part of the instructional program, for frequent review of each student's grouping status. (3) Alternative strategies for ability grouping should be explored and exploited for their usefulness in promoting learning. (4) Favorable self-concept should be a goal in itself, but it is also a supportive factor in learning. An attitude of firm confidence and hope by the teacher, fundamental to effective learning, should

be conveyed to every student. (5) Teacher training should include an emphasis on welcoming diversity in children, especially with regard to language and customs of minority groups, and on teaching children to prize it in each other. (6) Finally, steps should be taken as early as possible in each local situation to promote unitary school populations in each district and in each classroom. Action to improve instruction by any of the alternative strategies to ability grouping will be effective in proportion to the extent to which they can be applied before a district or city has become almost completely an ethnic and/or a socioeconomically limited population.

REFERENCE

Findley, W. G., & Bryan, M. N. Ability Grouping: 1970. Athens, Georgia: Center for Educational Improvement, College of Education, University of Georgia, 1971. (96 pp.) Free on request.