DOCUMENT RESUME

ED 113 414 UD 015 552

AUTHOR Beard, Jacob G.; And Others
TITLE Relative Achievement Levels of White and Black Children Before and After Desegregation.
PUB DATE Apr 75
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association annual meeting (Washington, D.C., April 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS *Academic Achievement; *Caucasian Students; *Integration Effects; *Minority Group Children; Negro Youth; Racial Differences; Racial Integration; School Integration; Statistical Data; Test Results
IDENTIFIERS Florida; *Performance Contrasts

ABSTRACT

Changes in the relative academic achievement of Florida's black and white children over the last 13 years during which desegregation was taking place are investigated. The availability of achievement data for the entire black and white population, along with the 13-year time span between observations are seen as principal advantages. Data show that the academic achievement gap between white and black children was smaller in 1974 than in 1961. The amount of decrease in the gap is stated to be small, yet of practical significance for the four subject areas examined: reading, vocabulary, math computation, and math problems solving. Of these areas, convergence in the black and white distribution was greater for math computation than for the othersubtest areas examined. The use of different tests in 1961 and 1974 is held to make assessment of changes in the absolute performance level of black or white children difficult. It is suggested that the study be viewed as an evaluation of a social action program rather than a scientific experiment. (Author/AM)

RELATIVE ACHIEVEMENT LEVELS OF WHITE AND BLACK CHILDREN BEFORE AND AFTER DESEGREGATION [1, 2]

Jacob G. Beard, Kathleen Burk, and Ron Nungester
Florida State University

The purpose of this study was to determine whether the interval separating the academic achievement level of black and white students in the State of Florida decreased over the time span, 1961 - 1974. During that time period, Florida's school system, as well as that of many other states, changed from a racially segregated system to a desegregated one.

Desegregation has been undertaken to remedy a general condition which was believed by many to be inherently unacceptable, and by the courts to be unconstitutional in that racial segregation among schools is in itself a denial of equal opportunity. Although improvement in academic achievement of black students was not the primary reason for desegregating the schools, the effects of desegregation on the academic achievement levels of both black and white children has been a concern of educators and of citizens in general.

Carlson (1972) quoted James Coleman as insisting that

> School integration ... is the most consistent mechanism for improving the quality of education of disadvantaged children. Integration alone reduces the existing gap between black and white children by 30 per cent. All the other school factors together don't add up to nearly that much.

Two frequently cited studies have examined system-wide effects of desegregation on academic achievement. Stallings (1959) examined the results of desegregation on black and white children in the Louisville,

---

[1]Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D. C., April, 1975.

[2]The study was suggested by Dr. A. A. Abrahams, Florida A&M University, Tallahassee, Florida.

Kentucky schools. He found that both races gained in achievement after desegregation, as measured by standardized tests, but that blacks gained more than whites. Stallings attributed these gains to motivation. It should be noted that the schools involved in the study had been desegregated a relatively short while, from four months to one school year, that the climate in Louisville was "favorable" to desegregation, and that Louisville white children were not assigned to black teachers at that time.

Hansen (1960) reported that, in the five years following desegregation of the Washington, D. C., schools, achievement generally improved in all subjects and grade levels. However, the data used in his study do not provide clear descriptions of black performance before and after desegregation. The major inference which can be drawn from Hansen's study is that, given accompanying major improvements in the quality of education, school desegregation was accomplished without apparent losses in pupil achievement.

The present study deals with the broad question: Has the relative academic achievement of Florida's white and black children changed over the last thirteen years during which desegregation has taken place? This study should be viewed as an evaluation of a social action program rather than a scientific experiment. Cohen (1970) described such evaluative studies in the following way: "...evaluating social action programs is only secondarily a scientific enterprise. First and foremost, it is an effort to gain politically significant information on the consequences of political acts." Those interested in a study of situational determinants of black performance might well examine the review prepared by Katz (1964).

Principle advantages of this study over others surveyed were the availability of achievement data for the entire black and white populations of a large state, and the substantial time span, 13 years, between observations. The 1961 data were collected at a time when almost all schools were completely segregated. On the other hand, almost all students were attending desegregated schools in 1974 and most had attended such schools throughout their school careers.

