     RESEARCH FINDINGS ON ABILITY GROUPING ARE INCONCLUSIVE
BECAUSE NEITHER HETEROGENEITY NOR HOMOGENEITY HAS BEEN
DEFINED WITH SUFFICIENT CLARITY. THE TENDENCY IN THESE
STUDIES HAS BEEN TO STRESS THE PERFORMANCE OF THE PUPILS IN
SUCH CLASSES RATHER THAN THE PERFORMANCE OF THE CLASS AS A
WHOLE. IN A STUDY OF 181 CLASSES (4,705 PUPILS) HOMOGENEITY
WAS MEASURED BY THE STANDARD DEVIATION OF CLASS PERFORMANCE
ON THE FIRST TWO METROPOLITAN READING TESTS GIVEN IN TWO
SUCCESSIVE YEARS. GROWTH WAS THEN DETERMINED BY THE
DIFFERENCES IN CLASS MEANS ON THE TWO TESTS. THE SAME
SUBJECTS WERE TESTED IN GRADE THREE AND FOUR, AND WERE
DIVIDED INTO HIGH, AVERAGE, AND LOW LEVELS OF ACHIEVEMENT AND
DEGREE OF HOMOGENEITY. (A STANDARD DEVIATION OF 6.0 THROUGH
8.9 MONTHS CHARACTERIZED "AVERAGE HOMOGENEITY.") FINDINGS
SHOW AN INCONSISTENT GROWTH PATTERN--(1) ON THE WORD
KNOWLEDGE SUBTEST, MEAN GROWTH WAS PRACTICALLY IDENTICAL FOR
THE AVERAGE AND LOW HOMOGENEITY CLASSES, AND (2) ON THE
READING SUBTEST, THE LOW HOMOGENEITY CLASSES SHOWED GREATER
GROWTH THAN THE AVERAGE OR HIGH CLASSES. EVIDENCE OF
INCONSISTENCY WAS ALSO EVIDENT WHEN VARIOUS COMBINATIONS OF
INITIAL ACHIEVEMENT LEVEL AND CLASS HOMOGENEITY WERE
ANALYZED. THEREFORE, NARROWING THE RANGE OF ABILITY IN
CLASSES DOES NOT IPSO FACTO IMPROVE PUPIL ACHIEVEMENT.
PROGRAMS DESIGNED SPECIFICALLY FOR THE SEVERAL ABILITY LEVELS
ARE NEEDED AS A CONCOMITANT OF ABILITY GROUPING. THIS ARTICLE
WAS PUBLISHED IN "THE URBAN REVIEW," VOLUME 2, FEBRUARY 1967.
(NH)

ED013842

UD002 9

Ability Grouping — What Good Is It?

by Joseph Justman

If one were to ask an elementary school supervisor why he uses ability grouping in organizing his school at the beginning of each year, he would probably cite a number of reasons—pupil achievement is better, teachers find it easier to teach classes showing a narrow range of ability, the slower children do not become a hindrance to those who learn more readily, etc. Yet, when the research in the field is examined, the findings are generally inconclusive.*

To some degree, the conflicting results obtained in the scores of studies which have been conducted over the past 40 years is understandable. In most instances, the conditions under which the studies were conducted differed markedly. Moreover, most of the studies in the area of ability grouping compare the performance of pupils enrolled in "homogenous" and "heterogeneous" groups.

What is a "homogeneous" group? In most instances, the designation is a convenient administrative label. It is not a generic term. Whether or not a class is truly homogeneous depends on the spread of ability in the total population from which the class is drawn. It is not inconceivable that a so-called "heterogeneous" class drawn from a population with a narrow range will actually show less variation in ability than a so-called "homogeneous" class drawn from a broad-range population.

There is another shortcoming characteristic of the research in the field of ability grouping. Most of the studies tend to focus their attention not on the performance of the homogeneous or heterogeneous classes that have been formed, but on the performance of the children enrolled in such classes. The individual pupil, rather than the class, is the unit of analysis. The findings of a typical study are reported in the following terms: "Pupils enrolled in homogeneous groups, as contrasted with matched pupils enrolled in heterogeneous groups, tend to...." Somewhere in the program of analysis, the class has disappeared.

In view of the shortcomings noted above, there appears to be need for a study of ability grouping in which homogeneity would be strictly defined, and in which the class, rather than the pupil, would be the unit of analysis. Such a study is reported below. Homogeneity is defined in terms of the standard deviation of *class* performance on an initial test, and growth is measured in terms of differences in class means on initial and final tests.

