DOCUMENT RESUME

ED 124 672 UD 016 124

AUTHOR Acland, Henry
TITLE Secondary Analysis of the Emergency School Assistance Program.
INSTITUTION Rand Corp., Santa Monica, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE Dec 75
CONTRACT 300-75-0106
NOTE 95p.; For Executive Summary, see UD 016 123

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Gains; Achievement Rating; Changing Attitudes; Evaluation; Federal Programs; *Followup Studies; *Grade 10; High School Students; Integration Effects; Integration Litigation; Integration Studies; *Negro Achievement; Negro Education; *Negro Youth; *Performance Factors; Program Effectiveness; Racial Composition; Racial Factors; Racial Integration; School Integration; Secondary Education; *Success Factors
IDENTIFIERS Emergency School Aid Act; *Emergency School Assistance Program; ESAA; ESAP

ABSTRACT

The original evaluation of the effect of Emergency School Assistance Programs (ESAP) in the second year, indicated that black male tenth graders in schools receiving the emergency aid did significantly better on a test of academic performance than did similar students in non-funded comparison schools. No program effects were reported for other groups of students. The analysis reported here accounts for the programs effect on black males. The orientation of this investigation is exploratory and not confirmatory in nature, and this is reflected in all stages of the work. The aim is to search for clues, not clear-cut certainties, about the existence and causes of program success. Basically five issues are addressed as follows: (1) did ESAP alter the racial climate of the funded schools; (2) are there relationships between student educational performance and non-cognitive variables that are dependent on his race and sex; (3) what did ESAP buy; (4) can teacher and student non-cognitive variables be related to particular educational programs; and, (5) can educational achievement levels be linked to non-cognitive variables and these in turn to educational programs. It is suggested that ESAP changes the attitudes and behaviors of white males and black females in a positive direction. (Author/AM)

SECONDARY ANALYSIS OF THE EMERGENCY SCHOOL ASSISTANCE PROGRAM

PREPARED FOR THE U.S. OFFICE OF EDUCATION,
DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

HENRY ACLAND

R-1868-HEW
DECEMBER 1975

Rand
SANTA MONICA, CA. 90406

The work upon which this publication is based was performed pursuant to Contract 300-75-0106 with the U.S. Office of Education, Department of Health, Education, and Welfare. Views or conclusions contained in this study should not be interpreted as representing the official opinion or policy of the Department of Health, Education, and Welfare.

Published by The Rand Corporation

PREFACE

The Emergency School Assistance Program (ESAP) provided financial aid to schools during 1970-1971 and 1971-1972 to help them in the process of achieving "successful desegregation." The original evaluation of the program's effect, in the second year, indicated that black male 10th graders in schools receiving the emergency aid did significantly better on a test of academic performance than did similar students in unfunded comparison schools. No program effects were reported for other groups of students: 10th grade black females, 10th grade whites (both sexes), and 5th graders (all four race-sex groups).

The analysis reported here is an attempt to account for the program effect on black males. Could the effect be better explained by what the schools did with ESAP money or by changes brought about in the attitudes and behavior of students or teachers?

The National Advisory Council on Equality of Educational Opportunity (NACEEO) asked the Office of Education to pursue this interesting result with the aim of finding any useful information about the operation of ESAA, the successor to ESAP, which the council monitored. The National Council knew that this secondary analysis would be exploratory, not confirmatory, in nature. This orientation is reflected in all stages of the work; the aim is to search for clues, not clear-cut certainties, about the existence and causes of program success.

The Office of Education contracted with Rand for the task of reanalysis. The work has been reported in two parts. The first reevaluated the finding that ESAP resulted in better performance for black male students. The second, reported here, seeks to explain how this effect came about.

SUMMARY

At a time when many schools in the southern states were going through the transition from segregated to desegregated status, the federal government initiated the Emergency School Assistance Program (ESAP) to help them through a difficult phase. The program supplied financial assistance to target schools, which, on average, increased their annual budget by an estimated 3 percent. This funding paid for all sorts of activities that were authorized by the regulations with the common intention of meeting "special needs incident to the elimination of racial segregation."

The program was administered, in part, on an experimental basis so that within a set of comparable pairs of schools some, chosen at random, received the funding while others did not. It was therefore possible to investigate the program's effect by comparing experimental and control schools. The original evaluation (NORC, 1973) showed that black male 10th graders performed significantly better, on average, than similar students in the control schools.

There are five issues addressed here:

(1) Did ESAP alter the racial climate of the funded schools?

NORC had suggested, but not proved, that the higher levels of performance for black males was attributable to improvements in the racial climate of the ESAP-funded schools. This change cannot be detected with these data (Section II). But if the students are divided into four groups (black males, black females, white males, white females) it is found that ESAP had a differential effect on students' attitudes and perceptions of their teachers and schools. However, these race and sex linked effects do not impinge on the black males, so the higher achievement levels of black males cannot be attributed to the intervening effect the program had on their attitudes. It is suggested that there was an indirect effect that worked on white students who, in turn, influenced black males; but this affords a tenuous explanation of how ESAP worked.

(2) Are there relationships between student educational performance

and noncognitive variables (such as attitude to school work) that are dependent on his race and sex?

The data include a large number of noncognitive variables--measures of a student's attitude to school, teachers, and fellow students, and his feelings of self-esteem. In Section IV these noncognitive variables are related to achievement levels for the four race-sex groups. The strength and direction of these relationships was found to vary across the four groups, and in some instances the variability promised to explain the higher performance levels of black males. Specifically, black students who felt race was not linked to IQ or who felt blacks were "smart" tended to do better on the achievement test. However, other parts of the analysis could not link either ESAP funding or specific educational programs to these noncognitive variables. Therefore, while this part of the analysis is suggestive of reasons for the performance levels of different race and sex groups, the clues could not be connected to a larger picture that included ESAP.

(3) What did ESAP buy?

Consistent with the diverse activities authorized by ESAP regulations, schools were found to have used program money for all kinds of programs and services. Analysis of this information in Section V points to several problems in specifying the content of the program in action. These include the vague terms used to define educational activities, the mixture of programs that were funded, and the lack of correspondence between different accounts of how the money was used. The most important conclusion, however, is that it is not possible to define a distinctive program in action. The activities that ESAP funded look much like the kinds of things that go on in the control schools. Though it would be possible, in principle, to differentiate experimental and control schools in terms of the extent or size of the funding, data were not available for this purpose. Because of this, the remainder of the analysis is a general investigation of relationships between educational programs and student outcomes; this investigation cannot be tied to ESAP per se.

(4) Can teacher and student noncognitive variables be related to particular educational programs?

Programs and activities are identified in Section V, and these

dimensions of the school environment are related to selected measures of attitude and behavior for teachers and students in Section VI. Generally, specific programs do not have a substantial or consistent effect on teachers or students, but there are two important exceptions. Both teachers and students tend to have more positive or favorable attitudes and self-reported behavior when their schools have one of two kinds of programs: inservice training that emphasizes race relations, or programs in intergroup relations for students. The effects are mainly concentrated among white students, however. Thus, one of the main objectives of ESAP, the improvement of race relations in schools, is attained to a greater extent in schools that run the kinds of programs that might be expected to achieve those ends. The size of these effects should not be exaggerated, but the clues afforded by these results are encouraging.

(5) Can educational achievement levels be linked to noncognitive variables and these in turn to educational programs?

If certain educational programs can be associated with differences in student attitudes, can both be linked to educational achievement? The question is whether the degree to which achievement is related to these noncognitive characteristics of students can be attributed to the presence of programs involving intergroup relations. There appears to be no connection of this kind, for although achievement levels are associated with noncognitive differences among students, the existence of those relationships cannot be attributed to particular educational strategies (Section VII). This section also contains an analysis of districts where students in the experimental school scored much higher (or lower) than those in the control school. This shows the "successful" ESAP schools tended to be smaller and situated in fairly rural areas. The way ESAP was implemented in the district could account for this difference.

ACKNOWLEDGMENTS

The author would like to acknowledge the help of John Pincus and Robert Crain in reviewing the work and providing helpful suggestions for the execution of the analysis. Robert York, of the U.S. Office of Education, was consistently fair in monitoring the research. Two members of the National Advisory Council on Equality of Educational Opportunity, Frederick Mosteller and Jacquelyne Jackson, followed the work closely and contributed perceptive comments. Finally, Carol Frost and Sue Pollich were indispensable and efficient in handling data preparation chores.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS ix

Section

I. INTRODUCTION 1
The Data Set 4
Integration and Disturbance 6

II. DID ESAP ALTER THE RACIAL CLIMATE OF THE SCHOOL? 10
Relations Between Groups 11
The Principals' Attitudes 14
The Teachers' Attitudes 15
Students' Attitudes 16

III. THE INTERACTIVE EFFECT OF ESAP ON NONCOGNITIVE OUTCOMES 18

IV. RELATIONSHIPS BETWEEN STUDENT NONCOGNITIVE VARIABLES AND ACHIEVEMENT SCORES 28

V. WHAT DID ESAP MONEY BUY?
The ESAP Director Survey 37
The Principal Survey 40
The Teacher Survey 40
Program Types 41
A Note on Program Delivery 43
Why Did Schools Choose Their Particular Program? 45

VI. RELATIONSHIP BETWEEN PROGRAM TYPE AND PROGRAM OUTCOMES: TEACHER AND STUDENT NONCOGNITIVE VARIABLES 48
Teacher Outcomes 48
Student Outcomes 56

VII. THE RELATION OF STUDENT ACHIEVEMENT TO ESAP, EDUCATIONAL PROGRAMS, AND STUDENT ATTITUDES 69
Analysis of Outliers 69
Educational Programs, Noncognitive Outcomes, and Student Achievement 76

VIII. REVIEW 84

BIBLIOGRAPHY 86

I. INTRODUCTION

The Emergency School Assistance Program (ESAP) provided financial aid to schools that were in the process of becoming racially desegregated or had recently made this transition. The program's objectives were broad ones:

> The purpose of the emergency assistance to be made available...is to meet special needs incident to the elimination of racial segregation and discrimination among students and faculty in elementary and secondary schools by contributing to the costs of new or expanded activities...designed to achieve successful desegregation and the elimination of all forms of discrimination in the schools on the basis of race, color, religion or national origin. (Part 181.2, Title 45, Code of Federal Regulations.)

This was to be achieved by financial assistance, which, for the schools studied here, amounted to about $10,000 per school. The funds were used for all kinds of things, in keeping with the broad range of activities authorized by the bill: teacher training, remedial education, curricular innovations, purchase of teaching equipment, community programs, remodeling buildings, and hiring professionals such as guidance counselors or teacher aides. The impression is that the program's aims lacked specificity and the means of attaining them were correspondingly diffuse. But though the program had a good deal of flexibility in action, legislators perceived a specific, short-term need faced by desegregating schools, as evidenced by congressional hearings. Funds were to help with a temporary difficulty.