METHOD

The State of Florida has conducted a state-wide every-pupil testing program at the junior high or middle school level since 1957. The program was offered in the fall of the ninth grade until 1971 when it was changed to the spring of the eighth grade. The results were available by race, black or white, for the year 1961, when the battery consisted of the School and College Ability Tests (SCAT) and the Metropolitan Achievement Tests (MAT), and for the year 1974 when the battery consisted of both whole and adapted tests from the Comprehensive Test of Basic Skills (CTBS). The number of students by race and years are shown in Table 1.

TABLE 1

Number of Students by Race and Year

Race	Number of Students	
	1961	1974
Black	12,501	28,036
White	65,405	96,148

Four content areas, two mathematical and two verbal, were selected for the analysis. The paired tests are listed below.

1961	1974
Verbal (SCAT)	Reading Vocabulary (CTBS)
Reading Comprehension (MAT)	Reading Comprehension (CTBS)
Math Computation (MAT)	Math Computation (CTBS)
Math Problem Solving (MAT)	Math Concepts and Problem Solving (CTBS)

Although the subtests of each pair were not identical, they were judged to be similar enough in content and test format to permit meaningful analyses.

Four methods of describing the data were used: (a) comparison of means, (b) comparison of mean P-values, (c) comparison of frequency distributions, and (d) ridit analysis. Since the population data were available, it was not necessary to utilize inferential techniques.

The means could not be compared directly because the tests were different in 1961 and 1974 and some subtests contained substantially different numbers of items. It was arbitrarily decided to express the difference between white and black means in terms of majority group standard deviation units. That is, the difference between means was divided by the white standard deviation for that subtest and that year. In order to make the mean performance of each group more interpretable, the means were divided by the number of items to yield a mean P-value, or mean item difficulty. Frequency polygons were also constructed to show the amount of overlap in the frequency distributions for the two groups in 1961 and in 1974.

In addition to the above analyses, a ridit analysis was used to express the shift in the differences between the two groups. The term "ridit" is derived from the initial letters of the phrase "relative to an identified distribution." Ridit analysis was developed by Bross (1958) and has been used in biometrics. The ridit analysis begins with the calculation of ridits for the reference group, which are identical to percentile ranks when the percentiles are based on the mid-point of the raw score

interval. Given the frequency distribution of any other group over the same categories, a mean ridit for this comparison group can be calculated. The mean ridit is computed by finding the sum of the products of the observed frequency of each category of the comparison group and the corresponding ridit from the reference group, then dividing by the total frequency of the comparison group. This mean ridit can be interpreted as the probability that a subject randomly selected from the comparison group will have scored higher than a randomly selected subject from the reference group. For example, if the mean ridit for the comparison group is .60, then the probability is .60 that a comparison group subject will have a higher score than a reference group subject. This application of ridit analysis is discussed in Fleiss (1973).

RESULTS

Table 2 presents the means, standard deviations, and differences between black and white means expressed in white standard deviation units for each test. The lengths of the intervals separating the means of blacks and whites in terms of white standard deviation units were smaller in 1974 than in 1961 for each of the four types of tests. The greatest increase in the relative achievement of blacks was in math computation, in which the difference between means in 1961 was approximately two white standard deviations and in 1974 was approximately one standard deviation. The smallest increase in the achievement of blacks relative to whites was in vocabulary. The decrease in the difference between vocabulary means was less than one-half standard deviation.

TABLE 2

Means, Standard Deviations, Differences between Means in White Standard Deviations, and Number of Items

Test	Mean		Standard Deviation		Differences between means	Number of items
	White	Black	White	Black		
Reading						
1961	30.92	17.82	7.43	7.94	1.80	44
1974	13.78	9.18	3.91	3.92	1.18	20
Vocabulary						
1961	35.31	16.11	11.80	8.17	1.63	60
1974	25.05	14.41	8.62	6.82	1.23	40
Computation						
1961	33.72	18.93	7.44	8.44	1.99	45
1974	31.89	21.31	10.82	9.19	.98	48
Problem Solving						
1961	33.14	19.10	8.08	7.40	1.74	48
1974	17.84	10.19	6.68	4.90	1.15	30

The mean item difficulty values for blacks and whites on each test and the differences between black and white difficulty values are presented in Table 3. With the exception of vocabulary, the mean P-values were smaller for whites in 1974 than in 1961. This could have occurred either because the test administered in 1974 were more difficult or because the achievement of whites has decreased since 1961. While the 1974 tests seemed to be more difficult for whites, the average performance of blacks increased from 1961 to 1974 on all tests except Mathematics Problem-Solving. The reading and math computation tests were slightly more difficult for whites in 1974 than in 1961, but less difficult for blacks in 1974. The vocabulary test was easier for both groups in 1974, and the problem solving test was more difficult for both groups in 1974. The differences between the mean difficulty levels for blacks and whites were smaller in 1974 than in 1961 for all tests.