Procedure

In a study conducted in the New York City schools, parallel forms of the Metropolitan Reading Test were administered in May of two successive years to all third-grade classes and to all fourth-grade classes in more than 75 schools. Those classes that had remained virtually intact (no more than two pupils had left or been added to the class) over the period of one year which had elapsed were identified. Because of mobility and of pupil absence on the date of testing, test data were not available for both years for every pupil. Classes in which data were not available for at least 20 pupils, and for at least 75 per cent of the pupils on register, were dropped. These restrictions effectively eliminated classes that were abnormally small, and classes for which only partial data were available.

on register, were dropped. These restrictions effectively eliminated classes that were abnormally small, and classes for which only partial data were available.

After these restrictions had been applied, data for a total of 4,705 pupils enrolled in 181 classes drawn from 42 schools remained available for further analysis. These classes were divided into three groups, based on mean initial test scores. Since the initial test was administered in May, when normal achievement would be represented by a grade score of 3.9, all classes in which the mean initial reading grade fell between 3.5 and 4.4 were classified as showing average achievement. The standard deviation at the initial testing was used to divide the classes in terms of homogeneity. A class was considered as showing average homogeneity if the standard deviation fell in the range from 6.0 through 8.9 months.

The Findings

A summary of the mean gains, in months, shown by the participating classes is presented in the following Table.

Mean Gains (Grade 3 to Grade 4) on Metropolitan Reading Test at Three Levels of Achievement and Homogeneity (in

Test 1 — Word Knowledge			
Homogeneity of Class		**High** 45 and over	**Average** 35 - 44
High 5.9 and Below	N	11	9
	Mean	90.7	11.2
Average 6.0 - 8.9	N	12	40
	Mean	18.1	14.1
Low 9.0 and Over	N	27	26
	Mean	14.7	12.6
Total	N	50	75
	Mean	16.8	13.2

Test 11 — Reading			
Homogeneity of Class		**High** 45 and over	**Average** 35 - 44
High 5.9 and Below	N	12	17
	Mean	19.1	11.2
Average 6.0 - 8.9	N	14	38
	Mean	13.4	11.0
Low 9.0 and Over	N	21	20
	Mean	17.1	13.9
Total	N	47	75
	Mean	16.5	11.8

The 181 classes, taken as a group, gained 13.1 months ledge and 12.5 months in Reading over the one year period and final testing. As one would expect, mean gains in achieve be positively associated with initial reading level. Classes achievement showed greater mean growth than those with achievement; the mean growth shown by the latter, in turn, that of classes with low initial achievement. This trend was n sections of the achievement test.

The same generalization could not be advanced, however, were divided into subgroups showing high, average, and l When this was done, mean growth in Word Knowledge was for classes showing average or low homogeneity, while on test, the mean growth of classes showing low homogeneity wa of classes with average or high homogeneity.

Lack of a consistent growth pattern was even more evide combinations of initial achievement level and class homog sidered. For example, greater growth in Word Knowledge wa with high initial achievement as class homogeneity increa showing low initial achievement, however, greater growth wa increasing heterogeneity. In the case of classes showing avera ment, the greatest growth was noted in classes with an avera and the least growth in classes with high homogeneity.

A similar pattern of inconsistency was noted in the Read the achievement test. For those classes showing high initial greatest mean gains were made by classes that were classified geneity category. In the case of classes showing average ini the greatest gains between initial and final tests were observ low homogeneity. For classes showing low initial achieve mean gains were noted in classes with average homogeneity.

ED013842

UDoo2 975

Grouping — What Good Is It?

astman

to ask an elementary school superv:sor why he uses ability group-
nizing his school at the beginning of each year, he would probably
ber of reasons—pupil achievement is better, teachers find it easier
sses showing a narrow range of ability, the slower children do not
indrance to those who learn more readily, etc. Yet, when the re-
ne field is examined, the findings are generally inconclusive.*

degree, the conflicting results obtained in the scores of studies
been conducted over the past 40 years is understandable. In most
he conditions under which the studies were conducted differed
Moreover, most of the studies in the area of ability grouping com-
erformance of pupils enrolled in "homogenous" and "heteroge-
ups.

"homogeneous" group? In most instances, the designation is a con-
ainistrative label. It is not a generic term. Whether or not a class
ogeneous depends on the spread of ability in the total population
the class is drawn. It is not inconceivable that a so-called "hetero-
lass drawn from a population with a narrow range will actually
ariation in ability than a so-called "homogeneous" class drawn
d-range population.

another shortcoming characteristic of the research in the field of
ping. Most of the studies tend to focus their attention not on the
of the homogeneous or heterogeneous classes that have been
on the performance of the children enrolled in such classes. The
pupil, rather than the class, is the unit of analysis. The findings
study are reported in the following terms: "Pupils enrolled in
us groups, as contrasted with matched pupils enrolled in hetero-
ups, tend to...." Somewhere in the program of analysis, the class
ared.

f the shortcomings noted above, there appears to be need for a
ility grouping in which homogeneity would be strictly defined,
h the class, rather than the pupil, would be the unit of analysis.
y is reported below. Homogeneity is defined in terms of the stan-
on of class performance on an initial test, and growth is measured
ifferences in class means on initial and final tests.

