DOCUMENT RESUME

ED 239 001 UD 023 304

AUTHOR Stephan, Walter G.
TITLE Blacks and "Brown": The Effects of School Desegregation on Black Students.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE 83
NOTE 46p.; Paper prepared for the National Institute of Education Panel on the Effects of School Desegregation. For related documents, see UD 023 302-308.
PUB TYPE Information Analyses (070) -- Viewpoints (120)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Gains; *Black Students; *Desegregation Effects; *Effect Size; Elementary Secondary Education; Mathematics Achievement; Meta Analysis; Outcomes of Education; Program Effectiveness; Program Evaluation; *Racial Relations; Reading Achievement; Research Methodology; Research Reports; School Desegregation; *Self Esteem; Standardized Tests
IDENTIFIERS Brown v Board of Education

ABSTRACT

This paper analyzes 19 studies that were considered by the National Institute of Education's panel on the effects of school desegregation in order to determine desegregation's effect on black achievement, black self-esteem, and racial relations. It is stated that, overall, the results show that reading achievement improves somewhat as a result of desegregation but math achievement does not. The author notes that there is a basic problem in evaluating desegregation programs in that there is great diversity among programs, which accounts for the diverse results obtained in different studies. It is suggested that desegregation has not increased black self-esteem, and in some cases has decreased it; tentative conclusions suggest that racial relations have not been improved by desegregation. According to the author, research, and particularly long-term research, is still needed on the effects of desegregation on students and communities, as well as studies that examine under what conditions desegregation does work. (CMG)

Blacks and Brown: The Effects of School Desegregation on Black Students*

Walter G. Stephan

New Mexico State University

Written for the National Institute of Education Panel on the Effects of School Desegregation

*The author wishes to thank Deanna Nielson for her assistance in preparing this article.

The Effects of Segregation and Desegregation

It is important to put the question of the effects of desegregation on Black achievement in historical context. To do this I would like to quote from social scientists and other expert witnesses who testified in the Brown (1954) trial. It is clear from their testimony that the social scientists believed that segregation had a negative impact on Black achievement in at least three ways.

First, the fact that segregated Black schools were inferior to white schools in terms of the quality of the facilities and per pupil expenditures was thought to lead to low levels of achievement. Prior to Brown it was not uncommon for Southern states to allocate from 2 to 5 times as much money per pupil for white students as was allocated for Blacks (Ashmore, 1954; Thompson, 1975). Also, Black schools in the South had teachers who were less well trained and who were paid about half as much as teachers in white schools (Ashmore, 1954). Conditions in Black schools were often appalling. Consider the findings of Matthew Whitehead who testified about the schools in Clarendon County, Carolina, during the Briggs vs. Elliot (1951) case.

> "The total value of the buildings, grounds, and furnishings of the two white schools that accommodated 276 children was four times as high as the total for the three Negro schools that accommodated a total of 808 students. The white schools were constructed of brick and stucco; there was one teacher for each 28 children; at the colored schools, there was one teacher for each 47 children. At the white high school, there was only one class with an enrollment as high as 24; at the Scott's Branch high school for Negroes, classes ranged from 33 to 47. Besides the courses offered at both schools, the curriculum at the white high school included biology, typing, and bookkeeping; at the black high school, only agriculture and home economics were offered. There was no running water at one of the two outlying colored grade schools and no electricity at the other one. There were indoor flush toilets at both white schools but no flush toilets, indoors or outdoors, at any

of the Negro schools--only outhouses, and not nearly enough of them" (Kluger, 1976, p. 332).

Second, it was thought that the "badge of inferiority" that segregation represented led Black students, and their teachers, to have low expectations regarding their capacities to achieve. These low expectations were believed to lead to low achievement. This argument can be traced through the testimony of several social scientists. David Krech said:

> "Legal segregation, because it is legal, because it is obvious to everyone, gives ... environmental support for the belief that Negroes are in some way different from and inferior to white people" (Kluger, 1976, p. 362).

In another trial Horace English testified that:

> "If we din it into a person that he is incapable of learning, then he is less likely to be able to learn ... There is a tendency for us to live up to-- or perhaps--I should say down to social expectations and to learn what people say we can learn, and legal segregation definitely depresses the Negro's expectancy and is therefore prejudicial to his learning" (Kluger, 1976, p. 415).

Third, in addition to reducing expectancies, segregation was also thought to reduce the motivation to learn among Black students. Brewster Smith testified that:

> "Segregation is, in itself, under the social circumstances in which it occurs, a social and official insult and ... has widely ramifying consequences on the individual's motivation to learn" (Kluger, 1976, p. 491).

And Louisa Holt argued that:

> "The fact that segregation is enforced ... gives legal and official sanction to a policy which is inevitably interpreted both by white people and by Negroes as denoting the inferiority of the Negro group ... A sense of inferiority must always affect one's motivation for learning since it affects the feeling one has for one's self as a person" (Kluger, 1976, p. 421).

In the original Brown (1951) decision this line of reasoning was sufficient to convince Judge Huxman that:

> "Segregation of white and colored children in public schools has a detrimental effect upon the colored children. The impact is greater when it has the sanction of the law; for the policy of separating the races is usually interpreted as denoting the inferiority of the Negro group. A sense of inferiority affects the motivation of a child to learn. Segregation with the sanction of law, therefore, has a tendency to retard the educational and mental development of Negro children and to deprive them of some of the benefits they would receive in a racially integrated school system" (Kluger, 1976, p. 424).

To summarize, it was because segregation was associated with inferior schools and led to low levels of expectancy and motivation in Black children that it was believed to cause low levels of achievement. At the time little or no data existed on the relative achievement levels of Blacks and whites in segregated schools. Thus, the argument rested on reason, not fact.

Because the Brown trials were concerned with the negative effects of segregation, minimal consideration was given to the anticipated effects of desegregation. In fact, desegregation as a remedy for segregation was rarely mentioned (Kluger, 1976). The social scientists' arguments concerning the effects of segregation implied that removing the "badge of inferiority" represented by segregation would increase the academic expectancies and motivation of Blacks and that these increases, along with improved facilities and instruction, would lead to higher achievement.

Subsequent theorizing about the effects of segregation and desegregation on Black achievement has elaborated on these basic notions. For instance, the U.S. Commission on Civil Rights' study of Racial Isolation in the Public Schools suggested that:

> "Negro children suffer serious harm when their education takes place in public schools which are racially segregated, whatever the source of such segregation may be. Negro children who attend predominantly Negro schools do not achieve as well as other children, Negro and white. Their aspirations are more restricted than those of other children and they do not have as much confidence that they can influence their own futures" (1967).