TABLE 3

Mean Item Difficulty Values and Differences between Difficulty Values of Blacks and Whites

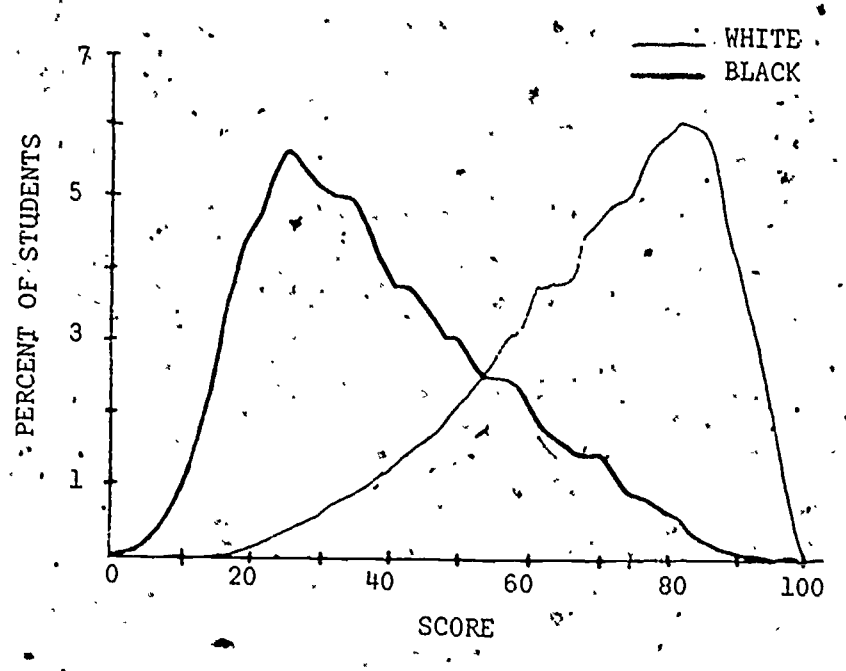
	1961			1974		
Test	White	Black	Difference	White	Black	Difference
Reading	.703	.398	.305	.689	.459	.230
Vocabulary	.589	.269	.320	.626	.360	.266
Computation	.749	.421	.328	.664	.444	.220
Problem Solving	.690	.398	.292	.595	.340	.255

Distributions of the black and white scores for both years and each test are presented in Figure 1. Percentages are plotted instead of frequencies so that areas under the curve are equal for black and white distributions.

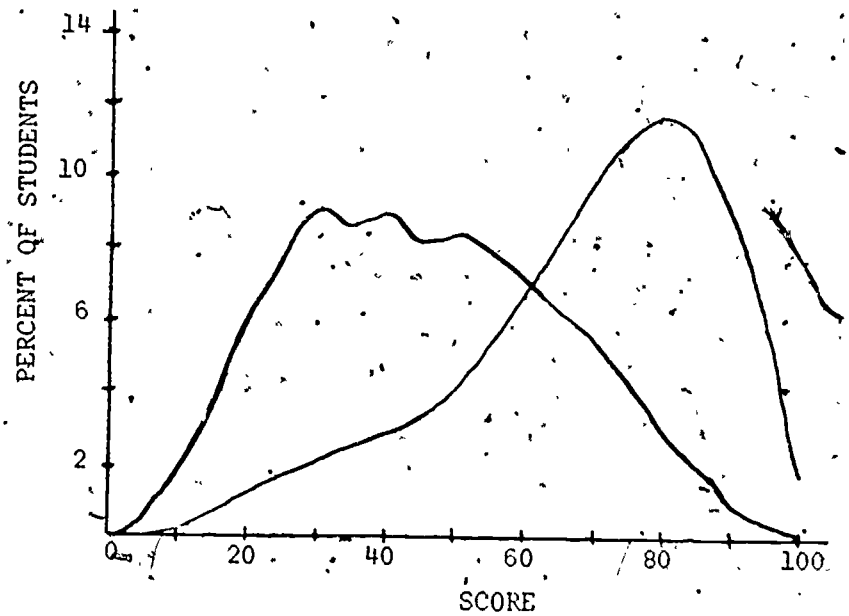
Although the distributions are not directly comparable due to the different characteristics of the tests used, the graphs indicated that there was more overlap of the black and white distributions in 1974 than in 1961 for each type of test. However, the reasons for the increased overlap in 1974 differ according to the type of test. For example, in reading the white distribution is about the same for both years and the black distribution appeared to be more similar to the white distribution in 1974. In math computation, the white distribution changed in 1974 to result in a greater amount of overlap with the black distribution, which appeared to change only slightly. Because different tests were used in 1961 and 1974, it was not possible to make any interpretations about changes in the absolute performance level of blacks or whites.