To what extent was this notion of a disturbance or crisis justified? Preliminary to the main analyses reported here, the data collected for the evaluation of ESAP were used to see if schools did experience a short-term crisis associated with desegregation. The analysis, reported at the end of this Introduction, deals with the way teachers and students perceive the quality of race relations in the school. There is a small but discernible tendency for the quality of the social environment to decline and then recover as a school makes the transition to desegregation. But it also seems the onset of this disturbance is delayed,

suggesting that schools take time to catch up to the problems associated with changes in racial composition. This analysis, then, gives qualified support for the common sense notion that schools experience temporary disturbance. In retrospect the assumption of ESAP, that schools were, in fact, disrupted by desegregation is modestly supported by these data.

A major investigation of the program's effects was mounted by the National Opinion Research Corporation (NORC, Vols. 1 & 2, 1973), and the data collected in that evaluation are the basic material for this report. The evaluation sought to explain how ESAP money had been spent and what effect it had had on schools, teachers, and students. The most important finding was that students in ESAP-funded schools scored higher on an achievement test than similar students in comparison schools. The program had an effect on educational performance. However, this effect was found for one group of students only: black male 10th graders. It did not have the same effect on black females or whites of either sex, nor did it have this effect on the sample of 5th graders.

The National Advisory Council on Equality of Educational Opportunity (NACEEO) asked the Office of Education to pursue this finding further, and the contract for this work was awarded to Rand. The Advisory Council's interest in the reanalysis was prompted by the need for information that might help guide ESAA, the successor to ESAP, which it was the Council's task to monitor. The Council did not expect definitive answers. For one thing, the program effect was fairly small, so any associated relationships were likely to be weak. For another, there were assorted problems of measurement that imposed limits on any analysis, notably that seeking to identify the content of the Emergency School Assistance Program. Therefore the secondary analyses are carried out as a free-ranging and open-ended inquiry; the aim is to search for any useful information about the way ESAP worked, and more particularly how it succeeded. It is an exploratory, not a confirmatory, inquiry.

Rand's secondary analysis had two parts. The first was a reinvestigation of the program effect on the achievement scores of black male 10th grade students. The results of this stage of the work show that these students did score above similar students in comparison schools, though the new estimates are somewhat smaller than those reported in the

original evaluation. It appears black male students in the ESAP-funded schools score, on average, one test point above those in the comparison schools. The second stage of the Rand work, reported here, sought to account for the special program effect on this subset of students. The original data are reanalyzed here with the aim of finding out why the program was associated with better educational performance for one group of students, but not for others.

The first part of the analysis (Sections II and III) pursues a possibility raised in the original evaluation that ESAP altered the racial climate of the schools, and this in turn helped some students perform better on the achievement test. Racial climate is measured in a number of ways, including statements of attitude on the part of teachers, students, and school principals; their self-reported behavior, perceptions of others, and observations about the quality of race relations. These are used to investigate two propositions: (1) that ESAP schools had more favorable racial climates, and (2) that that climate had a specialized effect on some subgroups of students and not others in the way they felt about school and how they got on there.

This is followed by examination of the relationship between student achievement levels and their attitudes, self-concept, perceptions of their teachers, and other measures that are called "noncognitive." The purpose is to search for differential relationships for the four race and sex groups to see if achievement is influenced by these noncognitive variables in ways that could explain the performance of black male students (Section IV).

The next step was to make a more detailed analysis of the program in action. ESAP money was used in all kinds of ways; because of this it is natural to ask if some uses, activities, or programs were more effective than others. Accordingly, in Section V, I look at the different uses made of ESAP money and discuss the difficulties of identifying the school-level activities with certainty. Then, in Sections VI and VII, I investigate the effects of different program types on a range of outcomes, including teachers' attitudes toward desegregation and toward minority students and students' feelings about school and their educational performance.

THE DATA SET

In 1971-1972, the year of the evaluation, the program involved 452 school districts and varying numbers of schools within each of these districts. This secondary analysis is concerned with a subset of school districts in which ESAP was administered on a randomized basis. There were 103 districts included in this part of the program; 48 of them had a pair of high schools. In each of these 48 districts, school administrators selected a pair of high schools that had comparable populations of students. One school from each pair was then randomly chosen to receive ESAP funding, and the unfunded school became the control for comparison purposes. Much of the secondary analysis involves comparisons between the experimental, or treatment, schools and the control schools.

The 48 school-pairs studied here are the total set of experimental and control schools. A school-pair was eliminated from the previous analysis either because of conventional missing data problems or because there were no students of a given race or sex.

THE IMPLEMENTATION OF ESAP

A potential problem involving the experimental design is that it does not appear to have been perfectly executed. Some experimental schools did not receive the treatment while some control schools did receive it. Information was obtained from the ESAP director, who was responsible for ESAP at the local level and reported on the use of ESAP funds in each school. He was also asked whether the schools had received these funds in the first place. Two of the 48 experimental schools were found not to have obtained the program money, nor did they have any ESAP-funded activity such as inservice training or new personnel, equipment, or resources.

For the control schools, seven of the 48 were reported to have received ESAP money, and in six of these seven the ESAP director also reported at least one activity or program financed by ESAP. This breakdown is reported, for control schools only, on the following page.

To take the most stringent view, 19 control schools had some ESAP-funded activity. Since one of these is the school matched to one of the two experimental schools that violated the experimental design, there are 20 school-pairs out of 48 that do not comply with the experiment protocol.

TALLY OF ESAP-FUNDED PROGRAMS.
(control schools only)

ESAP Funding	None	At Least One	Total
None recorded	28	13	41
Some funding	1	6	7
Total	29	19	48

A case can be made that these data are ambiguous. First, it is possible there was some confusion about which schools were being asked about by the interviewer. This is likely because a good deal of school information was obtained by leaving questionnaire forms for later completion, often by the director's assistant. Second, some ESAP-funded activities were available on a district-wide basis. For example, an inservice training scheme run from a central location would draw teachers from a number of different schools, perhaps including the control schools. This would explain the 13 control schools that did not receive ESAP funding, yet participated in ESAP-funded activities. To check this possibility I looked to see if the programs recorded for these 13 schools were the kinds that might logically be offered on a district-wide basis. For instance, did these schools tend to have programs like inservice training or community projects? Or were they just as likely to have purchased audio-visual equipment or to have installed tutoring programs, the kinds of things that are more logically delivered at the school level? There is a problem of making a neat distinction between district-level and school-level activities, but to some extent this can be done, and subsequent analysis showed no tendency for the 13 control schools to have taken part in the first rather than the second kind of program.

In view of the importance of the experimental design and the possible violation of that design, key analyses that will be reported here were replicated eliminating the 20 school pairs where the experimental design may have been mismanaged. The results of these analyses indicated there was no bias involved in including all the schools, so analyses reported here are based on the maximum data set of 48 school-pairs. The decision

to use all possible schools was based on both the empirical finding that the schools violating the design do not bias the results and on the uncertainty about the quality of the information reported by the ESAP director.

INTEGRATION AND DISTURBANCE

It was suggested earlier that integration goes with a period of temporary disturbance in the schools. A comparison of schools that had been integrated for different lengths of time made it possible to test this supposition. The principals in experimental and control schools were asked when desegregation had had the "greatest effect on change of racial composition of the student body." Schools where this change had been greatest in the current school year (1971-1972) were separated from those where it had taken place in the two previous years (1969-1970 and 1970-1971) and these from schools where desegregation had taken place earlier yet, making three levels of "recency of desegregation." Two possibilities were investigated. First, it might be that most recently desegregated schools would be worst off, followed by schools desegregated next longest, followed by the schools desegregated for the longest period. This is based on the idea of a steady recovery from the immediate effect of racial mixing. The second possibility is a little more complicated, involving the notion of a delayed reaction. Here, the effect of desegregation takes a year or so to become evident, followed by recovery. If this is how things work, schools desegregated for one to two years should be worst off, and schools that have been desegregated for shorter or longer periods would be relatively better off. The two possibilities can be diagrammed as shown at the top of p. 7.

Measures of the school environment are derived from principal, teacher, and student questionnaires and cover, among other things, their perception of the relations between races, of teachers' attitudes to desegregation, students' attitudes to their schoolwork, and students' attitudes toward students of the opposite race. A total of 67 measures were used. Simple comparisons were made among the schools desegregated for different lengths of time; these comparisons showed very few

Variable Indicating Quality of School Environment	First Hypothesis			Second Hypothesis		
Favorable score			x			x
		x		x		
Unfavorable score	x				x	
Year school desegregated	1971	1969/ 1970	1968 or before	1971	1969/ 1970	1968 or before

statistically significant results, but the trend of the mean scores for schools is consistent with the second hypothesis about the effect of desegregation. Schools appear to experience a delayed onset of the effects of desegregation, which became evident after one or two years, followed by recovery. To repeat, few of these differences were found to be statistically significant; the justification for the conclusion is based on the similarity of results for a substantial number of variables, not on the size of the differences involved. Therefore, the results of the analysis are not reported, except that descriptions of variables conforming to the second pattern are listed in Table 1, which also shows the total number of variables used in the analyses.

Although the results are not strong, they are consequential because they confirm the common sense observation of the association between integration and disturbance, so it is important to test alternative explanations of the findings. One contending explanation is that desegregation had taken place earlier in rural areas, and urban districts followed later. Schools desegregated the longest were therefore more likely to be rural and to experience fewer tensions and difficulties as a result of their rurality rather than because of the onset of desegregation. Conversely, schools that had been desegregated a short time were more likely to be urban schools so that the measures of the school environment would be picking up the influence of the urban

environment. Therefore the analyses were repeated controlling for the percent of the population in the county living in communities over 2,500. The introduction of this control did not alter the original conclusion; the pattern of results could not be explained in terms of the different locations of schools.