Jencks and his colleagues (Jencks, Smith, Aclard, Bane, Cohen, Bintis, Heyns & Michelson, 1972, pp. 97-98) offered four reasons why desegregation should improve Black achievement. First, they cited the anticipated positive effects of improvements in school and teacher quality. Second, they cited the knowledge that may be acquired from white peers who have been socialized into middle class white norms--the lateral transmission of values hypothesis (for evidence that this does not occur see Miller, 1981). Third, Jencks et al. suggested that teachers in desegregated schools may expect more from Blacks and this may lead Blacks to learn more. Fourth, desegregation may lead Blacks to expect that they have a better chance of making it in society which may motivate them to work harder and learn more (for a synthesis of many of these arguments see Linsenmeier & Wortman, 1978).

Achievement Tests

All of the studies to be considered in this analysis of the effects of desegregation on Black achievement employed standardized achievement tests. Any understanding of the results of these studies requires that some consideration be given to the nature of these tests. Achievement tests were developed to measure what students have learned. They consist of items that sample the general body of knowledge that schools are expected to teach. The items that are selected are those that discriminate best between students who have

learned a great deal and those who have not. Items which sample knowledge that everyone learns are not included. This restricts the type of knowledge sampled to that which is not always learned or taught.

The tests usually take one to three hours to complete. During this period students at the junior high school level attempt to answer approximately 85 multiple choice questions per hour. The content areas covered most thoroughly (and the only ones reported in most desegregation studies) are math and verbal skills. Some tests deal with science and social studies, but use less extensive coverage for these topics. Thus, these tests examine only a very restricted domain of achievement. This domain, verbal and math skills, is clearly important, but so too are other domains of achievement that are not measured. Among these other domains are knowledge of our political, economic, and legal systems, and knowledge of the history of our society and other countries.

Scores on these tests correlate reasonably well from year to year and they correlate reasonably well with tests designed to measure aptitude and intelligence (Jencks et al., 1972, p. 60; Wallach, 1976). However, neither achievement tests nor those designed specifically for the purpose are especially good at predicting college grades or later success in life (Jencks et al., 1972, p. 57).

The test that has been most extensively scrutinized in this regard is the Scholastic Aptitude Test (SAT) developed by the Educational Testing Service. More than 2,000 studies have examined the ability of this test to predict future academic performance. The

results indicate that the SAT correlates about .30 to .40 with first year college grades (Lord & Campos, cited in Linn, 1982). SAT scores do not correlate as well with overall college grades (Humphreys, 1968) nor do they predict whether or not students will finish college (Astin, 1970). Also, there is little relationship between SAT scores (or similar measures such as the GRE) and later success after college (Marston, 1971; McClelland, 1971). In sum, the SAT and most standardized achievement tests have high content and construct validity, but only low to moderate predictive validity.

We must be extremely cautious in interpreting the meaning of achievement scores. They reflect the amount of standard curriculum materials in the domain of math and verbal skills that students have learned. Thus, achievement scores may serve as an indicator of the quality of the math and verbal skills programs at the schools the students are attending, although the same material may be acquired in the home, from peers, or from the mass media. To the extent that desegregation has an effect on achievement scores, it may be caused by changes in the quality and amount of instruction in math and verbal skills, changes in the quality of the student body, or changes in the students' motivation to learn. The changes that do occur probably should not be interpreted as an indication that the students will subsequently be more or less successful in institutions of higher education or in economic terms.

I do not mean to imply that test scores are not important, but I believe they are often important for the wrong reasons. Scores on achievement tests are used as criteria to determine what tracks students will be assigned to and whether students will be admitted to

college. They are also important because students and teachers perceive the scores as an indication of ability and individual worth. In this way, these tests may place inappropriate limits on the aspirations and self evaluations of low scoring students and they may lead teachers to have low expectations for low scoring students (For evidence that teachers have low expectations see Mercer, Iadicola, & Moore, 1980).

Because these tests measure what students have learned, anything that affects how much material they are taught or their capacity to assimilate what is presented will affect achievement test scores. Curriculum changes, differences in styles of presentation and testing, and disruptions that influence the capacity of teachers to teach or students' ability or desire to learn are likely to have a negative impact on what students learn. Because many of the studies reported in the literature cover only the initial phases of school desegregation they are very likely to be affected by these factors. In particular, the learning environment is apt to differ from the students' previous experiences, especially for minority students. Some of these differences may be beneficial in the long run such as more demanding teachers, more competitive classmates, and greater diversity in the student body, but these factors may initially have negative effects on achievement. Other factors such as tension and conflict between groups, negative comparisons with better prepared students who are often higher in social class, and dealing with teachers who have little experience teaching minority group students probably have a negative impact and continue to do so.

Although achievement tests are designed to measure what students

have learned, scores on these tests are also affected by other factors. Most important among these other factors is the situation in which the tests are administered. In particular, high anxiety levels have a negative effect on performance, except for the very best students. It is possible that Black students taking these tests in desegregated schools experience more anxiety than Blacks in segregated schools. This is likely to be the case to the extent that achievement is emphasized in desegregated schools and the Black students feel academically inferior to or threatened by the white students.

Achievement tests are "speeded" which means that students have a time limit that is too short for many of them to finish all the items. This too may create anxiety, it also means that a premium is placed on motivation and attentiveness. Students who are not motivated to do well or who do not try hard will not score well on these tests. Lapses of attention that amount to 5 minutes during the testing hour will mean failing to answer about 7 questions (at the junior high level). This could affect the outcomes by more than 50 points (on tests that have a range of 200-800 with an average of 500). The tests are most likely to yield accurate results when the conditions of testing do not elicit high levels of anxiety and the students are motivated to do well and are attentive. While these factors would be expected to influence measures of achievement both before and after desegregation it would not be surprising to find that they had a more negative impact after desegregation.

The race of the examiner can also affect test performance. Blacks often perform better when the examiner is Black rather than

white (e.g., Katz, Roberts, & Robinson, 1965). It is frequently the case that as students move from segregated to desegregated schools the race of the examiners changes from Black to white. Regrettably, we have no information on the degree to which such factors actually have affected the results of the studies we are reviewing, but they should lead us to be cautious about interpreting these studies.

The Studies in the N.I.E. Study Set

Anderson

This early study examines an unusual early desegregation plan in which students in the numerical minority in a given school could transfer to schools in which their group was in the majority. Thus, students could transfer from desegregated to segregated schools. The study was done in Nashville in 1963. It followed students from the 2nd to the 4th grade. The Metropolitan Achievement Tests were used to measure reading and math achievement. The sample size was adequate (N = 34 in the desegregated group), but not large. It is possible that some of the students in the desegregated group were exposed to one year of desegregation prior to being pretested in the second grade. It appears from the report that this problem probably affected less than one-sixth of the students in this group.

Beker

Like most early studies, the desegregation that was examined in this study (1964) consisted of voluntary transfers. The study was done in a large Northern city. Two grade levels were included (grades 2 and 3). The sample sizes were very small and may yield unreliable results (N = 7 - 25). The study is a Fall-to-Spring

comparison of reading and math abilities done during the first year of desegregation (measured with the Stanford Achievement Test).