Ridit analysis was used for each test to determine the probability of a randomly selected black having a higher test score than a randomly selected white. If the distributions of blacks and whites were the same, the value of the ridit would be .50; i.e., the probability of a black with a higher score than a white would be the same as the probability of a white with a higher score than a black. Ridits for each test are presented in Table 4.

In 1961, ridits ranged from .096 for vocabulary to .120 for reading. All of the ridits for 1974 were higher than those for 1961. This indicated



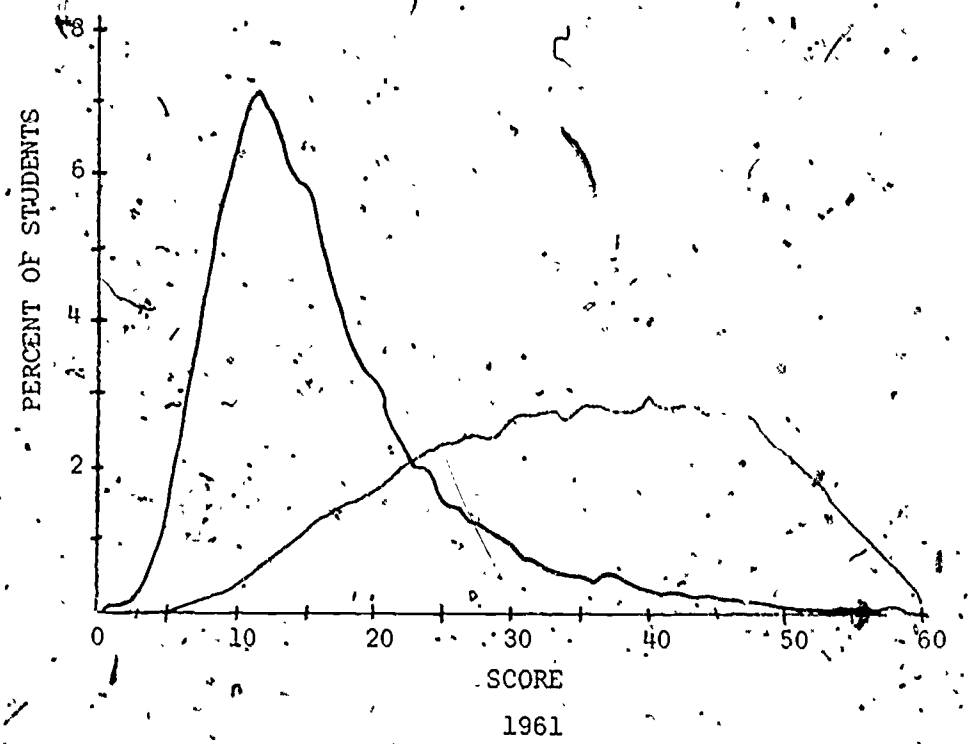
1961

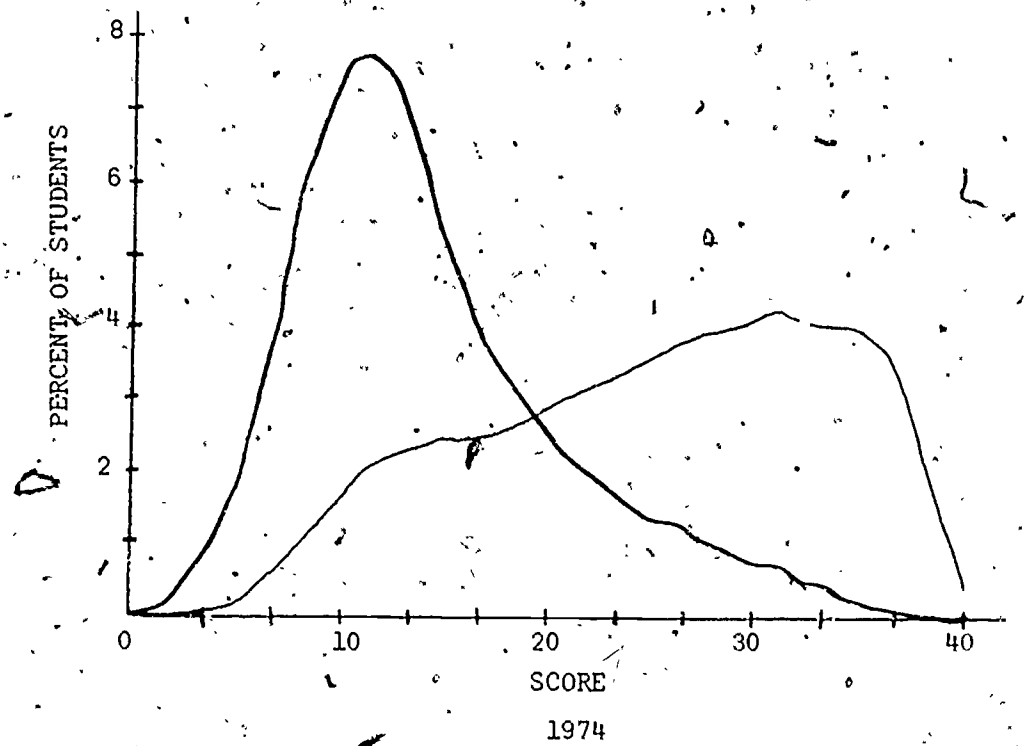


1974

READING

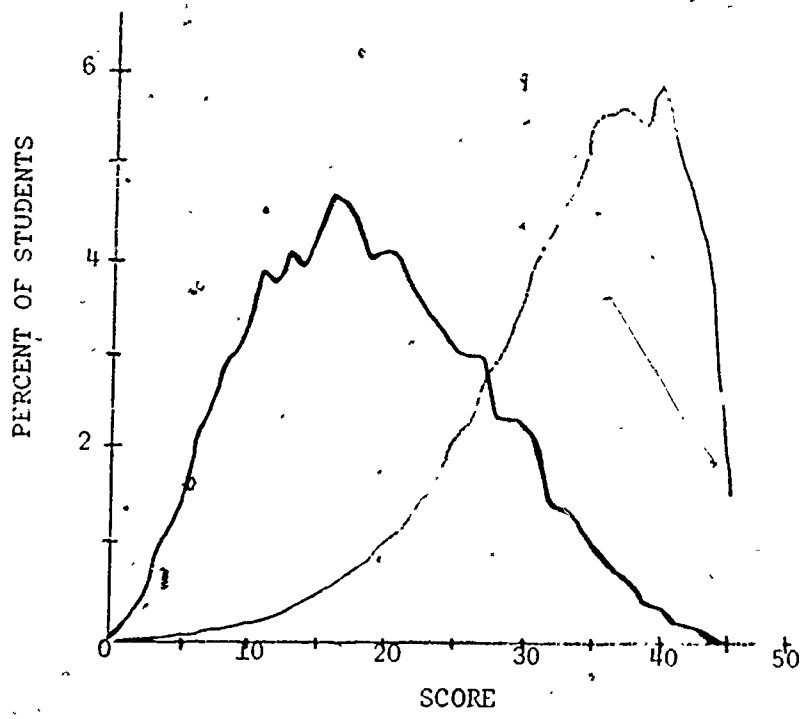
Figure 1. Frequency distributions of black and white test scores for 1961 and 1974





VOCABULARY

Figure 1. (continued) Frequency distributions of black and white test scores for 1961 and 1974.