Table 1

SUMMARY OF ANALYSES RELATING THE LENGTH OF TIME A SCHOOL HAS BEEN DESEGREGATED TO THE RACIAL CLIMATE OF THE SCHOOL

Teacher data (24 variables used)

Reports a greater amount of fighting than before desegregation

Reports white students becoming less prejudiced

Describes contact between whites and blacks as friendly

Thinks blacks would be better off in integrated schools, not segregated ones

Thinks whites would be better off in integrated schools, not segregated ones

Feels it is proper to let students know how they feel about race relations

Tends to have class discussions about race

Says teachers tend to be fair to black students

Reports students tend to be favorable to desegregation

Reports white teachers in school tend to be favorable to desegregation

Reports black teachers in school tend to be favorable to desegregation

Thinks the civil rights movement has done good rather than harm

Thinks that blacks and whites should be allowed to intermarry

Says that student dances have not been eliminated because of possible racial problems

Says that student elections have not been eliminated because of possible racial problems

Principal data (17 variables compared)

Thinks that whites would be better off in integrated, not segregated schools

Thinks the civil rights movement has done more good than harm

Reports black teachers' attitudes to desegregation being favorable in his school

Table 1--continued

Student data (26 variables compared)
Says he feels he belongs in the school
Says most of his teachers like the idea of blacks and whites going to the same school
Tends to disagree that "white students complained favoritism is being shown to black students"
Tends to disagree that "tensions have made it hard for everyone"
Reports few serious problems in the way things are working out between black and white students in the school
Says that among the three students he talks to most in school, at least one is of the opposite race
Had called a student of the opposite race on the phone
Had helped a student of the opposite race with homework
Had asked a student of the opposite race for help with own homework
Would prefer to be in a racially mixed school
Would like to have more friends of the opposite race
Says he is usually comfortable around students of the opposite race
Disagrees that students of the opposite race are dumb
Thinks that race is not linked to smartness
Reports that he tends to be happy "these days"

Note: Schools were divided into those where the effect of desegregation was reported greatest during the present academic year (1971), during the two years previous to that (1970, 1969) and earlier (1968 or before). Mean scores on measures of school climate were computed for each of these three groups. The total number of variables analyzed is shown at the top of each section. The body of the table contains descriptions of the variables that show the score for the second group (desegregated 1970, 1969) indicates the school to have a less favorable racial climate than either the first group (desegregated during the school year 1971) or the third group (desegregated before 1969).

II. DID ESAP ALTER THE RACIAL CLIMATE OF THE SCHOOL?

ESAP was intended to assist schools in the process of integration. Integration involves more than racial mixing; it also refers to the quality of the relationships between racial groups, the degree of harmony in the school, the level of hostility, students' attitudes toward one another, and the expectations they have of themselves and of the other group. These intangible qualities are sometimes talked about collectively as the "racial climate" of the school. Since ESAP was expected to help schools integrate, it is reasonable to ask if it had an effect on climate; did schools that received program funding have more favorable environments than the unfunded control schools?

The second reason for raising this question was suggested by the authors of the original evaluation. Their summary contains the speculation that black male 10th graders scored higher in the experimental schools "because of improvements in the schools' racial climate affecting the motivation of these students" (NORC, 1973, p. iv). As the report made clear, the evidence for this suggestion was incomplete; in fact, the analysis of ESAP's influence on student attitudes toward integration showed no effect of the kind they suggested. But only one measure was used (NORC, pp. 55-58). Much the greater part of the data on student, teacher, and principal attitudes and behavior was not used to answer the question: Are there any differences between experimental and control schools that could establish the link between the program funding and the higher performance of black males?

In this section, I make straightforward unadjusted comparisons between ESAP and control schools. Since the purpose is to look at overall program effect, the students are not divided here by race and sex. The most important data are obtained from principal, teacher, and student surveys. This information is supplemented by data obtained from the school counselor and from an observer (the person from the evaluation team who administered the school surveys).

An interview was administered personally to each principal in each experimental and control school. Since there are 48 experimental and

48 control schools, the results of this survey are presented as "number of principals responding."

Ten 10th-grade teachers were sampled from each school and given a questionnaire to answer on their own. Since there are unequal numbers of teachers in experimental (495) and control (489) schools, comparisons of these data are presented as percentages (for example, "x percent of experimental school teachers said...").

Student questionnaires were given out in class by the evaluation staff, and every effort was made to have students complete it when the teacher was out of the room because some questions deal with students' perceptions of their teachers. Here too, there are slightly different numbers of students (2665 experimental, 2621 control) so results are presented as percentages.

Data from each source are classified into four broad areas: relations between different groups with an emphasis on their behavior, attitudes of the principal to integration and to minority groups, attitudes of the teachers to integration and toward minorities, and attitudes of students to one another and their feelings about themselves.

RELATIONS BETWEEN GROUPS

Possibly the simplest way of assessing the quality of relations between groups is to ask an outsider for his immediate reaction. Such an outsider is the member of the evaluation team who visited the school to administer various questionnaires. While there, he recorded his sense of the "general atmosphere." Most schools, of both kinds, were found to be relaxed rather than tense (30 experimental and 31 control); there is no important difference between the experimental and control schools. This assessment is backed up by the teachers who usually indicated that problems associated with desegregation were "minor" (80.9 percent experimental and 74.8 percent control). However, students offer a different account; they were much more likely to say tensions had "made it hard for everyone" (47.7 percent of students in the experimental schools answered this way and 48.6 percent of the students in control schools). The difference between students' and teachers' reports may well have to do with the wording of the questions--tensions

are not the same as problems--but still, there is a good chance different groups perceive the same situation in different ways.

General questions have the disadvantage that people can read their own meanings into them. This is harder when questions are specific. For example, the principals were asked how many days the school was closed during the year because of racial tension. Only two principals (both of experimental schools) said this had happened. Asked about the previous year, rather more said the school had closed (3 experimental and 10 control). The difference between experimental and control schools here cannot be attributed to the recency of desegregation; the experimental schools tend to have been desegregated longer, but not by much. Teachers were also asked about the suspension of normal school activities because of racial problems. They reported on the elimination of dances (34.5 percent experimental and 31.4 control) and of student elections (7.7 percent experimental and 6.4 percent control). More dramatic was information about whether teachers were attacked by students (2 experimental schools and 5 control), and about the amount of fighting that went on among students (in 13 experimental and 12 control schools, fights led to the need for medical treatment). The level of overt hostility seems just about equal in both types of schools; this is supported by the students' own account, which may bring out (or possibly exaggerate) details that adults would rather forget (16.7 percent of students said they had been involved in a fight during the year in experimental schools and 15.7 percent in controls). Students also reported quite a lot of interracial fighting, though they tend to report this from the point of view of their own "side." In experimental schools, 36.7 percent of the students said blacks had attacked whites, and in control schools, 36.7 percent said the same.

Harder to interpret than the levels of violence in the school are the statistics on discipline and control. Yet this is an important area in which the relationship between adults and students is expressed and indirectly tells us about the level of tension in the school. Expulsion is the most serious action schools take to deal with student problems. It is not used often; in fact, it was used so little in these schools that it is a poor index for the purpose of this analysis. (Of 10th grade

students in experimental schools 1 percent had been expelled that year and 0.4 percent in controls--black students only). Dropout rates, which may reflect an informal expulsion, are also very similar for the two kinds of school (6.3 percent of 10th graders dropped out of experimental schools and 7.2 percent in controls). Absenteeism is another credible measure of the degree to which students find school an enjoyable or rewarding experience; and here, too, experimental and control schools look almost identical (two experimental school principals and three control said absenteeism had increased for black students during the previous year). The students' own report supplies the same message. In experimental schools, 54.6 percent said they had "stayed away from school just because they didn't want to come," and 54.3 percent in control schools. They were also similar when it comes to being sent to the school office for breaking rules (46.1 percent of students in experimental schools had been at least once during the year and 47.9 percent in control schools).

Finally, there are various hints about the extent to which black and white students mix and form friendships. Generally, students keep themselves apart. Most often they said the three people they talked to most in school were of the same race as themselves (78.4 percent in experimental schools and 80.4 percent in controls), and they were unlikely to have made social contact outside school (25.8 percent in experimental schools had phoned a student of the other race and 28.3 percent in control schools). However, the informal segregation was far from complete. About one third of students said they had asked for help with homework from someone of the other race (34.5 percent experimental and 32.2 percent control). Teachers support this account; asked to characterize the quality of contact between blacks and whites they most usually said there were a "few" interracial friendships (52.6 percent of experimental school teachers and 46.3 percent control). Added to that, the outside observer usually said he had seen some interracial groups around the school during recess or after school (23 of the experimental schools and 22 of the controls). The most important point in all this is not the level of interaction but the difference between experimental and control schools. While some comparisons show students in experimental schools are better off than those in controls (11 out of 21 comparisons favor experimental

schools), it is also true that control school students are better off than experimental school students in several other comparisons (10 out of 21).

The only tenable summary of balanced results like these is that the quality of relations is systematically neither better nor worse in the experimental schools. In no single comparison is there a dramatic difference between the two kinds of schools. Yet if ESAP did not change the quality of relations, perhaps it altered the way principals and teachers felt about the issues of race and integration.

THE PRINCIPALS' ATTITUDES

The principal is in a unique position to define what constitutes acceptable behavior and to change people's ideas and attitudes. This is not to say he does exert such powerful influence, but he might. Therefore his attitudes toward desegregation and minority group students are important. One indicator of his enthusiasm for desegregation was obtained directly by asking how well he liked it; most said they did (33 of the experimental school principals and 38 of the controls). But it seems this feeling is not perceived by others, since only about a third of the teachers said they thought the principal liked desegregation (33.4 percent experimental and 31.6 percent controls); students were even more doubtful (16.5 percent said their principal liked it in experimental schools and 16.6 percent in control schools). Some of the difference is probably attributable to the way the questionnaires were filled in. Principals did it face-to-face with an interviewer who was identified as an ESAP evaluator; teachers and students filled theirs in anonymously so they may have been more candid.

Principals were also asked whether students would be better off in schools that were racially segregated. This reveals equally positive inclinations; four of the principals of experimental schools said blacks would be better off and five of the control school principals. Not surprisingly, they answer differently when asked about whites; but still, less than half say white students would be better off in segregated schools. (14 for experimental school principals and 15 for controls). Again, the results are mixed and do not support the contention that there were clear differences between experimental and control schools.

THE TEACHERS' ATTITUDES

Principals of the two classes of schools are evidently quite alike. But perhaps they are figureheads, without direct influence on the day-to-day life of the students. It is much harder to argue that the classroom teachers are similarly remote, so their attitudes may be even more significant. However, the teachers in experimental and control schools, like the principals, appear to be much alike. For instance, asked whether blacks would be better off in all black schools, 29.7 percent of experimental teachers felt they would, and 25.8 percent of control school teachers, which suggests control school teachers are somewhat better disposed to integration than experimental school teachers. Asked the same question about whites, the difference is reversed (41.5 percent of experimental school teachers say whites would be better off and 41.4 percent of control school teachers). Other groups also reported on what they felt teachers' attitudes to be. For instance, principals tended to say their teachers liked desegregation (28 of experimental school principals said white teachers liked it and 32 of the control school principals said the same). The students were less likely to think their teachers liked desegregation (10.4 percent of experimental school students said so and 9.8 percent of the control school students).