Bowman

This is one of the longer studies in the set. It runs from 1967 to 1970. A group of students was followed from grades 1 to 3 and another group from grades 3 to 5. The sample sizes were of moderate size (around 50 total at each grade level), but adequate. The students participated in the program voluntarily and it took place in a medium sized Northern city (Syracuse). Different tests, the Iowa Tests of Basic Skills and New York State's Tests, were used to measure achievement at the pretest and the post-test levels which makes changes in test scores somewhat difficult to interpret.

Carrigan

I did not calculate effect sizes for this study because I believe the control group cannot be used to assess the effects of desegregation. In this study the control group was attending desegregated schools (50% Black). Since this control group had already received the "treatment" of desegregation they provide a check primarily for maturation effects. Any changes in this group may be a consequence of ongoing exposure to desegregation which means that the differences occurring in this group are not a proper control for the differences in the "desegregated" group. Also, the "desegregated" group actually started out in a somewhat desegregated school (80% Black) so this is not an optimal group to measure the effects of desegregation.

Clark

This is one of the small number of studies in the set that was

done in the South. It is a study of a majority-to-minority transfer program that took place in 1969-1970. The sample size is adequate (N = 108 for desegregated group), but the duration of the study is brief, extending from Fall to Spring. This is the only study in the set that includes rural students. It covers only the sixth grade and provides both a test of reading and math (SCAT).

Evans

This study was done in Fort Worth during the 1971-1972 school year. The Iowa Tests of Basic Skills were given to 4th and 5th grade students in the Fall and Spring of that year. The court-ordered desegregation plan involved clustering elementary students and busing Black students (in grades 3-5) to achieve a degree of racial balance. The sample sizes were larger than in most of the other studies in this set (N = 179-393).

Iwanicki and Gable

I excluded this study because the "predesegregation" group had already been attending desegregated schools for a full academic year at the time of the "pretest." Thus, the predesegregation comparison is actually a cross-sectional comparison between a segregated control group and a group of students that has been desegregated for one year. This means that the measure of the effects of desegregation is a measure of the effects of the second year of desegregation. Since all of the other studies that I have included measured the first year of desegregation, including this study with the others may yield an inaccurate picture of the effects of desegregation. This would be particularly true if desegregation had a greater impact on achievement during the first year than during subsequent years.

Klein

This is a Fall-to-Spring examination of the effects of desegregation done in a small city (35,000) in the South. The students were in the tenth grade. The sample size was adequate (N = 38 in the desegregated group), but not large. The study was done in 1965. The desegregation plan was a voluntary one involving Black students who transferred from segregated Black schools to white schools. The tests used were the Math and English Cooperative Exams.

Laird and Weeks

This is an early study of the effects of desegregation (1964). It was done in a large Northern city (Philadelphia) over a 1-1/2 year time span. Desegregation was brought about by overcrowding in a segregated Black school. Parents in this school could request to transfer their children to white schools so desegregation was voluntary. Students in grades 4-6 were tested on the district's own verbal and math tests. The sample size at each grade level is modest (22-39), but acceptable.

Rentsch

This study was done on a voluntary desegregation plan in Rochester, New York, and covers a 2-year time period. There were adequate sample sizes (N = 27 to 33) to calculate effects in grades 3-5. The students were tested on reading and math skills (apparently using a test developed by the District). The students who attended the desegregated schools had previously attended schools that were 90% minority. Attrition was fairly high in this group (56%). Although this study provided analyses of both matched and

unmatched samples of segregated and desegregated students, I decided against using the analyses of the matched groups because the sample sizes were small (N = 9 - 13).

Savage

This study covered a longer time period than many of the others, 2 years, and it is one of the minority of studies that were conducted in the South (Richmand, Va.). Also, it is one of the relatively small number of studies examining senior high school students. The sample size is adequate (N = 42 in the desegregated group) to calculate reliable means for math and reading achievement on the Sequential Educational Progress Test. The study was conducted between 1969 and 1971 and examined a voluntary desegregation plan involving minority to majority transfers.

Sheehan and Marcus

This study was done in Dallas, Texas, and covers a 1-1/2 year period. It involves court ordered busing and it was done recently (1976-1978). In these regards it is more representative of urban desegregation programs than most of the other studies in the set. The fourth grade students were measured with the Iowa Test of Basic Skills. The sample size is very large (nearly 2,000). One drawback is that the degree of desegregation varied considerably within the desegregated sample (from 5% to 65% Black).

Slone

This is a study of the second year of school desegregation. Desegregation occurred during the 1963-1964 school year. The first measure of achievement was gathered in April 1965 and the second in March 1966. The predesegregation school was multi-ethnic (90%

minority, but only about 70% Black) and thus this study differs from the other studies of desegregation. Also, the "segregated" control group was attending a school that was 40% white. Since the predesegregation levels of achievement cannot be determined, the effects of desegregation cannot be evaluated.

Smith

This is a long-term study , covering 3 school years. It was conducted in Tulsa, Oklahoma. The students were pretested in seventh grade and post-tested in ninth grade. The sample size is larger than in most studies (N = 274). The Stanford Achievement Tests were used to measure math and verbal skills. The desegregated students were attending naturally integrated junior high schools. Unfortunately, no information was provided on the degree of segregation in Tulsa's elementary schools, but it is probably reasonable to assume a high level of segregation given that the study began in 1965.

Syracuse

This study of fourth grade students measured reading achievement (Stanford Achievement Test) in the Fall and Spring of the 1965-1966 school year. The number of students in the desegregated group was small, but adequate (N = 24). The type of desegregation program the students participated in is not specified in the report.

Thompson and Smidchens

This study of natural desegregation in the elementary schools of Ann Arbor was eliminated from the analyses because the students had been attending desegregated schools for 2 years before the predesegregation measures were obtained. Thus, this study lacks a true predesegregation measure. In addition, the "segregated" control

group was 58% white.

Van Every

This study was done in Flint, Michigan, and involves desegregation produced by locating a low-cost housing project in a previously all white neighborhood. The study covers a 2-year period, following students from the fourth to the sixth grade. The sample size is somewhat small (desegregated group = 22). The study was completed in 1969. The Science Research Associates tests for reading and math were used.

Walberg

This is a study of the Boston Metro Project in which urban Black students at all grade levels were voluntarily bused to suburban white schools. The performance of these Black students on the Metropolitan Achievement Tests for reading and math were compared to the performance of their siblings who remained in segregated Black schools. The study was conducted during 1968-1969. The sample sizes for the desegregated groups are moderate (N = 61-144), those for the segregated groups are smaller (N = 14-53), but still reasonably adequate.

Zdep

This is a study of a voluntary transfer plan in which urban Blacks could attend suburban schools. The students were very young (grade 2). The Metropolitan Readiness Test was used to measure reading and math ability in the Fall and during the Spring of the first year of desegregation. The study was done in 1968. The sample size was quite small and may not yield reliable results (N = 12 in the desegregated group). The report does not indicate where the

study was done.