conducted in the New York City schools, parallel forms of the
n Reading Test were administered in May of two successive years
ade classes and to all fourth-grade classes in more than 75 schools.
s that had remained virtually intact (no more than two pupils had
added to the class) over the period of one year which had elapsed
ied. Because of mobility and of pupil absence on the date of test-
were not available for both years for every pupil. Classes in which
t available for at least 20 pupils, and for at least 75 per cent of the
egister, were dropped. These restrictions effectively eliminated
vere abnormally small, and classes for which only partial data were

were dropped. These restrictions effectively eliminated classes
normally small, and classes for which only partial data were avail-

e restrictions had been applied, data for a total of 4,705 pupils
181 classes drawn from 42 schools remained available for further
ese classes were divided into three groups, based on mean initial
ince the initial test was administered in May, when normal achieve-
be represented by a grade score of 3.9, all classes in which the
reading grade fell between 3.5 and 4.4 were classified as showing
evement. The standard deviation at the initial testing was used to
asses in terms of homogeneity. A class was considered as showing
nogeneity if the standard deviation fell in the range from 6.0
months.

of the mean gains, in months, shown by the participating classes
in the following Table.

Mean Gains (Grade 3 to Grade 4) on Metropolitan Reading Test Achieved by Classes at Three Levels of Achievement and Homogeneity (in Months)

Test 1 — Word Knowledge

Homogeneity of Class		Achievement Level of Class			
		High 45 and over	Average 35 - 44	Low 34 and Below	Total
High 5.9 and Below	N	11	9	30	50
	Mean	20.7	11.2	9.0	12.0
Average 6.0 - 8.9	N	12	40	24	76
	Mean	18.1	14.1	10.2	13.5
Low 9.0 and Over	N	27	26	2	55
	Mean	14.7	12.6	12.9	13.6
Total	N	50	75	56	181
	Mean	16.8	13.2	9.7	13.1

Test II — Reading

Homogeneity of Class		Achievement Level of Class			
		High 45 and over	Average 35 - 44	Low 34 and Below	Total
High 5.9 and Below	N	12	17	25	54
	Mean	19.1	11.2	9.1	12.0
Average 6.0 - 8.9	N	14	38	26	78
	Mean	13.4	11.0	11.2	11.5
Low 9.0 and Over	N	21	20	8	49
	Mean	17.1	13.9	9.5	14.6
Total	N	47	75	59	181
	Mean	16.5	11.8	10.1	12.5

The 181 classes, taken as a group, gained 13.1 months in Word Know-
ledge and 12.5 months in Reading over the one year period between initial
and final testing. As one would expect, mean gains in achievement tended to
be positively associated with initial reading level. Classes with high initial
achievement showed greater mean growth than those with average initial
achievement; the mean growth shown by the latter, in turn, was greater than
that of classes with low initial achievement. This trend was noted on both sub-
sections of the achievement test.

The same generalization could not be advanced, however, when the classes
were divided into subgroups showing high, average, and low homogeneity.
When this was done, mean growth in Word Knowledge was virtually identical
for classes showing average or low homogeneity, while on the Reading sub-
test, the mean growth of classes showing low homogeneity was higher than that
of classes with average or high homogeneity.

Lack of a consistent growth pattern was even more evident when various
combinations of initial achievement level and class homogeneity were con-
sidered. For example, greater growth in Word Knowledge was shown by classes
with high initial achievement as class homogeneity increased; with classes
showing low initial achievement, however, greater growth was associated with
increasing heterogeneity. In the case of classes showing average initial achieve-
ment, the greatest growth was noted in classes with an average homogeneity,
and the least growth in classes with high homogeneity.

A similar pattern of inconsistency was noted in the Reading subsection of
the achievement test. For those classes showing high initial achievement, the
greatest mean gains were made by classes that were classified in the high homo-
geneity category. In the case of classes showing average initial achievement,
the greatest gains between initial and final tests were observed in classes with
low homogeneity. For classes showing low initial achievement, the largest
mean gains were noted in classes with average homogeneity.

Conclusions

It is very clear that reducing the range of ability in these classes was not associated with increased achievement in reading. The lesson for the school administrator is equally clear—homogeneous grouping is not a panacea for educational ills. The school administrator who looks to homogenous grouping as a means of improving pupil achievement will find the process of little value unless definite programs, specifically designed for the several ability levels into which they group their classes, are developed. Grouping by itself, without curricular modification as a concomitant, will not give rise to the desired outcome of improved pupil performance.

*Miriam Goldberg and others. The Effects of Ability Grouping.
New York: Teachers College Press, 1966.

Dr. Joseph Justman is acting director of the Bureau of Research for the Board of Education of the City of New York.