Teachers were asked about black students' academic potential, something to which great importance has been attached by the research community. They were asked to judge what proportion of blacks had the potential to attend the largest university in their state. Very few said that even half the black students in their school had this future (7.0 percent experimental school teachers and 6.3 percent in control schools). Teachers were also asked to assess the proportion of blacks who were "discipline problems," and here too there is little difference between experimental and control schools; about half the teachers in both types of schools said at least one in 20 of their black students were discipline problems (53.4 percent experimental and 55.1 percent control).

Again, although there are individual comparisons that show the teachers in the two kinds of schools were different from one another, the differences tend to cancel out. Taken together there is little indication that teachers in ESAP-funded schools had more, or less,

favorable attitudes toward integration or to their students than teachers in control schools.

STUDENTS' ATTITUDES

The students were not asked about desegregation directly; instead they were asked whether they would prefer to attend a racially mixed school. In experimental schools, 49.8 percent said they would and in control schools, 47.6 percent: rather less. Comparable differences show up on other variables too. For example, asked if they would like more friends of the other race, 58.3 percent of the experimental school students said yes and 55.6 percent of the control school students. Very slightly fewer students in experimental schools say they are never uncomfortable around students of the opposite race (32.2 percent and 32.7 percent). However, they are just about equal when it came to saying if students of the other race are dumb (29.4 percent and 29.2 percent) and if race has anything to do with smartness (81.6 percent in experimental schools said it does not and 81.6 percent in control schools).

A different angle on the quality of the school atmosphere can be obtained from students' answers to questions about how much they like school and how they feel they are treated there. This brings the discussion back to where it began. One index of student perception is the view they have about the fairness of the rules in the school. Only half felt the rules were fair (53.6 percent in experimental schools and 52.2 percent in controls), and a lot thought they got punished for no good reason (44.1 percent and 45.6 percent). More generally, the 10th graders were not especially enthusiastic about going to school (43.6 percent said they were pleased to go to school in the experimental schools and 44.6 percent in controls). And when they got there, only a few felt they belonged in the school in some way (21.5 percent and 24.2 percent). Most of these differences favor the experimental schools and indicate the program improved things for students. Students in experimental schools were also more optimistic about thinking they could complete college (65.9 percent and 64.8 percent), although they ranked themselves, in comparison with their classmates, in exactly the same manner (28.1 percent say they are above average in experimental schools and 28.1 percent in controls).

In conclusion, the best evidence that ESAP made schools more harmonious places comes from the analysis of student attitude variables. Students in the experimental schools were slightly more likely to prefer desegregated schools, to prefer more friends of the other race, and to feel comfortable around students of the other race. But even here there is contradictory evidence, showing control school students were happier about going to school, for example. In the other three areas explored here, experimental and control schools seem very much alike. To the extent that comparisons from different sources can be fitted together, the most sensible conclusion is that ESAP schools were not clearly different from control schools, they did not have markedly better relations between races, their principals and teachers were not better disposed to integration or to minorities, and their students were also very similar.

III. THE INTERACTIVE EFFECT OF ESAP ON NONCOGNITIVE OUTCOMES

The previous section looked at the effect of ESAP on teachers, principals, and students without subdividing them in any way; but it is both possible and likely that different groups responded in different ways to ESAP. Since the analyses of achievement scores had indicated higher achievement levels for black male students in experimental schools, the question was raised whether ESAP had differential effects on other outcome variables for the students. For instance, suppose black male students in experimental schools had greater confidence in their own ability than black male students in control schools, and suppose that this difference is not found for groups of different race and sex. Then one link will be suggested in the chain of events between ESAP funding and educational achievement. Of course, it would then be necessary to show that students' self-confidence was related to their educational performance, but at least the first step would show a differential program effect on a noncognitive quantity.

The analytic approach is simple but the interpretation of results fairly complicated. The sample of students is divided into four subgroups: black males (N=906), black females (N=1037), white males (N=1538), and white females (N=1611). For any given outcome--say, student self-confidence--the difference between the means for experimental and control schools is tested within each of the four subgroups. In Table 2 each line shows the mean scores and the significance level of the difference symbolized in the conventional manner. Most of the variables are dichotomous, so the numbers to the right of the decimal point can be interpreted as percentages. For example, taking the first comparisons, the value of 1.616 means 61.6 percent of the black males in the experimental schools said they were in a school club or team. In the control schools, 65.8 percent of the black males said the same. But not all variables are of this kind, and care should be taken to check the coding of each variable in the left hand margin.

Before I turn to the results themselves, consider the kinds of findings that are pertinent to the inquiry.

Table 2

MEANS OF EXPERIMENTAL AND CONTROL SCHOOL STUDENTS, ON NONCOGNITIVE OUTCOMES, BY RACE AND SEX

	Black Male		Black Female		White Male		White Female	
	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont
(1) *Level of interaction in school*								
If student is a member of a school club or sports team (1=no/2=yes)	1.616	1.658	1.553	1.559	1.614	1.625	1.668	1.648
If student had talked with a school counselor during the year (1=no/2=yes)	1.525	1.537	1.576	1.550	1.580	1.530*	1.615	1.559*
Student talked with teacher about his interests outside of school (1=no/2=yes)	1.546	1.563	1.505	1.530	1.636	1.599	1.624	1.652
Is there an adult the student could go to if in trouble or upset? (1=no/2=yes)	1.678	1.738*	1.748	1.708	1.660	1.668	1.718	1.728
(2) *Students' feelings about school*								
Student feels he belongs in the school (1=disagree/2=agree)	1.734	1.745	1.753	1.679*	1.792	1.768	1.826	1.809
Student says he feels happy (Range: 1-3, high score=happy)	1.783	1.828	1.788	1.738	1.943	1.931	2.067	2.049
Student says he is glad to go to school in the morning (1=no/2=yes)	1.509	1.518	1.572	1.496*	1.315	1.313	1.437	1.502**
Student says he hates school (1=yes/2=no)	1.720	1.704	1.712	1.687	1.569	1.606	1.676	1.712

Table 2--continued

	Black Male		Black Female		White Male		White Female	
	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont
Amount of time spent on homework (Range: 1-5. 5=most time)	3.159	3.021	3.688	3.569	2.521	2.598	3.382	3.398
Found either schoolwork or homework interesting during the week (1=no/2=yes)	1.591	1.635	1.745	1.680*	1.390	1.386	1.540	1.531
(3) *Getting into trouble*								
Says he had been in a fight during the school year (1=yes/2=no)	1.779	1.799	1.908	1.905	1.730	1.748	1.926	1.930
Had been sent to the office for breaking the rules during the year (Range: 1-3)	2.111	2.031	2.442	2.418	2.112	2.069	2.586	2.617
Says he spent days away from school just because he didn't want to come (Range: 1-5. 5=never...1=16 or more days)	4.475	4.473	4.641	4.589	4.045	4.029	4.293	4.291
Student says he is blamed for things that are not his fault (1=yes/2=no)	1.575	1.539	1.674	1.631	1.534	1.551	1.743	1.756
Student says rules in the school are fair (1=disagrees/2=agrees)	1.430	1.461	1.453	1.366**	1.558	1.520	1.625	1.667
Says that when punished it is for no good reason (1=disagrees/2=agrees)	1.524	1.545	1.432	1.460	1.500	1.461	1.341	1.399*
(4) *Attitudes to students of the other race*								
If student could choose, he would go to a racially mixed school (1=no/2=yes)	1.655	1.659	1.624	1.636	1.350	1.273**	1.468	1.475

Table 2--continued

	Black Male		Black Female		White Male		White Female	
	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont
Would like more friends of the other race (1=no/2=yes)	1.765	1.746	1.798	1.750	1.434	1.383*	1.489	1.494
Feels uncomfortable with students of the other race (Range: 1=4. 4=no)	3.024	2.974	3.227	3.258	2.629	2.597	2.864	2.814
Describes students of the other race as dumb (1=yes/2=no)	1.672	1.706	1.746	1.705	1.543	1.505	1.680	1.702
Describes students of own racial group as dumb (1=yes/2=no)	1.779	1.763	1.812	1.822	1.879	1.902	1.898	1.907
Says whites are smarter than blacks (1=yes/0=other response)	0.021	0.011	0.006	0.011	0.343	0.328	0.152	0.160
Says blacks are smarter than whites (1=yes/0=other response)	0.056	0.053	0.036	0.043	0.004	0.004	0.000	0.003
Says color doesn't have anything to do with smartness (1=agree/0=other response)	0.904	0.916	0.957	0.934	0.648	0.666	0.845	0.836
(5) *Interaction with students of the other race*								
Says three students talks to most are of the same race (1=yes/2=not all same)	1.348	1.339	1.281	1.253	1.158	1.137	1.145	1.131
Had called a student of the other race on the phone (1=no/2=yes)	1.375	1.392	1.287	1.352*	1.197	1.198	1.223	1.248

Table 2--continued

	Black Male		Black Female		White Male		White Female	
	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont
Had helped a student of the other race with homework (1=no/2=yes)	1.521	1.510	1.546	1.559	1.524	1.451**	1.656	1.633
Had asked for help with homework from student of other race (1=no/2=yes)	1.378	1.357	1.439	1.422	1.240	1.194*	1.358	1.350
(6) *Perception of teachers' attitudes to desegregation and minorities*								
Teachers seen as liking blacks and whites going to the same school (1=yes/0=other)	0.167	0.151	0.160	0.108*	0.077	0.081	0.056	0.078
Teachers seen as disliking blacks and whites at same school (1=yes/0=other)	0.167	0.146	0.154	0.177	0.169	0.202	0.154	0.145
Principal seen as liking blacks and whites at the same school (1=yes/0=other)	0.241	0.205	0.233	0.189	0.138	0.157	0.101	0.140*
Principal seen as disliking blacks and whites at same school (1=yes/0=other)	0.141	0.162	0.124	0.166	0.142	0.149	0.125	0.116
Teachers seen as being unfair to whites (1=yes/2=no)	1.855	1.816	1.844	1.852	1.598	1.579	1.585	1.637*
Teachers seen as being unfair to blacks (1=yes/2=no)	1.357	1.327	1.281	1.269	1.867	1.873	1.821	1.827
White students seen as complaining of favoritism to blacks (1=yes/2=no)	1.691	1.729	1.692	1.728	1.488	1.489	1.428	1.506**

Table 2--continued

	Black Male		Black Female		White Male		White Female	
	Exp	Cont	Exp	Cont	Exp	Cont	Exp	Cont
Black students seen as complaining of favoritism to whites (1=yes/2=no)	1.503	1.522	1.442	1.416	1.482	1.490	1.325	1.377*
(7) *Students' self-esteem*								
Student's rating of his ability compared with others in his class (Range: 1-5. 5=high)	3.079	3.188*	3.107	3.151	3.290	3.316	3.403	3.314**
Student's assessment of ability to complete college (Range: 1-5. 5=could complete)	3.606	3.694	3.642	3.650	3.641	3.619	3.731	3.647
Student says he will go to college (1=no/2=yes)	1.471	1.510	1.552	1.555	1.529	1.505	1.532	1.541
Student says he does not have much to be proud of (1=agree/2=disagree)	1.571	1.619	1.622	1.575	1.738	1.724	1.749	1.782
N based on	468	438	507	530	775	763	824	787

* = significant at the .05 level.
** = significant at the .01 level.
*** = significant at the .001 level.