In summary, the desegregation in these studies was typically voluntary (66% of the cases), the cities it occurred in were generally medium to large, the region was more often the North than the South, the schools the students attended were more frequently elementary schools than secondary schools ($\bar{X}$ grade level = 5.5), Blacks were very much in the minority in most of these schools, and most of the studies were conducted prior to 1970 ($\bar{X}$ = 1968).

Effect Sizes

The principal measune of interest to be extracted from these studies is the size of the effects of desegregation on the verbal and math achievement of Black students. To calculate these effect sizes the formulas proposed by Glass (1977) were employed. In calculating these effect sizes I have taken into consideration the duration of the study.

All of the studies included in the study set employ quasi-experimental designs in which one group of students is tested before and after desegregation. The results for these students are compared to those of a group of students who remain in segregated schools and who are pretested and post-tested at the same time as the desegregated group. The generic formula to obtain effect sizes in standard deviation units for this design is to calculate the difference between the desegregated and segregated groups at the pretest and divide this score by the standard deviation for the segregated group.

1) $\frac{\bar{X}_1 - \bar{X}_2}{S.D._2}$ = pretest difference

This score indicates the degree of pretest equality between the two groups. A similar score is then obtained for the post-test scores.

2) $\frac{\bar{X}_1 - \bar{X}_2}{S.D._2}$ = post-test difference

To derive an overall effect size the pretest difference (1) is subtracted from the post-test difference (2). This formula yields an index of the magnitude of the effects of desegregation in units that can be compared across studies.

The use of the standard deviation of the control group (the segregated group in this case) to calculate effect sizes was proposed by Glass (1977). It would be possible to use in place of this standard deviation a pooled standard deviation comprised of the average of the standard deviations of the experimental and control groups on the assumption that this would yield a more stable estimate of the standard deviation. This more complex approach would be justified if the standard deviations of the experimental and control groups differed substantially. This appears not to have been the case in the present set of studies. In no instance (on the pretest or the post-test) were there significant differences between the mean standard deviations of the segregated and the desegregated groups. Thus, it seemed reasonable to employ the simpler formula advocated by Glass.

In this set of studies the duration of desegregation varies considerably. In order to obtain an index of the effects of desegregation during the first year of desegregation I first divided the effect size (E) by the duration (D) of the study to yield an effect size per month. In calculating the duration of the study I used the total number of months the study covered and subtracted 3

months for each summer vacation period that was included. Thus, the duration measure reflects only the number of months the students actually spent in school. Next, I multiplied effect size per month by 8 to obtain an index of the effect size per year.

$$\frac{E}{D} \times 8 = \text{effect size per year}$$

The primary value of this index of effect size is that it avoids including together in subsequent analyses studies that vary in duration from 4 to 36 months. These scores were calculated separately for verbal and math achievement to determine if desegregation had differential affects on the two basic areas covered by achievement tests. Since some studies included more than one grade level I calculated effect sizes for each grade and for each study as a whole so that comparisons could be made using grade or study as the unit of analysis. The effect sizes for grade are presented in Table 1.

Using this procedure for calculating effect size per year assumes that desegregation has linear effects over time, at least over the first 3 years of desegregation. This is the easiest, and I believe, the most defensible assumption to make in dealing with the effects of desegregation over the first few years of desegregation. There are other plausible relationships, however. For instance, it might be predicted that if desegregation had positive effects, most of the benefits would accrue to the students during the initial year or two of desegregation after which little additional benefit would be derived. Alternatively, desegregation might be expected to have negative effects on achievement initially because of the negative conditions under which it so frequently occurs. Later, after

adjustments have been made, desegregation might be predicted to have beneficial effects. The curvilinear nature of these predictions makes them difficult to apply to the present studies. In this set of studies the assumption of linearity appears to be reasonable in the case of math where the correlation between the duration of the study and the effect size was marginally significant ($r = .48$, $p < .10$). In the case of reading, the correlation was not significant ($r = -.17$, ns.). Krol's (1978) study of effect sizes for achievement is consistent with the assumption that the effects over time are linear.

The manner in which the results of these studies are presented is highly variable. In some studies the means and standard deviations necessary to calculate effect sizes using the generic formula are reported, but in others the effect sizes must be calculated using F tests, T tests, analyses of difference scores or analyses of covariance. Strictly speaking none of the latter calculations is precisely comparable to the generic formula, since the derived standard deviations are calculated from the overall variance. In cases where only covariance analyses are available, the effect sizes are almost certainly overestimated. This means that the average effect sizes across this group of studies are only approximate estimates.

Using studies as the unit of analysis, the average effect size for the first year of desegregation (8 months) was .17 verbal achievement, while the average effect size for math achievement was .00 (Table 2). Using the effect size for each grade as the unit of analysis, the effects are .15 for reading and .00 for math. Dropping the four studies from the sample set that I excluded has little

effect on the results. Using studies as the unit of analysis, the mean effect size for verbal achievement including all the studies in the set is .14 and for math it is .04. These results appear to indicate that verbal achievement improves somewhat, but math achievement shows little effect as a result of desegregation. The difference between the X for reading achievement and the $\bar{X}$ for math achievement is marginally significant ($t = 1.96$, $p < .08$, Table 4).

One way to convey the magnitude of these effect sizes is to consider what it would mean in terms of a test, such as the SAT or the GRE, that has a $\bar{X}$ of 500 and a standard deviation of 100. The effect for verbal achievement would translate into a 17 point increase as a consequence of the first year of desegregation. The math effect would translate into no improvement. Another more approximate way of thinking about these figures would be to consider what the effects of desegregation are on the average percentile ranking of Black students on a standardized test. If desegregation improved verbal achievement .17 standard deviation units, this would raise the average percentile rank of Blacks about 5 percentage points during the first year of desegregation. For math there would be no changes in percentile rank due to desegregation.

Why would desegregation affect the reading achievement of Blacks and not their achievements in math? One possibility is that reading achievement may be improved by direct exposure to the language usage and vocabulary of white students and teachers. Learning middle-class vocabulary and syntax may aid test performance. Such an improvement would not be due to any changes in the quality of teaching, or changes in expectancies or achievement motivation, but simply to

being able to understand the tests and the content of the questions better. Similar improvements would not be expected for math because there is no parallel to this type of indirectly learned information in the case of math. Here no improvement would be expected unless there were changes in the quality of instruction or the students expectancies or achievement motivation increased.