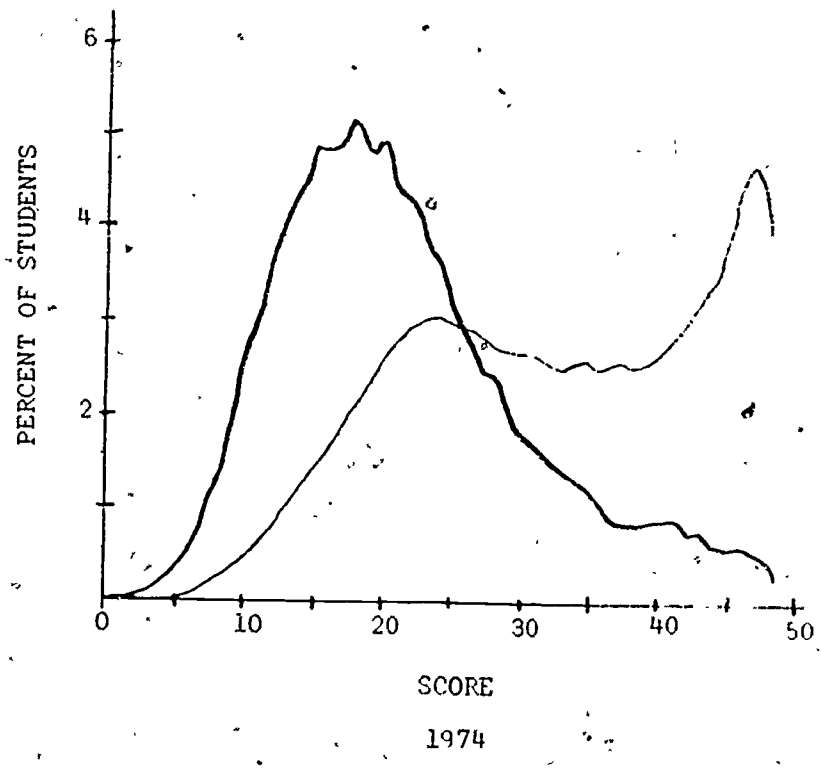
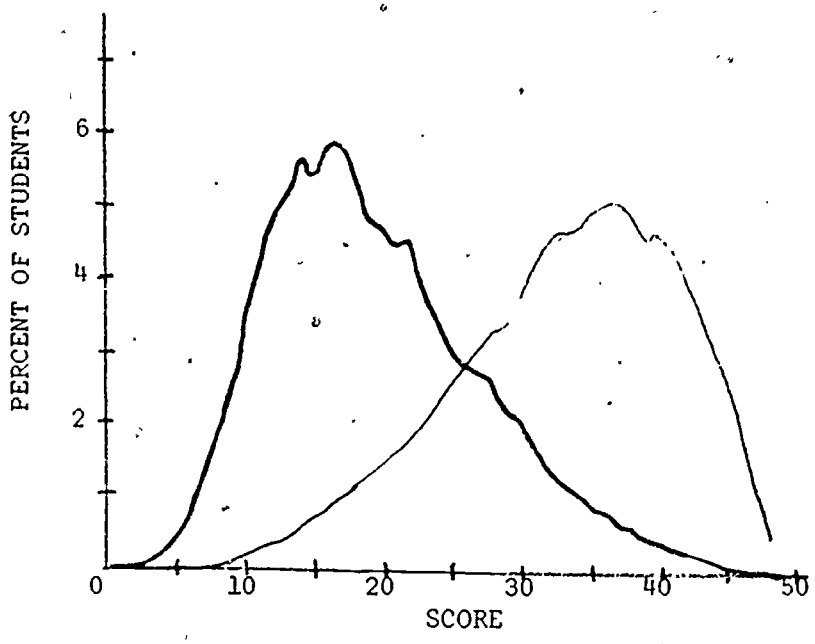
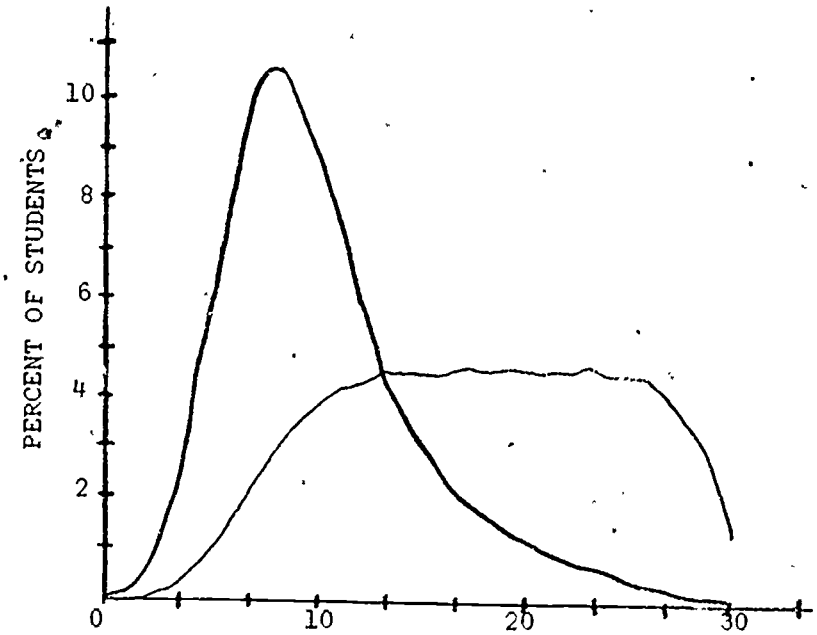