(1) The first class of finding would be one where black males in experimental schools have a more "favorable" score than those in control schools, but the difference does not show up for the other race-sex groups. This might mean having higher self-esteem, as in the example given above, or more frequent contact with teachers, or a more positive view about students of the opposite race.

(2) Black males in experimental schools are no different from those in control schools with respect to the outcome variable. But at the same time, some or all of the other race-sex groups are worse off in the experimental schools than in controls. Suppose black males have the same level of self-confidence whether they are in experimental or control schools. Further, imagine students in the other race-sex groups had lower levels of self-confidence if they were in the experimental schools. Then it can be argued that black males in experimental schools enjoy a relative advantage, and one that might account for their higher level of achievement.

(3) The third class of finding is more complicated still. Here black males are worse off on some noncognitive outcome if they are in experimental schools, but the outcome measure turns out to be negatively related to achievement. For example, black males may be less self-confident in experimental schools, while students of the other race-sex groups are more self-confident if they are in the experimental schools. Further, imagine the unlikely event that higher levels of self-confidence are associated with lower levels of achievement. Then this would indicate a possible explanation for the superior performance of black males in experimental schools.

The results of the analysis are presented in Table 2. The dependent, or outcome, variables are put into seven groups on the basis of the kind of attitude or behavior they measure. The first four variables deal with students' interaction with the school and, more particularly, the level and ease of their communication with the adults in the school. The supposition is that the more students are involved with their teachers or counselors, the more they are integrated with the school and the greater their chance of doing well.

The second group, student's feelings about the school, deal directly

with students' attitudes toward the school and toward school work. Some measures are readily interpreted; if a student says he hates school, it probably means he doesn't like being there, but others are more ambiguous. For example, the amount of time a student spends on homework may indicate his enthusiasm for school, but it must also reflect the teachers' demands, regardless of what the student feels. Further attitude measures deal with the amount of trouble students get into and the extent to which they find the rules fair.

Since ESAP was aimed at improving race relations, students' attitudes to those of a different race (fourth part) are obviously important, as well as their level of interaction with them (fifth part). These two classes of variables contain measures of attitude to other racial groups, the amount of self-reported contact, and an index of stereotyping.

Variables in the sixth group deal with students' perceptions of teachers' attitudes toward desegregation and minorities. Of course, this will reflect both the teachers' actual behavior and the way students respond to and interpret that behavior. Finally, the last part contains four measures of students' self-esteem or feeling of self-worth.

The results will be discussed in terms of the three kinds of findings outlined above. The first question is whether black males in experimental schools scored higher on these variables than similar students in control schools when an experimental-control difference was not found in the other race-sex groups. It turns out there are two statistically significant differences between experimental and control schools (out of 40 comparisons), but in both cases the control school students, not the experimental school students, score highest. This eliminates the most likely area for future inquiry, since the program does not appear to have any positive effect on the noncognitive outcomes for black males.

There are similarly straightforward comparisons for the other three groups. Black females: There are six variables for which the experimental-control difference is statistically significant, and for five of these the difference favors the experimental school. White

males: All of the five significant comparisons show experimental school students scoring in a more favorable direction than control students. White females: Here the situation is reversed, since six out of eight significant comparisons favor the control schools. The results for the white males and white females suggest a question for future study--that these changes had an indirect effect on black male achievement. For instance, it is imaginable that the program improved white male attitudes to integration in a way that made the school environment less threatening for black males, and this in turn made it easier for them to do their work and score better on tests. This inquiry will be reported in Section VII.

The next kind of finding is that black males in experimental schools score at the same level as those in control schools, while experimental school students in the other race-sex subgroups score at a lower level than the control school students. Of course, if all the remaining three subgroups showed this pattern, the result would have added significance. However, none of the variables analyzed in Table 2 showed this pattern. A more modest standard is to compare the results for black males with those for black females, which will expose the sex-related interaction, and then with results for white males, exposing the race-related interaction effect. One interaction pattern fits the first condition, none fit the second. The one that does show up is a variable measuring interracial contact. Students were asked if they had called a student of a different race on the phone. There is no difference between black males in experimental and control schools, but there is a significant difference between black females; the control school female black students report higher levels of interaction than experimental school female black students. If this variable is a determinant of achievement levels, the finding is of importance, since black male students in experimental schools then have a relative advantage over black females in these schools. The importance of this noncognitive variable will be explored in the next section.

The final class of finding is more complicated still, since it requires black males to be worse off, on the noncognitive variables used here, if they are in the experimental schools; and the differences

between experimental and control students for the other three race-sex groups are either insignificant, or positive in favor of experimental schools. This scenario also requires that the noncognitive variable in question is negatively correlated to achievement. Here, too, one variable follows the required pattern. Students were asked if there was an adult whom they could turn to if they were upset or in trouble. A significant difference is reported in favor of control school students for black males. Further, there is no important difference between experimental and control group students for the other three race-sex groups on this variable. This indicates a noncognitive variable of potential importance, and I shall look at the association between this variable and student achievement scores in the next section. In this case, the association has to be negative to support an explanation of the higher achievement scores of black males.

In conclusion, although there are no instances in which black male students in experimental schools score higher than control students (while other race-sex groups do not), a few noncognitive variables show a promising pattern of results for explaining the relative difference in performance between black males and the other subgroups. The relationship between noncognitive outcomes and achievement will be dealt with in the next section. A huge number of comparisons are made between experimental and control schools in Table 2; some of these are going to be statistically significant, and among those will be found some fitting the patterns that are interesting to this analysis. In other words, we might expect some positive results by chance, but the results are not at all consistent in the sense that they do not turn out the same way for variables that measure similar quantities.

IV. RELATIONSHIPS BETWEEN STUDENT NONCOGNITIVE VARIABLES AND ACHIEVEMENT SCORES

In search of the causes of higher black male achievement in ESAP schools, the analysis investigates the way noncognitive variables relate to achievement scores for the four race-sex groups. If the relationships vary across subgroups the finding might suggest ways of explaining ESAP's effect on black male achievement. For example, it might be that students' self-confidence was more strongly related to educational performance for black males than for other subgroups, perhaps because they are more vulnerable to disturbances in the school than other groups and profit most from any improvement in their self-confidence. To test this and other propositions, the sample is divided into the four sex and race groups. Within each of these groups, zero order correlations are produced for each of the noncognitive variables and student achievement levels. The general question is whether, for a given variable, these correlations show significant variation across the four race-sex groups.

This raises the question of how to judge "significant variation." For one thing, the reliability of the achievement test is not exactly the same for each group (it is .85 for black males, .84 for black females, .91 for white males and .90 for white females). So reliability will account for some of the observed variation. However, it turned out that the corrections had an inconsiderable effect on the pattern of correlations. The question then is how to judge if a set of four correlations is similar or different. There are two approaches: one based on tests of statistical significance, the other on the relative values of r^2. If the first standard is applied, it takes only small differences between two correlations for them to be judged unlike one another. For example, a correlation of .07 is statistically significant at $p = .05$, when $N = 906$ (as it does for black males), while a correlation of .10 is significant at $p = .01$, for the same number of cases. The difference in statistical significance levels indicates the two correlations are appreciably different. However, the predictive

efficiency of these correlations should also be considered. From this point of view, a correlation of .07, which implies the independent variable might account for one half of 1 percent of the variance in the dependent variable, does not seem much different from one of .10, which accounts for 1 percent of the variance. The issue then is what difference in r^2 is going to be considered a quantum for the purpose of the analysis? An arbitrary choice was made of one percentage point of the total variance--that is, two correlations are considered to be different from one another if one accounts for an additional 1 percent of the dependent variable variance. This means a correlation of .345 is judged different from one of .330; so is a correlation of .122 different from one of .070. These examples illustrate the point that as correlations get bigger, the required difference gets smaller. In the analysis, both of these standards are used. Statistical significance levels will be used to decide if correlations are different from zero; and r^2 will be relied on to judge if two correlations are substantially different from one another.

In reporting on Table 3, I first look at the leads suggested by the analysis in Section III. There, two noncognitive variables were identified as being relevant to the explanation of the higher achievement scores for black males. The first is a measure of the students' contact with others of a different race, defined by whether they called someone on the phone. Although no significant difference was found between black males in experimental and control schools, black females in control schools had higher levels of interracial contact than did those in experimental schools. If it can be found that this noncognitive measure is positively related to achievement, it will suggest that black males in experimental schools were better off than black females in experimental schools, at least in the sense that they had a relative advantage over them. The variable turns out to be positively correlated with achievement (Table 3, Part 5), and the correlations for three of the four groups, though all smaller than 0.1, are also statistically significant. This suggests one lead that could account for the relative superiority of black males in experimental schools. However, using the values of r^2, the correlation for black males (r = .03, r^2 = .0009) is

Table 3

ZERO ORDER CORRELATIONS BETWEEN MEASURES OF STUDENTS' NONCOGNITIVE VARIABLES AND THEIR ACHIEVEMENT SCORES.