In this set of studies, the magnitude of the effect sizes is unrelated to the region in which the studies were done, the size of the cities in which the studies were done, and the size of the samples (Table 3). There is a marginally significant negative correlation between the grade the students were in when they were desegregated and the size of the effect for reading achievement ($r = -.33$, $p < .10$). The relationship between grade and effect size is not significant for math ($r = .22$, ns). For reading this suggests that younger students benefited more than older students from desegregation. One explanation for this relationship is that exposure to white students (and in some cases, white teachers) may benefit students who have had little previous direct or vicarious contact with whites. This benefit probably consists of exposure to the type of vocabulary that achievement tests measure. Older students who have had more direct and vicarious contact with whites may benefit less from exposure to whites in desegregated schools because they have had more exposure to white middle-class language usage and vocabulary.

The correlation between the year the study was done and the size of the effect for reading is also marginally significant ($r = -.49$, $p < .10$, using studies as the unit of analysis). The correlation

between the year the study was done and math achievement is not significant ($r = -.32$). It is not clear why this effect exists for reading. One possibility is that the early studies tended to be of voluntary desegregation where only select students participated. These desegregation programs may have made special efforts to help the incoming students and these students were probably highly motivated to succeed. In contrast, students in mandatory desegregation programs and later voluntary programs may have received less special treatment and may not have been as motivated to learn. However, the effects of special treatment would be expected to affect both reading and math, and there was no relationship for math, although the direction of the correlation is the same.

It also appears that the effect size for reading was larger in school districts where the desegregation was voluntary rather than mandatory ($\bar{X} = .21$ voluntary, $\bar{X} = -.03$ mandatory). While this difference is statistically significant ($t = 3.15$, $p < .05$, using studies as the unit of analysis and the corrected effect sizes as the dependent measure), the number of districts in which desegregation was mandatory is so small ($n = 2$) that these results may not be reliable. The effect for math was not significant ($t = .25$, ns). The most likely explanation for these effects is that the students who participated in desegregation voluntarily were more motivated to get to know other students. This informal contact would have enabled them to acquire verbal skills that could have affected their test performances, but it would not have enabled them to acquire math skills that affect test performance.

I would like to argue that none of the relationships regarding

effect size, grade, year, city size, region, or type of desegregation should be regarded as conclusive because the effect sizes themselves are unreliable. Even the overall effect sizes that were obtained may not be meaningful. Given the variability in the effect sizes in these studies, the confidence limits are rather broad. The 95% confidence limits (the range within which the true population X is likely to fall, with only a 5% probability of being mistaken) for verbal achievement are .04 to .30, and the 95% confidence limits for math achievement are -.09 to +.09. Thus, in the case of reading achievement we can be reasonably confident that desegregation has an effect, although it may be very small indeed. In the case of math, desegregation appears to have no effects.

There are other reasons why the average effect sizes should be regarded with more than a little caution. In those studies involving multiple grades it is possible to examine fluctuations in the standard deviations of the students' achievement scores. For instance, in Rentsch's study the range in standard deviations for the verbal scores is 9.57 to 13.14, and the range for math scores is 6.52 to 13.37. Obviously, when these standard deviations are used to calculate effect sizes (using the generic formula) the magnitude of the effect size will depend on the standard deviation that is used. If the standard deviations are unstable, then the effect sizes will be correspondingly unstable. The lack of stability in standard deviations tends to be a problem with the studies where the sample sizes are small.

One reason that the studies with small samples have variable standard deviations consists of sampling problems (e.g., non-random

sampling). Fluctuations in standard deviations within studies may also occur as a consequence of variable conditions during test administration. Anyone who has given tests to elementary students is aware of how difficult it is to maintain standardized procedures. Large sample sizes compensate somewhat for this variability in testing conditions, but most of the studies reviewed here did not use large samples.

Even if the standard deviations were stable, the small sample sizes of many of these studies would result in means that may not be accurate. In order to be accurate to within .5 standard deviation units of the true population $\bar{X}$, a sample size of 15 is required. To be accurate to within .1 standard deviation units, requires a sample of 384. Thus, the mean values reported in the studies with small sample sizes are not likely to be measured accurately enough to provide reliable effect sizes. If there were a sufficient number of these samples, the errors of measurement would cancel each other out, but the number of samples is not large enough in this set of studies to lead to confidence in the summary figures concerning effect sizes. Also, the substantial variability in effect sizes suggests that the mean effect size may be distorted by extreme scores and indeed the effect size for verbal achievement is lowered to .13 if the median is used as a measure of central tendency rather than the mean. If the effect sizes were corrected for the unreliability of the achievement tests this would also lower the estimate of the verbal achievement effect size.

Another reason that the average effect sizes should be viewed with caution concerns methodological problems with the studies.

While these studies were chosen because they are the best ones available, they are not without their defects. The list of potential defects is a long one. Threats to internal validity include those already mentioned, small sample sizes, non-random samples, and fluctuations in standard deviations (suggesting unreliability of measures). In addition, the quality of the measures of achievement varies (some use measures developed within the district, others use tests standardized on white populations), attrition varies considerably across studies and threatens the validity of studies where it is high, and the segregated control groups are often of uncertain comparability to the desegregated groups.

Threats to external validity are comprised primarily of concerns with the non-representativeness of these samples of Black students and of this group of studies. Only students who are in desegregated schools at the end of the study are included in the post-test and often in the pre-test $\bar{X}$'s. Usually students who stay in the program are not compared to those who drop out to determine if they are different. Thus, we cannot be confident that the samples of desegregated students in these studies are representative of Black students generally. Also, the studies are mostly of voluntary desegregation in medium to large northern cities. The degree to which it is appropriate to generalize these results to mandatory desegregation in other regions of the country or to small cities and rural areas is unclear.

Glass (1977) in discussing meta-analyses as a research method suggests that "Respect for parsimony and good sense demands an acceptance of the notion that imperfect studies can converge on a

true conclusion" (p. 356). His argument relies on an example in which a set of studies are "similar in that they show a superiority of the experimental over the control group" (p. 356). However, this argument may not apply as forcefully to a set of studies, such as those on the effects of desegregation on Black achievement, in which the results are variable rather than similar. Under these circumstances, the variability in results may be interpreted in terms of methodological problems as parsimoniously as in terms of more substantive causes.

A Basic Problem in Evaluating Desegregation

Perhaps the most fundamental oversight of the social scientists involved in the Brown trial was in not giving due consideration to the manner in which segregation would be eliminated. They were not alone in this oversight, even the lawyers for the NAACP did not consider this problem in detail until after the first Brown decision in 1954. The Justices of the Supreme Court were vague in their recommendations saying in the second Brown decision in 1955 only that segregation should be ended with "all deliberate speed" (Kluger, 1976, pp. 714-747). When desegregation began to be implemented 10 years after Brown, the forms it took were as varied as the communities in which it took place. I believe it is this complexity more than any other factor that accounts for the diverse results that have been observed in studies of the effects of desegregation on achievement. The diversity of desegregation programs is so great as to render the word without a precise meaning.