Figure 1. (continued) Frequency distributions of black and white test scores for 1961 and 1974.



1961



1974

Figure 1. (continued) Frequency distributions of black and white test scores for 1961 and 1974.

that the distributions of blacks and whites were more similar in 1974 than in 1961, although the distributions were still very different in 1974. The highest probability of randomly selecting a black with a higher score than a randomly selected white in 1974 was .231 for math computation. The corresponding probability in 1961 was .103. The increase in the ridit for math computation was consistent with the analysis of the differences between means, indicating that greater convergence of performance levels has occurred in math computation than in vocabulary, reading, or math problem solving.

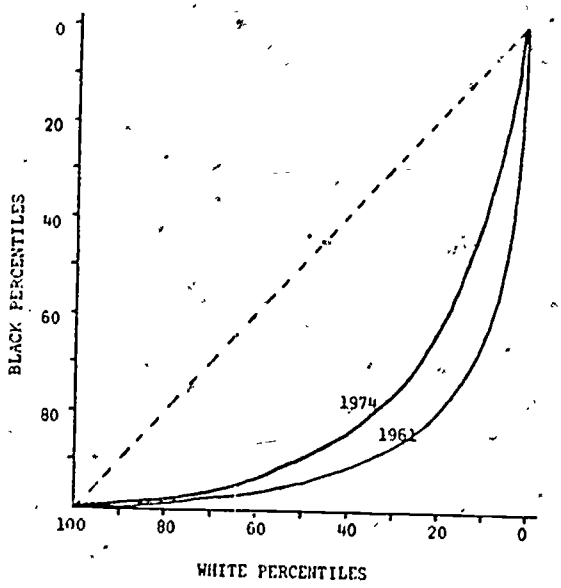
TABLE 4

Ridits by Test and Year

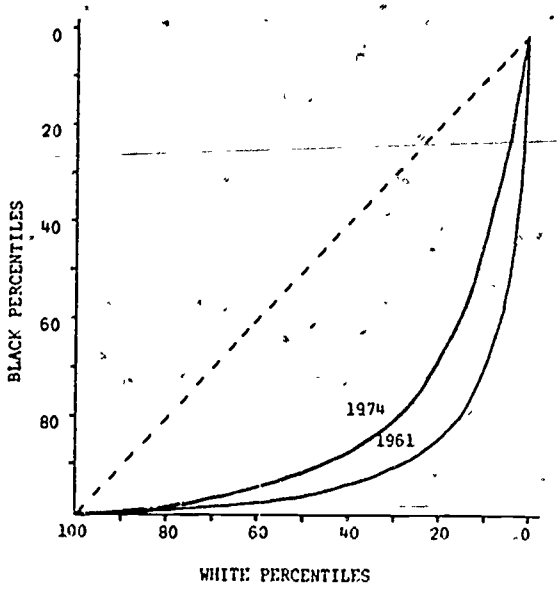
	TEST			
Year	Reading	Vocabulary	Computation	Problem Solving
1961	.120	.096	.103	.110
1974	.203	.175	.231	.180

A graphic representation of the information obtained from ridit analysis is presented in Figure 2. In each of the graphs, the ordinate represents black percentile ranks and the abscissa represents white percentile ranks. At each raw score, the percentile rank in the black distribution is plotted against the percentile rank in the white distribution. Any point on the curve illustrates the percentile rank received by a white and a black in their respective distributions for the same raw score. If black and white performance were the same, a straight diagonal line, as shown on the graphs, would appear. As the distribution of blacks and whites become more similar, the curve approaches the straight diagonal line. Comparison of the graphs for 1961 and 1974