Noncognitive Variables	Black Male	Black Female	White Male	White Female
(1) *Level of interaction in school*				
If student is a member of a school club or sports team (1=no/2=yes)	.07*	.15***	.15***	.24***
If student had talked with school counselor during year (1=no/2=yes)	.07**	.10**	.04	.04
Student talked with teacher about outside interests (1=no/2=yes)	-.01	.05	.05	.06*
Is there an adult student could go to in trouble (1=no/2=yes)	-.01	-.03	.03	-.00
(2) *Students' feeling about the school*				
Student feels he belongs in the school (1=disagree/2=agree)	.06	-.12***	.08**	.07**
Student says he feels happy (high score=happy)	-.04	.07*	.12***	.11***
Says he is glad to go to school in the morning (1=no/2=yes)	-.06	-.07*	.08**	-.02
Student says he hates school (1=yes/2=no)	.04	.09**	.20***	.12***
Amount of time spent on homework (Range:1-5. 5=most time)	.01	.11***	.06*	.02
Found either schoolwork or homework interesting during week (1=no/2=yes)	-.05	-.07*	.07*	.03*

Table 3--continued

Noncognitive Variables	Black Male	Black Female	White Male	White Female
(3) *Getting into trouble*				
Says he had been in a fight during the school year (1=yes/2=no)	.04	.08**	.16***	.07**
Had been sent to the office for breaking the rules during the year (Range: 1-3)	.10**	.13***	.15***	.09***
Says he spent days away from school just because he didn't want to come (Range:1-5. 5=never...1=16 or more days)	.08*	.02	.13***	.10***
Student says he is blamed for things that are not his fault (1=yes/2=no)	.13***	.21***	.17***	.05
Student says rules in the school are fair (1=disagrees/ 2=agrees)	.04	.04	.14***	.08**
Says that when punished it is for no good reason (1=disagrees/ 2=agrees)	-.04	-.06	-.01	-.05
(4) *Attitudes to students of the other race*				
If student could choose, he would go to a racially mixed school (1=no/2=yes)	.14***	.19***	.19***	.15***
Would like more friends of the other race (1=no/2=yes)	.08*	.08*	.19***	.08***
Feels uncomfortable with students of the other race (Range: 1-4. 4=no)	.15***	.12***	.13***	.03
Describes students of the other race as dumb (1=yes/2=no)	.15***	.17***	-.02	-.01

Table 3--continued

Noncognitive Variables	Black Male	Black Female	White Male	White Female
Describes students of own racial group as dumb (1=yes/2=no)	.12***	.13***	.12***	.08**
Says whites are smarter than blacks (1=yes/0=other response)	-.01	-.02	.01	.10***
Says blacks are smarter than whites (1=yes/0=other response)	-.08*	-.10**	-.08**	-.02
Says color doesn't have anything to do with smartness (1=agree/0=other)	.13***	.11***	.01	-.10***
(5) *Interaction with students of the other race*				
Says three students talks to most are of the same race (1=yes/2=not all same)	-.10**	-.10**	-.04	-.04
Had called a student of the other race on the phone (1=no/2=yes)	.03	.07*	.09***	.08**
Had helped a student of the other race with homework (1=no/2=yes)	.05	.14***	.17***	.05*
Had asked for help with homework from student of other race (1=no/2=yes)	.08*	.05	.08**	-.02
(6) *Perception of teacher's attitudes to desegregation and minorities*				
Teachers seen as liking blacks and whites going to same school (1=yes/0=other)	-.01	.02	.01	.05
Teachers seen as disliking blacks and whites at same school (1=yes/0=other)	-.04	-.01	-.06*	.00

Table 3--continued

Noncognitive Variables	Black Male	Black Female	White Male	White Female
Principal seen as liking blacks and whites at same school (1=yes/0=other)	.07*	.08*	.05	.02
Principal seen as disliking blacks and whites at same school (1=yes/0=other)	-.01	-.03	-.06*	.02
Teachers seen as being unfair to blacks (1=yes/2=no)	.03	-.02	-.02	-.11***
Teachers seen as being unfair to whites (1=yes/2=no)	.08*	-.03	.11**	.01
(7) *Student's self-esteem*				
Student's rating of his ability compared with others in his class (Range: 1-5)	.32***	.39***	.46***	.45***
Student's assessment of ability to complete college (Range 1-5. 5=could complete)	.33***	.39***	.47***	.48***
Student says he will go to college (1=no/2=yes)	.25***	.28***	.40***	.41***
Student says he does not have much to be proud of (1=agree/2=disagree)	.16***	.21***	.12**	.18***
N based on	906	1037	1538	1611

* = significant at the .05 level.
** = significant at the .01 level.
*** = significant at the .001 level.

not appreciably different from the correlation for black females ($r = .07$, $r^2 = .0049$) in that the larger correlation potentially accounts for less than one-half of 1 percent of additional variance.

The second variable that was identified as having possible importance in accounting for the performance of black males is a measure of their interaction with adults in the school. The comparison showed that black males were more likely to say there was an adult they could turn to in trouble if they were in control schools than if they were in experimental schools. Further, there were no significant differences between the two types of schools for the other three race and sex groups. If it is now found that this noncognitive variable is negatively correlated with achievement, it will suggest a possible chain of events that led to the higher performance of black males in experimental schools. These correlations are shown in Table 3, Part 1. In three of the four groups the sign of the correlation is negative, which suggests it is either a disadvantage to feel you can go to an adult if you need to, or else that the kinds of students who either want or need to have this kind of contact with adults are less likely to do well in school work. In any case, although the signs of the correlations are negative, they are also all indistinguishable from zero. Therefore, although the correlations suggest a causal sequence that is important, it is possible that the true value of these correlations is zero or positive.

The second way of inspecting Table 3 ignores the findings of Section III and looks for any evidence of relationships that are different for the four race-sex groups. Given this aim, the search is for positive correlations that are stronger for black male 10th graders. Or, where the noncognitive variable is coded so that a low score is favorable, the search is for negative correlations that are weaker for the black males.

Students' attitudes to those of the other race (Part 4) seem to fall into the first category. That is, "favorable" responses to these questions are more strongly associated with achievement for black males than they are for the cher race-sex groups. Take, for example, the last variable in Part 4, if the student says that race has nothing

to do with being smart. Disagreement is interpreted as evidence of a more favorable attitude toward others and toward oneself, and this is anticipated to mean a better disposition toward school work and to possible academic success. In any event, it is found that the favorable answer to this question is more strongly related to achievement for black males than it is for the other three groups. This means the belief that race and smartness are unconnected is more important for this subgroup than the others. It should be added that this is race-linked rather than sex-linked. Black males are little different from black females. But blacks (of both sexes) seem to be quite different from whites (of either sex).

This finding is more or less supported for other measures of attitude toward one's own race and the opposite race. Disagreeing that either one's own ethnic group or the other one is "dumb" seems to be more closely associated with high achievement for blacks than it is for whites (Part 4). But though this suggests that self-image and image of the other ethnic group and, perhaps, the facility of intergroup relations are more important to blacks than to whites, there is contradictory evidence that cannot be ignored. For example, the idea that self-image is more important for blacks than for whites is undermined by the correlations reported in Part 7, which show self-esteem is almost always less closely associated with achievement for blacks than for whites. Equally, it does not help blacks if they feel that blacks are smarter than whites (Part 3), though predictably enough it hinders them if they feel whites are smarter. Therefore there has to be some uncertainty that self-image or image of the other race is going to explain differences in achievement levels of the four race-sex groups. However, this is the most promising line of inquiry opened up by this part of the analysis and deserves further attention..

The obvious question is whether ESAP funding can be linked with these noncognitive variables for black males. The answer is found in Table 2, Part 4, which shows clearly that control school students, not experimental school students, had more favorable scores on the two most important attitudinal variables (Student says race not linked to IQ, and does not describe own race as "dumb"). In other words there is no

chance of linking ESAP funding to these noncognitive variables, even though the linkage between them and student achievement could possibly explain the differences in achievement level among race-sex groups. The second question is whether particular educational programs might account for the differences in student attitudes. This will be pursued in Section V.

V. WHAT DID ESAP MONEY BUY?

Up to this point the analyses reported here have treated ESAP as a program consisting of additional funding aimed at helping schools make progress with desegregation. This is the most comprehensive way of looking at the program. But it also ignores another level of reality. Money was translated into actions--resources and programs in the schools--and these activities will be the subject of the rest of this investigation. Specifically, what were the federal funds used for in schools? And were some programs more effective than others? The destination of ESAP money was decided, in part, at the local level, and, not surprisingly, different school districts concocted different schemes. The activities authorized by the bill put very little off limits (notably, ESAP money could not be used for busing where it was already in use to fulfill the requirements of a court-ordered desegregation plan).

It is extraordinarily difficult to specify the content of the program in simple and precise terms. ESAP paid for remedial education, the development of instructional methods, the repair of buildings, efforts to involve the community in the school, training for teachers, hiring teacher aides, secretarial assistance--the range was wide. This presents obvious problems to the analyst who wants to make accurate generalizations about the nature of the program. The problems are compounded by the terms used by the evaluators. For example, how certain is it that "inservice training programs" really have enough in common to be sensibly defined as a single set of activities? Or, more simply, what did "comprehensive planning" consist of? Again, when ESAP paid for counseling, what did this involve and which students benefitted? Working at this distance from the program there is no hope of retrieving the answers; the best that can be done is to cross-check different sources of information collected by the evaluators. The definition of program types will be reported on below.

Yet, there is a more basic point, one that turns this part of the secondary analysis into a general investigation of the effects of

different program types rather than an investigation of the effects of ESAP-funded programs. To make the point concretely, I shall anticipate the task of analyzing the effect of inservice training programs. From the point of view of the evaluation of ESAP, the relevant issue is whether ESAP-funded inservice training programs were more (or less) effective than other ESAP-funded activities. In other words, how does this use of program funds compare with other uses? The first step in such an analysis would be to identify schools with ESAP-funded in-service training programs. This can be done with available data. But having got to this point it cannot be assumed that the other schools do *not* have inservice training. Other experimental schools may have this kind of program, funded by other sources. So too may the control schools have inservice training, funded either by ESAP or other sources. So unless it is claimed that the ESAP-funded inservice training is qualitatively different from programs funded by other means, the analysis cannot go forward. There is no evidence that ESAP-funded programs were special in any respect; therefore, the question has to be rephrased. Instead of ESAP-related activities, the objective of the analysis must be to investigate the effects of different programs regardless of funding source.

It is important to stress that by altering the question the investigation is no longer strictly an investigation of ESAP. It becomes an inquiry into the relationship between educational programs and outcomes, such as student attitudes, without regard to the origin of the programs. It is also important to emphasize that this analysis cannot capitalize on the strengths of the randomized design because program types were not randomly distributed among districts. So it is only reasonable to suspect particular circumstances were associated with particular program types. For example, it might be that schools that had experienced a lot of racial tension will be more likely to choose inservice training with an emphasis on intergroup relations. A school with lots of slow learners might be more likely to use ESAP money for remedial education; some of these possibilities will be investigated in the second part of the section. There is a reasonable possibility that program type is confounded with other school level factors, and these

factors may well account for variations in outcomes. The point is that the analysis is hampered by the usual constraints of confounded independent variables.

The issue addressed in this section is how best to characterize the program types. The information used for this purpose is obtained from three sources, the first two of which are much the most important.

THE ESAP DIRECTOR SURVEY

The ESAP director was most usually the person responsible for federal programs in the district, but sometimes it was the superintendent. He, or his assistant, was interviewed and asked about the use of ESAP funds and it is important to note that only the director survey refers to ESAP (see below). Further, most information the director supplied related to activities and programs in the schools, rather than in the district as a whole. And filter questions were added to check that these activities were ones that 10th grade students could have participated in.