Let me be specific about this complexity, although it is familiar to anyone who has studied the problem. Each community

starts with its own unique history of relations between the races including when Blacks and whites settled there, the origins of members of these groups, the social class structure of the groups, the degree of residential segregation and so on. The communities vary along such potentially important dimensions as size, region of the country, ratio of majority to minority group members, presence of suburbs and private schools to which whites may flee, and funding for public schools. The desegregation programs implemented in these communities have their own unique history of litigation and decision making by school boards and other public officials. The programs themselves vary in the techniques used to create desegregation, some programs are voluntary but most are not, the plans may involve voluntary cross-district busing, pairing, the use of magnet schools, the closing of some (usually Black) schools, and the mandated busing of students (usually Black students). The desegregation of teachers may or may not accompany the desegregation of students and the amount of preparation teachers are given for desegregation is variable. Additional curricular changes may occur at the same time as desegregation, the age of the students included in desegregation plans varies, the speed with which a plan is implemented varies, community opposition varies as does the amount of white flight, the ratio of majority to minority students differs from community to community as do the social class backgrounds of the students and the quality of their predesegregation educational experiences. As long as this list seems, it is surely incomplete. What these differences mean is that comparing the effects of desegregation across communities is extraordinarily difficult. It is possible to use

quantitative measures to examine the effects of some of the factors in this list, but the majority are more difficult to study and compare.

The Effects of Desegregation on Self Esteem and Race Relations

The social scientists who participated in the Brown trials believed that segregation had negative effects on the self esteem of Black students and on relations between the races, as well as having negative effects on achievement. One of the clearest presentations of their views comes from the statement that 35 social scientists filed as an Amicus Curiae brief in the Brown trial.

> Segregation, prejudices and discriminations, and their social concomitants potentially damage the personality of all children ... Minority group children learn the inferior status to which they are assigned ... they often react with feelings of inferiority and a sense of personal humiliation ... Under these conditions, the minority group child is thrown into a conflict with regard to his feelings about himself and his group. He wonders whether his group and he himself are worthy of no more respect than they receive. This conflict and confusion leads to self-hatred. ...
>
> Some children, usually of the lower socio-economic classes, may react by overt aggressions and hostility directed toward their own group or members of the dominant group. (Allport et al., pp. 429-430)

The social science brief and testimony in the individual trials leading up to Brown indicate that it was anticipated that ending segregation would remove the stigma of inferiority that was forced on Black children.

Self esteem. The effects of desegregation on self esteem appear to be less favorable than the effects of desegregation on achievement. In my earlier review (Stephan, 1978), I found that desegregation led to decreases in the self esteem of Black students

in 5 of 20 studies and that there were no studies indicating that desegregation increased Black self esteem. As was true for the studies of the effects of desegregation on achievement, the majority of these studies have been concerned with the effects of desegregation over a period of 1 year or less. One study that examined the effects over a longer period of time found that while Black self esteem initially dropped, it rebounded to predesegregation levels during the second year (Gerard & Miller, 1975). Subsequent studies of Black self esteem, including my own (Stephan & Rosenfield, 1978), have not changed this picture much. My conclusions regarding the effects of desegregation on the self esteem of Black students are consistent with those of other investigators (e.g., Banks, 1976; Epps, 1975; Gordon, 1977; Shuey, 1966).

It appears that the social scientists who participated in Brown used an invalid assumption as a basis for their argument that desegregation would increase the self esteem of Black students. Undoubtedly segregation stigmatizes Black students, but this stigma is not reflected in the self esteem of Black students. Studies of segregated Blacks and whites show that Black students have self esteem levels that are similar to or higher than white students in more cases than they have lower self esteem (see Porter & Washington, 1979 and Stephan & Rosenfield, 1979, for reviews). These studies have employed questionnaire measures of self esteem rather than indirect measures such as the doll tests upon which the social scientists' statements in Brown were based. The indirect measures may have been tapping attitudes toward Blacks and whites as ethnic groups. There is considerable evidence indicating that young

Black children have less favorable attitudes toward Blacks than toward whites (Williams & Morland, 1976).

If segregated Black students do not have low self esteem, there is little reason to expect that desegregation would increase self esteem. In fact, there are several compelling reasons why decreases in self esteem might be expected. For instance, social comparison with white students who are academically better prepared than Blacks could lead Blacks to evaluate themselves negatively. Likewise, the loss of status and power that occurs when Blacks represent a minority of the student body in desegregated schools could also lower the self-esteem of Black students. In addition, negative evaluations by ethnocentric white students could adversely affect the self esteem of Blacks.

Attitudes. The social scientists in their brief were also hopeful that contact within the schools would improve intergroup relations.

> Under certain circumstances desegregation ... has been observed to lead to the emergence of more favorable attitudes and friendlier relations between races. ...There is less likelihood of unfriendly relations when change is simultaneously introduced into all units of a social institution ... and when there is consistent and firm enforcement of the new policy by those in authority. ...These conditions can generally be satisfied in ... public schools. (pp. 437-438).

The social scientists appreciated the fact that contact alone would not be sufficient to improve intergroup relations. Their statement notes several preconditions for favorable change; equal status between the groups, and firm, thorough implementation of desegregation. It is likely that they were aware of other relevant factors such as those mentioned by Williams (1947) a half dozen years

before the social science statement was drafted:

> "Lessened hostility will result from arranging intergroup collaboration, on the basis of personal association of individuals as functional equals on a common task jointly accepted as worthwhile." (Williams, 1947)

The data on the initial effects of desegregation on race relations suggest that the social scientists' caution was well founded. In an earlier review of the data, I found that desegregation increased Black prejudice toward whites in almost as many cases as it decreased prejudice (Stephan, 1978). The results for whites were somewhat more negative. Recent studies, including my own, which also indicated that desegregation does not improve race relations (Stephan & Rosenfield, 1978), have not led me to revise these conclusions (e.g., Bullock, 1976; Campbell, 1977; Patchen, 1982; Sheehan, 1980). The quality of these studies is not as high as the better achievement studies and there is such a small enough number of them that these conclusions can only be regarded as tentative. My conclusions are, however, generally consistent with those of other investigators (Armor, 1972; Epps, 1975; St. John, 1975, Schofield, 1978; Weinberg, 1970).

In the years since Brown the contact hypothesis has been elaborated and refined. These elaborations are helpful in understanding why desegregation often has not had a positive effect on race relations. Here are my own most recent statements concerning the conditions under which contact improves intergroup relations.

1. Cooperation within groups should be maximized and competition between groups should be minimized.
2. Members of ingroup and outgroup should be of equal

status both within and outside of the contact situation.