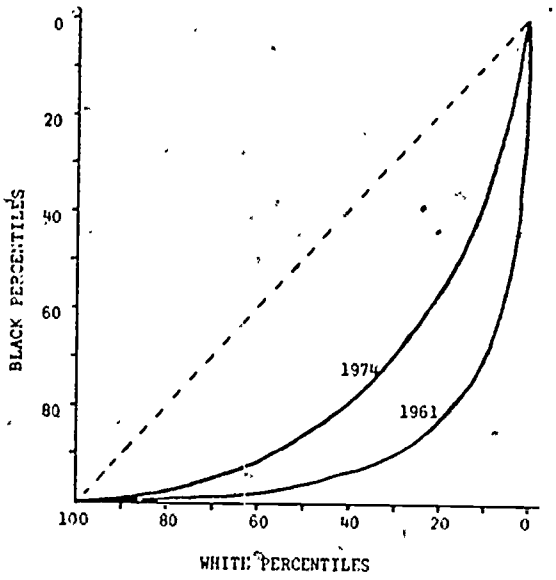
indicated that there was less discrepancy between black and white performance in 1974 than in 1961 in all of the types of tests. The greatest differences between the curves was evident in math computation. For example, on this test a black with a percentile rank of 80 in the black population in 1961 would have a percentile rank of 15 in the white population. A black with a percentile rank of 80 in 1974 would have a percentile rank of 40 in the white population.



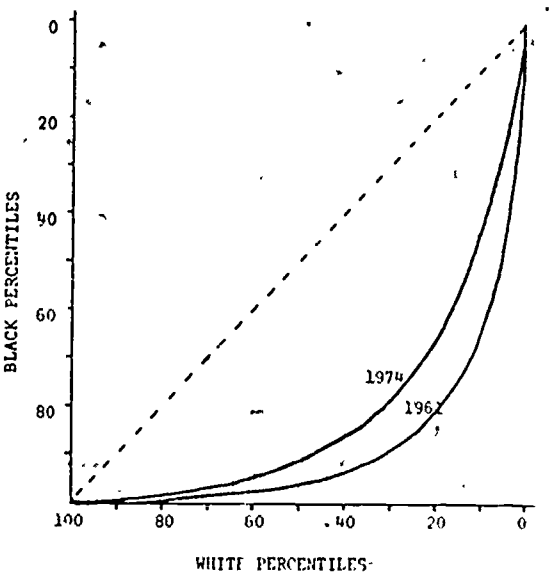
READING



VOCABULARY



MATH COMPUTATION



MATH PROBLEM SOLVING

Figure 2. Corresponding white and black percentiles at the same raw scores for 1961 and 1974

SUMMARY

The data consistently show that the academic achievement gap between white and black children was smaller in 1974 than in 1961. The amount of decrease in the gap is small but is of practical significance for the four subject areas examined: Reading, Vocabulary, Math Computation, and Math Problem Solving. Of the four areas examined, convergence in the black and white distributions was greater for math computation than for the other subtest areas examined.

Convergence in the black and white distributions could have been caused by improvement in the black group or by lowered performance in the white group. Examination of the frequency distributions suggested that the convergence for math computation might be at least partially caused by lowered white performance while convergence in reading appeared to result from improved black performance only. Unfortunately, the use of different tests in 1961 and 1974 made it difficult to assess changes in the absolute performance level of black or white children.

REFERENCES

Bross, I.D.J. How to use ridit analysis. Biometrics, 1958, 14, 18-38.

Carlson, K. Equalizing educational opportunity. Review of Educational Research, 1972, Vol. 42, No. 4, 464.

Cohen, D. K. Politics and Research: Evaluation of Social Action Progress. Review of Educational Research, 1970, Vol. 40, No. 2, 236.

Fleiss, J. L. Statistical Methods for Rates and Proportions. New York: Wiley & Sons, 1973.

Hansen, Carl F. The scholastic performance of Negro and White pupils in the integrated public schools of the District of Columbia. Harvard Educational Review, 1960, 30, 216-236.

Katz, I. Review of evidence relating to effects of desegregation on the intellectual performance of Negroes. American Psychologist, 1964, 19, 381-99.

Stallings, F. H. A study of the immediate effects of integration in scholastic achievement in the Louisville Public Schools. Journal of Negro Education, 1959, 28, 439-44.