	Informant asked if ESAP funds bought program?	Do data relate to the school or the district?	Is program available to 10th grade?
General program description (e.g., inservice training)	Yes	School	Yes
Equipment (e.g., AV equipment)	Yes	School	Yes
Specific inservice (e.g., remedial education)	Yes	School	No
Specialists (e.g., teacher aide)	Yes	School	Yes
General program description (teacher aides)	Yes	District	No

THE PRINCIPAL SURVEY

The principals, unlike the directors, were not asked about the use of ESAP money. Their answers provide information about the resources and programs in their schools regardless of funding source.

	Did ESAP buy program?	Do data relate to school or district?	Is program available to 10th grade?
Programs, courses, or personnel	No	School	Yes
Additional funds for equipment of different types	No	School	No
Existence and content of inservice training	No	School	No

THE TEACHER SURVEY

This is the least important survey for the purpose of program identification; it was not designed to find out what activities there were in schools. However, incidental questions do touch on the teachers' experience with a few programs, and this information will be used to cross-check the existence of programs in schools:

	Did ESAP buy program?	Do data relate to school or district?	Is program available to 10th grade?
Program in parent-teacher relations	No	School	No
Minority courses or texts	No	School	No
Teacher had had in-service program	No	School	Yes

PROGRAM TYPES

These data are used to define 11 dichotomous variables identifying program type (program presence contrasted with absence). These are different categories from those used in previous evaluations of ESAP (RMC, 1971; NORC, 1973), and there are three reasons for the creation of new variables. First, unlike previous investigators, I have tried to avoid creating composite indices. Since the terms used to identify ESAP activities were broad enough to start with, there seemed every reason not to compound the problem by putting several program types together in one variable. Besides, NORC's factor analysis of programs and activities (NORC, 1973, pp. 35-41) did not encourage the view that the data would yield readily interpretable composite indices. Second, with the limited set of 48 school pairs it turned out that many programs were not represented in the sample of schools. A variable for program type was discarded where there were four or fewer schools reported to have a particular program type. A great many of the variables turned out to fail this test; for the director survey, 41 out of 75 variables were not used for this reason. Third, the definition of program types was guided toward activities that might plausibly be used to improve educational achievement or alter racial attitudes.

The first program type, inservice training for teachers, like most of the other programs, is identified by both the director and the principal. The director survey indicated that there were 23 experimental schools and 11 control schools with inservice teacher training (see p. 42). The teacher survey, however, indicated 38 experimental and 34 control schools with inservice training. Different schools are identified by directors and principals, most likely because the principal gave information about programs of all kinds, not just ESAP funded activities.

The breakdown separates experimental and control schools to bring out two points. The director identifies a substantial number of control schools (11) as having ESAP-funded inservice training programs. There are 15 experimental schools (38 less 23) with inservice training funded by non-ESAP sources. The point has been made before: There is little sense in contrasting schools with ESAP-funded inservice training and those without. Many of the residual schools have inservice training,

and many of the control schools have ESAP-funded inservice training. The same point is implied in the following summaries.

	Experimental	Control
Inservice training with emphasis on intergroup relations		
Director survey	11	4
Principal survey	31	26
Inservice training with emphasis on remedial education		
Director survey	5	1
Inservice training with emphasis on teaching methods		
Director survey	10	4
Principal survey	38	36
Inservice training with emphasis on curriculum development		
Director survey	9	3
Principal survey	38	32
Provision of remedial education		
Director survey	10	3
Principal survey	40	36
Curriculum development		
Director survey	3	1
Principal survey	30	20
Programs to increase contact between school and community		
Director survey	10	5
Principal survey	29	15
Focus on relations between students		
Director survey	5	0
Principal survey	31	23

Courses or materials oriented toward minority groups	Experimental	Control
Director survey	3	0
Principal survey	13	17
Provision of teacher aides		
Director survey	34	35
Principal survey	29	23

These summaries bring out the point that ESAP money was spread over all kinds and varieties of educational programs. There is no single or dominant theme, except perhaps that the money seems to have been used for teacher training, especially training with an emphasis on intergroup relations, and on teacher aides. And in view of the size and timing of the grants in each school, it is likely these programs did not amount to profound, deep-cutting changes in the organization of the school but were more probably short-term remedies and responses to integration.

A NOTE ON PROGRAM DELIVERY

So far only the director and principal data have been used to determine what the ESAP money was used for. But since the matter is so important, it is worth taking one step further by looking at the teacher survey data. The purpose of doing this is to see if the teachers' answers corroborate the information given by director and principal.

Teachers were asked several questions about the resources in their school, particularly as they were affected. For example, they were asked if they had taken any inservice training during the year, how often they had discussions in class about race, and if they had learned about handling intergroup relations among students. They were also asked about the school in general--for example, whether the school had made efforts to get parents to visit the school during the year, or whether there had been special projects (plays or group discussions) that dealt with intergroup problems. Data from these questions were aggregated at the school level to give an average teachers' score for

each school. Thus, if all the surveyed teachers in a school reported they had had inservice training, then the aggregate score for that school would be 100; if half had received training, the score would be 50; and so on. These aggregate teacher responses can then be related to data obtained from the director and principal surveys. If the program was effectively delivered in the school, there will be reasonable agreement between the two sources of information. There is one caution. The program described by the director may not have been intended for the 10th graders or their teachers, so lack of consistency may not always point to uncertainty about the existence of a program.

In some areas, the director and principal data check out very poorly with the teacher information. For example, if schools are divided on the basis of the director data into those that had programs to develop parent-teacher relations and those that did not, we find that teachers in the first category of schools are *less* likely to report their school was "trying harder this year than it had in the past to get parents to visit the school or come to PTA or other parent groups." The aggregate teacher response (in agreement with the question) for the first kind of school was 18.0, and in the second kind of school, 30.6. If the same check is done using the principal data, the comparable figures are 28.4 and 28.7. In other words, the teacher reports imply ESAP made little difference to the program schools.

In other areas there is greater consistency. For example, schools where the director said there were ESAP-funded inservice training programs (available to the 10th grade teachers) had teachers who were more likely to report receiving inservice training than schools without this program. Schools were divided on the basis of directors' information about inservice training. In the schools with this program, 63.2 percent of the teachers, on average, said they had had some inservice training. Where there was no program, the average response fell to 51.8 percent. However, when the principals' information is used, this difference between program and nonprogram schools disappears (55.5 percent of the teachers said they had had training in schools without, and 55.9 percent said they had training in schools where programs were reported).

Better agreement is obtained about whether teachers received

inservice training with an emphasis on intergroup or race relations. When either the director or the principal says the school had a program of this kind, the teachers are more likely to say they had had some training in this area. Schools where the director claimed such a program had an average of 44.7 percent of the teachers saying they trained in intergroup relations (37.3 percent in schools without these programs). Schools where the principal claimed this kind of program had an average of 45.8 percent of the teachers reporting training in the area (27.7 percent of teachers in the schools without). This coherence is disturbed only by the fact that teachers in schools with inservice training in intergroup relations were less likely to report they had learned how to handle intergroup relations than teachers in other schools. But one further check gives weight to the positive side. Teachers were asked if they had class discussions about race. They were more likely to say they had these discussions if they were in a school that reported inservice training with an emphasis on race relations.

Finally, there is greater likelihood that principals and teachers will concur about the existence of courses in minority history or culture. Teachers were considerably more likely to say there was a course of this kind in the school (35 percent) if the principal indicated there was than they were if the principal said there was no minority course (teachers: 7.1 percent).

It is clear the teachers' information does not help a great deal in identifying what ESAP consisted of in the schools. Sometimes their reports supported what the director said ESAP funds had been used for; more usually they did not. This null finding is ambiguous, as pointed out in the beginning. Without much more precise information it is possible only to say that this conflicting evidence warrants real concern about the implementation of ESAP; it does not prove poor implementation.

WHY DID SCHOOLS CHOOSE THEIR PARTICULAR PROGRAM?

The ESAP money was used to buy a variety of different programs and services. It is reasonable to think that the choice reflected local needs; the supplementary funds bought things that the principals or the district administrators wanted most. If so, there should be some

relationship between the characteristics of the schools and the programs they got. To see if there were relationships of this kind, the program variables were cross-tabulated with a set of variables that characterize the schools. The latter included the school's racial composition, the principal's evaluation of the quality of his faculty, the principal's attitude to desegregation, the year desegregation had its greatest effect on the racial balance of the school, and the level of racial tension in the school during the previous year (indicated by whether the school was closed). Since the data were not collected with this analysis in mind, it is not surprising that this list is short. These few variables describe schools in only a limited way.

With one exception, the analyses provided no evidence that the schools' characteristics or special problems were related to the kinds of programs they had. The exception is that schools that had inservice training programs with an emphasis on race relations were more recently desegregated than other schools. This makes sense; a school just going through the process of desegregation is presumably in greater need of programs that deal with the relationships between ethnic groups. And the idea that there are special, temporary problems involved in desegregation received some modest support in earlier analysis (Section II). However, it would be equally sensible to expect that schools that had this kind of inservice training for their teachers would also have experienced greater racial tension, they might have principals that were more in favor of desegregation, or they might have a particular racial composition. But no association was found between these three variables and the presence of this kind of program (the criterion for deciding if two variables were associated was that χ^2 should exceed the value for $p = 0.2$).

The rest of the analysis revealed no other relationships between program type and school characteristics. For example, the principal's evaluation of the quality of his faculty is not associated with the probability that the school has an inservice training program, nor are programs that concentrate on remedial education more likely to be found in schools where principals feel that blacks' failure can be attributed to the restrictions imposed by white society rather than to students'

own failings. Schools that chose programs to help with the relationships between ethnic groups are not unusual in their racial composition, the year in which they desegregated, nor in the level of racial hostility in the school (measured by whether the school closed because of racial problems). Finally, schools that chose to spend money on teacher aides seem to have had just about the same student-teacher ratio as those that did not.

With the exception of the relationship between year of desegregation and inservice training programs focused on race relations, there are no revelations about the circumstances that led to the selection of particular programs. This most probably means that we do not have the appropriate measures to characterize the schools. However, it is then possible to infer that the schools did not get the most appropriate kind of program from the point of view of the principal, administrators, or students. ESAP funds were made available after the beginning of the school year, and the decision about fund usage was not entirely made at the school level; these facts give some weak support to this inference.