3. Similarity of group members on non-status dimensions appears to be desirable (beliefs, values, etc.).
4. Differences in competence should be avoided.
5. The outcomes should be positive.
6. There should be strong normative and institutional support for the contact.
7. The intergroup contact should have the potential to extend beyond the immediate situation.
8. Individuation of group members should be promoted.
9. Non-superficial contact (e.g., mutual disclosure of information) should be encouraged.
10. The contact should be voluntary.
11. Positive effects are likely to correlate with the duration of the contact.
12. The contact should occur in a variety of contexts with a variety of ingroup and outgroup members.
13. There should be equal numbers of ingroup and outgroup members. (Stephan, 1983)

Desegregation rarely occurs under conditions that would lead to improvements in race relations. Instead, desegregation often occurs after there has been considerable community opposition from parents, administrators, school boards, and teachers. Thus, institutional and normative support for the contact is frequently low; the atmosphere tends to be competitive rather than emphasizing cooperation in pursuit of common goals; the statuses of Blacks and whites often are

unequal both outside the school (due to social class) and within the school (due to unbalanced ratios of Blacks and whites); the Black students are often not as well prepared academically as the whites so stereotype confirming differences in academic competencies frequently occur; busing often limits out-of-school contact and the within school contact that does occur is more likely to be negative or neutral than positive and in most cases it will be superficial. Also, the contact is involuntary in the case of court-ordered desegregation.

Recent research on the use of cooperative interethnic groups in desegregated schools indicates that when the conditions specified above are met, intergroup relations and self esteem improve without any costs in terms of lowered achievement (e.g., Aronson, Stephan, Sikes, Blaney & Snapp, 1978; Cohen, 1980; Cooper, Johnson, Johnson & Wilderson, 1980; De Vries, Edwards & Slavin, 1978; Weigel, Wiser & Cook, 1975). Other intergroup relations techniques involving multi-ethnic curricula, discussions of race issues, and explicitly providing information about the cultures of different groups have also been found to improve intergroup relations in the majority of cases (see Stephan, 1983; and Stephan & Stephan, 1983, for reviews). What these studies demonstrate is that while simply mixing students of different groups in desegregated schools does not improve race relations, intergroup relations can be improved in desegregated schools by introducing special programs designed to achieve this goal.

Future Directions for Research in Desegregation

I would like to see research into techniques to improve

achievement, race relations, and self-esteem continue. In addition, there are several other areas where I think research should also be done. One of the major problems with nearly all desegregation research is that it only covers the effects of the first year of desegregation, or at most the first two or three years of desegregation. There are almost no studies of the long-term effects of desegregation. We need to know not only what the long-term educational effects of desegregation are, but we also need to know what the non-educational effects are. And we need to know the effects not only for whites and Blacks, but also for other ethnic groups as well. Does school desegregation reduce segregation in other realms, such as housing; do minority students who have attended desegregated schools get better jobs and do they get promoted at a faster rate than students who attended segregated schools; and is subsequent political participation increased as a result of attending desegregated schools.

Also, we need to know more about the effects of desegregation on the communities that have undergone it. For instance, how do people in communities with well established desegregation programs feel about desegregation now; are people who have attended desegregated schools more willing to send their children to desegregated schools than people who attended segregated schools; and what differences are there in the race relations of communities with well established desegregation programs compared to other communities.

A third set of questions concerns the factors associated with successful desegregation programs. When desegregation goes well, why does it work? One can imagine a wide variety of factors that could

be relevant, some having to do with the community in which it takes place, others having to do with the way the administrators and teachers respond to desegregation, and still others with the composition of the student body. The fact is that we know precious little about what differentiates successful from unsuccessful desegregation programs.

Desegregation in Perspective

It would be impossible to present a comprehensive evaluation of the effects of desegregation in this short article. Instead, I have attempted to confine myself to some of the effects of desegregation on students. However, the larger context in which desegregation occurs is of immense importance to an understanding of the meaning of desegregation.

In order to put desegregation in perspective, we must consider the role that it has played in influencing relations between the races in our society. Since 1954, vast changes in race relations have occurred; many overt forms of discrimination have been eliminated, levels of prejudice have decreased, most minority groups have made economic advances, political participation by minority group members has increased dramatically, and more minority group members are attending college. School desegregation has played a role in these economic, political, and social changes, but it is a role that is not well understood and is little studied. Any analysis that abstracts school desegregation from its social context is necessarily incomplete. Unfortunately, we are not now in position to perform such an analysis. Given the difficulty of answering even a limited question like the effects of desegregation on Black

achievement, it doesn't seem likely to me that we will be in position to do an adequate comprehensive evaluation of desegregation anytime in the near future.

As we acquire more information on the outcomes of desegregation, we will be in a better position to base policy decisions on data. However, for the present, it seems to me that we will have to continue to making major policy decisions about desegregation on the basis of competing values. Some of these values concern the goals of public education, in particular the degree to which the schools should concern themselves with intergroup relations and the preparation of students to participate in a pluralistic society. Other decisions that we will continue to have to make pit the importance of creating equal educational opportunities against freedom of choice and freedom of association. Perhaps most importantly we will have to decide whether we value the elimination of segregation enough to continue the 50 year battle against it. Social science may be of less value in making these crucial decisions than in making choices about the best ways of implementing these decisions.

Table 1

Effect Sizes

Study	Grade	Effect for Reading	Effect for Math
Anderson	4	.42	.24
Beker	2	.19	-.31
	3	.06	-.17
Bowman	3	.25	.21
	5	.00	-.10
Carrigan*	1	-.41	
	2	-.02	
	3	.30	
	4	-.13	
	5	.33	
	6	-.31	
Clark	6	.08	-.24
Evans	3	.02	.03
	4	.02	.03
	5	.02	.03
Iwanicki & Gable*	3	.00	
	5	.00	
Klein	10	.23	.33
Laird & Weeks	4	.22	.18
	5	.31	.18
	6	-.14	-.17
Rentsch	5	.07	.02
	6	.26	-.08
	7	.33	-.10
	12	.06	-.04
Savage	4	-.07	-.08
Sheehan & Marcus			
Slone*	5	.19	.22
Smith	9	-.01	.02
Syracuse	4	.55	
Thompson & Smidchens*	5	-.15	.04
Van Every	6	-.12	.14
Walberg	4	.15	.07
	6	.05	-.53
	8	.17	.24
	11	-.15	.14
Zdep	2	.66	-.15

*Excluded from analyses

Table 2

Means for
Uncorrected Effect Size and
Effect Size Corrected for Duration of Study

Using Classes as the Unit of Analysis

	Reading	Math
Uncorrected $\bar{X}$	.24	.04
S.D.	.39	.34
Corrected $\bar{X}$	.15	.00
S.D.	.22	.20

Using Studies as the Unit of Analysis

	Reading	Math
Uncorrected $\bar{X}$	.24	.06
S.D.	.35	.25
Corrected $\bar{X}$	.17	.00
S.D.	.22	.16

Table 3

Correlations of Corrected Achievement Scores with Grade, Year, City Size and Sample Size

	Reading		Math	
	By Classes	By Studies	By Classes	By Studies
Grade	-.33**		.22	
Year	-.42*	-.49**	-.10	-.32
City Size	-.18	-.21	-.18	.19
Sample Size	-.28	-.38	-.05	-.17

*p <.05
**p <.10

Table 4

Reading vs. Math*

	Reading Effect Size	Math Effect Size	t	df	p
Uncorrected (Studies)	.21	.06	1.33	13	ns
Corrected (Studies)	.14	.00	1.96	13	.08
Uncorrected (Classes)	.21	.03	2.27	24	.04
Corrected (Classes)	.12	.00	2.52	24	.02

*The Syracuse study is excluded from this analysis because it did not include math achievement.

References

Allport, F.H., et al. The effects of segregation and the consequences of desegregation: A social science statement. Minnesota Law Review, 1953, 37, 429-440.

Armor, D.J. The evidence on busing. The Public Interest, 1972, 28, 90-126.

Aronson, E., Stephan, C., Sikes, J., Blaney, N., & Snapp, M. The jigsaw classroom, Beverly Hills, Calif.: Sage Publications, 1978.

Ashmore, H.S. The Negro and The School. New York: Van Rees Press, 1954.

Astin, A.W. "Racial Considerations in Admissions," in David C. Nichols and Olive Mills, eds., The Campus and the Racial Crisis, Washington, D.C.: American Council on Education, 1970, p. 87.

Banks, W.C. White preference in blacks: A paradigm in search of a phenomenonx, Psychological Bulletin, 1976, 83, 1179-1186.

Campbell, B.A. "The impact of school desegregation." Youth & Society, 1977, 9, 79-111.

Cohen, E.G. A multi-ability approach to integrated classrooms. Paper presented at the American Psychological Association, Montreal, Canada, September, 1980.

Coleman, J., et al. "Equality of educational opportunity," Washington, D.C.: Department of Health, Education & Welfare, 1966.

Cooper, L., Johnson, D., Johnson, R., & Wilderson, F. The effects of cooperative, competitive and individualistic experiences on interpersonal attraction among heterogeneous peers. Journal of Social Psychology, 1980, 111, 243-252.

De Lone, R.H. Small Futures. New York/London: Harcourt Brace Jovanovich, 1979.

De Vries, D.L., Edwards, K.J., & Slavin, R.E. Biracial learning teams and race relations in the classroom: Four field experiments using teams--fames--tournaments. Journal of Educational Psychology, 1978, 70, 356-362.

Epps, E.G. The impact of school desegregation on aspirations, self-concepts and other aspects of personality. Law and Comtemporary Problems, 1975, 39, 300-313.

Ford, S.F., Campos, S. Summary of Validity Data from the Admissions Testing Program cited in R.L. Linn, "Admissions testing on

trial," American Psychologist, 1982, 37, 279-291.

Gerard, H.B., & Miller, N. School Desegregation. New York/London: Plenum Press.

Gordon, V.V. The Self-concept of Black Americans. Washington, D.C. University Press of America, 1977.

Hoyt, D.P. The Relationship Between College Grades and Adult Achievement: A Review of the Literature, ACT Research Report, Iowa City: ACT Research and Development Division, 1965.

Humphreys, L.G. "The fleeting nature of the prediction of college academic success." Journal of Educational Psychology, 1968, 59, 375-380.

Jencks, C., Smith, H.A., Bane, M.J., Cohen, D., Gintis, H., Heyns, B., Michelson, S. Inequality. New York/London: Basic Books, 1972.

Katz, I., Roberts, S.O., & Robinson, J.M. "Effects of task difficulty, race of administrator, and instructions on digit-symbol performance of Negroes." Journal of Personality and Social Psychology, 1965, 2, 53-59.

Kluger, R. Simple Justice. New York: Random House, Inc., 1975.

Krol, R.A. "A meta-analysis of comparative research on the effects of desegregation on academic achievement." Unpublished Ph.D. dissertation, Western Michigan University, 1978.

Linsenmeier, J.A.W. & Wortman, P.M. The Riverside school study of desegregation: A re-examination. Research Review of Equal Education, 1978, 2, 3-36.

Marston, A.R. "It is time to reconsider the Graduate Record Examination." American Psychologist, 1971, 26, 653-655.

McClelland, D.C. "Testing for competence rather than for intelligence." American Psychologist, Vol. 28, No. 1, January, 1973, 1-14.

Mercer, J.R., Iadicola, P., & Moore, H. "Building effective multiethnic schools: Evolving models and paradigms." In W. Stephan and J. Feagin, School Desegregation: Past, Present and Future. New York: Plenum, 1980.

Miller, N. "Changing views about the effects of school desegregation." In M. Brewer and B. Collins, Scientific Inquiry and the Social Sciences, San Francisco: Jossey Bass, 1981.

Schofield, J.W. School desegregation and intergroup relations. In D. Bartal & L. Saxe (Eds.), Social Psychology of Education:

Theory and Research, Washington, D.C.: Hemisphere, 1978.

Sheehan, D.S. "A study of attitude change in desegregated intermediate schools." Sociology of Education, 1980, 53, 51-59.

Shuey, A.M. The Testing of Negro Intelligence, (2nd ed.). New York: Social Science Press, 1966.

Stephan, W.G. "Intergroup relations." In G. Lindzey & E. Aronson (Eds.), The Handbook of Social Psychology, Reading, Mass.: Addison-Wesley, 1983.

Stephan, W.G. "School desegregation: An evaluation of predictions made in Brown vs. the Board of Education." Psychological Bulletin, 1978, 85, 217-238.

Stephan, W.G., & Stephan, C.W. "The role of ignorance in intergroup relations." Desegregation: Groups in Contact. New York: Academic Press, 1983.

Stephan, W.G., Rosenfield, D. "Effects of desegregation on racial attitudes." Journal of Personality and Social Psychology, 1978a, 36, 795-804.

Stephan, W.G., Rosenfield, D. "The effects of desegregation on race relations and self-esteem." Journal of Educational Psychology, 1978b, 70, 670-679.

Stephan, W.G., Rosenfield, D. "Black self-rejection: Another look." Journal of Educational Psychology, 1979, 71, 708-716.

St. John, N. H. School Desegregation: Outcomes for Children. New York: John Wiley & Sons, 1975.

Thompson, E.T. Plantation Societies, Race Relations, and the South: The Regimentation of Populations. Durham, N.C.: Duke University Press, 1975.

Wallach, M.A. "Tests tell us little about talent." American Scientist, 1976, 64, 57-63.

Weigel, R.H., Wiser, P.L., & Cook, S.W. "The impact of cooperative learning experience on cross-ethnic relations and attitude." Journal of Social Issues, 1975, 31, 219-244.

Weinberg, M. Desegregation Research: An Appraisal. Bloomington, Ind.: Phi Delta Kappa, 1970.

Williams, J.E., Morland, J.K. Race, Color, and the Young Child. Chapel Hill: University North Carolina Press, 1976.

Williams, R.M., Jr. The reduction of intergroup tensions: A survey of research on problems of ethnic, racial, and religious

group relations. New York: Social Science Research Council, Bulletin 57, 1947.