VI. RELATIONSHIP BETWEEN PROGRAM TYPE AND PROGRAM OUTCOMES: TEACHER AND STUDENT NONCOGNITIVE VARIABLES

Program-type variables identified in Section V are used here to investigate the effect these programs had on noncognitive outcomes. The purpose is to see if the existence of the programs in schools is associated with changes in attitudes and reported behavior of teachers and students. Program type variables are derived from the principal survey, which means the analysis is a general investigation of the relationship between educational activities and outcomes. As pointed out before, the analysis is not tied to the Emergency School Assistance Program.

TEACHER OUTCOMES

There are many teacher attitude variables, the dependent variables in this analysis, and there are 11 program type variables. Interrelating each program type with all possible outcome measures would have produced an indigestible array of results, so selection was called for. The outcome measures chosen were those I thought most closely represented the things ESAP was trying to change. They fell into three categories: teachers' attitudes to desegregation and to minority students, teachers' attitudes to the academic potential of students, and teachers' self-report on how they felt about teaching that year. The dependent variables were divided into these three groups because it seemed likely that different outcomes would be influenced by different program types. For example, it seemed likely that teacher training in race relations would have more of an effect on attitudes to desegregation than it would on how much teachers enjoyed their job. Of course, this kind of teacher training might have an effect on job satisfaction too, but parsimony directed attention to the best bets first. Accordingly, different selections of program type variables were related to each of the three sets of variables defined above. The first part of the analysis uses the simplest possible approach: a comparison of means of the outcome measures for schools with and schools without the program. The difference

between these means represents initial indication that the program is associated with a difference in outcome.

Teachers' attitudes to desegregation and to minority students are represented by 14 variables (Table 4, Part A) that deal with their feelings about racial mixing in schools ("Do you think blacks are better off in racially mixed schools?"), with their self-reported behavior in class ("Do you have discussions about race?") and with their perception of other teachers' attitudes to desegregation. These were identified as possible outcomes for three kinds of educational programs. The most obvious is teacher training in race relations, which is assumed to be directly aimed at changing teachers' attitudes concerning racial issues in the school. The other two, the presence of courses in minority history or culture and unspecified teacher training activity, though less obviously connected with attitude change, were assumed to have an indirect effect on the same class of outcomes.

General inservice training programs are not associated with any significant differences in teachers' attitudes. For example, they do not lead to more favorable teachers' attitudes toward minority students or to the value of desegregation, nor do they change teachers' openness with students about the matter of race. This is most probably because of the great variety of activities included under the classification "inservice training," each with somewhat different objectives. By contrast, teacher training specifically directed at race relations is associated with differences in attitude and reported behavior, which, from the point of view of the program, must be regarded as favorable. Teachers in schools with these programs are more likely to think white and black students profit from racial mixing, their personal attitudes to race relations are more favorable, and they see their fellow teachers, black and white, being better disposed to desegregation. When schools have courses in minority history or culture, the teachers are also more likely to have favorable attitudes on some of the same outcomes; the effect of this program is somewhat less pronounced than that of inservice training emphasizing race relations.

As always, these positive conclusions must be balanced against some familiar disclaimers. First, the effects are quite small ones,

Table 4

T-TEST OF MEANS OF TEACHER OUTCOME VARIABLES COMPARING PROGRAM AND NONPROGRAM SCHOOLS

Program Variables		Inservice Training with Special Emphasis on Race Relations	Presence of Courses in Minority History or Culture	Inservice Training, No Special Emphasis
Part A: *Attitudes Toward Desegregation and Minority Students*				
Teacher thinks black students better off in racially mixed schools (0=no/1=yes)	NP	.677*	.708	.719
	P	.753*	.754	.723
Teacher thinks white students better off in racially mixed schools (0=no/1=yes)	NP	.507***	.550**	.569
	P	.639***	.664**	.590
Proportion of white students teacher thinks as discipline problem. (Range 1-5. High=few)	NP	4.331	4.291	4.290
	P	4.271	4.310	4.298
Proportion of black students teacher thinks as discipline problem. (Range 1-5. High=few)	NP	3.796	3.829	3.749
	P	3.762	3.653	3 784
If teacher feels it is proper to let students know his opinion about race (0=no/1=yes)	NP	.675**	.719	.732
	P	.754**	.729	.719
If teacher has class discussions about race (0=no discussion/1=some)	NP	.471	.488	.451
	P	.530	.545	.522
Thinks amount of prejudice is exaggerated. (Range 1-4. 4=disagree)	NP	2.335	2.401	2.529*
	P	2.446	2.399	2.361

Table 4--continued

Program Variables		Inservice Training with Special Emphasis on Race Relations	Presence of Courses in Minority History or Culture	Inservice Training, No Special Emphasis
Part A: *Attitudes Toward Desegregation and Minority Students* (continued)				
Would like to live in integrated neighborhood. (Range 1-4. 4=disagree)	NP	2.029*	2.073	2.157
	P	2.158	2.178	2.089
Feels civil rights movement has done more good than harm. (Range 1-4. 4=agree)	NP	2.696	2.770	2.863
	P	2.872	2.866	2.779
Feels blacks and whites should not be allowed to marry. (Range 1-4. 4=disagree)	NP	2.448**	2.532*	2.598
	P	2.686	2.719	2.586
Teacher's estimate of number of teachers unfair to black students. (Range 1-5. 5=none)	NP	4.101	4.078	4.045
	P	3.998	3.958	4.040
Teacher's estimate of number of teachers unfair to white students. (Range 1-5. 5=none)	NP	4.351	4.356*	4.340
	P	4.275	4.181	4.297
Teacher's perception of white teachers' attitude to desegregation. (Range 1-5. 5=favorable)	NP	2.315***	2.467**	2.601
	P	2.711	2.718	2.529
Teacher's perception of black teachers' attitude to desegregation. (Range 1-5. 5=favorable)	NP	2.780***	2.955*	3.075
	P	3.185	3.145	2.995

Table 4--continued

Program Variables		Presence of Remedial or Tutoring Programs	Inservice Training with Emphasis on Teaching Methods	Inservice Training with Emphasis on Remedial Education
Part B: *Evaluation of Academic Potential and Performance*				
Teacher's estimate of proportion of black students performing adequately. (Range 1-4. 4=high)	NP	2.363	2.279	2.355
	P	2.358	2.383	2.360
Teacher thinks black failure to achieve equality due to lack of initiative, not society's restrictions (0=agree/1=disagree)	NP	.444	.530	.438
	P	.482	.462	.493
Teacher feels that students are learning more (0=no/1=yes)	NP	.243	.271	.281
	P	.288	.280	.272
Teacher's estimate of proportion of white students that could get to state university. (Range 1-4. 4=almost all)	NP	2.101	1.990	2.029
	P	1.996	2.028	2.016
Teacher's estimate of proportion of black students that could get to state university. Range 1-4. 4=almost all)	NP	1.466	1.325*	1.319**
	P	1.409	1.450	1.460

Table 4--continued

Program Variables		Presence of Counselors	Presence of Teacher Aides	Inservice Training with Emphasis on Teaching Methods
Part C: *Job Satisfaction*				
Teacher says she enjoyed teaching more this year than last (0=no/1=yes)	NP	.288	.361	.360
	P	.352	.329	.339

P = Program
NP = Non-program
* = significant at the .05 percent level.
** = significant at the .01 percent level.
*** = significant at the .001 percent level.

even though they are statistically significant. For instance, 75 percent of the teachers in schools with race relations inservice training and 68 percent of the teachers in schools that did not have this kind of program said black students would be better off in racially mixed schools, a difference of 7 percent. Though statistically significant, and undoubtedly of potential practical importance, this kind of difference cannot be called overwhelming. Second, in dealing with a sensitive issue like racial attitudes, respondents are likely to give socially "correct" answers rather than their true opinions, particularly if they have just been through a course to sensitize them to some of the issues connected with desegregation. In other words, there is an inevitable problem of validity with self-completed questionnaires. We cannot be sure, just because teachers told us so, that white teachers were really more favorably disposed to desegregation. Still less can we conclude that their attitudes, as perceived by other teachers, were connected with their behavior. Nevertheless, we are looking for clues rather than certainties, and we have to strike a balance between agnosticism and naivete.

A final problem with these results is that they are heir to the usual limitations of post hoc observational surveys; in other words, the observed differences might well be attributed to other causes than the one under consideration. Specifically, the differences in teachers' attitudes might be attributable to other background characteristics of these teachers rather than to the program in the school. Accordingly, regression analyses were run that are analogous to the t-tests reported in Table 4 yet that allow for control of other independent variables. Two were chosen: teacher age and length of experience with desegregated schools. The clear conclusion of this analysis was that the controls for these characteristics did not significantly alter the relationships between program type and teacher attitude.

Another possibility was that characteristics of the school might explain the observed differences. One promising counterexplanation of the effect of inservice training in race relations was based on the earlier finding that recently desegregated schools were more likely to have this kind of program. Also teachers' perceptions of the racial

climate in the school were related to the recency of desegregation. Therefore the differences in attitudes that are apparently due to the existence of a program might really be a reflection of the stage in the desegregation process different schools have reached. Accordingly, a second series of regressions were run where controls were introduced for the time the school had been desegregated. Two dummy variables were entered identifying schools desegregated during 1971 and those desegregated during 1969 and 1970. The introduction of these controls, like the earlier analysis, did not alter the original conclusion that the program was associated with more favorable attitude outcomes.

The second part of the analysis (Table 4, Part B) looks at the way teachers evaluate the academic potential and performance of their students. The outcomes here cover teachers' estimation of the students' ability to get into college and of the proportion of black students performing "up to scratch," their opinion about students' performance in general during the year and about the reasons black students do not achieve "equality."

These outcomes were related to three kinds of programs. The most promising one identified schools with programs in remedial education. The other two were both inservice training activities, one that concentrated on curriculum development and the other on teaching method. None of these programs seemed to make a consistent difference to teachers' attitudes about students' ability or capability. There are two statistically significant results, but they are overwhelmed by the other nonsignificant differences. This apparently discouraging result quite probably indicates the ambivalence implicit in remedial programs; for although they may encourage more positive attitudes on the part of teachers, the fact that "underachievers" are getting specialized treatment in the school may encourage a negative attitude. Those predisposed to think black students are unlikely to succeed may be confirmed in their views if they see these students getting extra help.

The final piece of analysis (Table 4, Part C) looks at a measure of job satisfaction: teachers' responses to a question about how much they enjoyed teaching during the current year compared with last year. This, though far from ideal, was the best available measure of their disposition

to their job. Three program types were related to this outcome; two identified the provision of aides and counselors in the schools, and the third identified teacher training that focused on teaching methods.

None of these program types was associated with statistically significant differences in the dependent variable. Again, this may be because teachers do not appreciate the efforts of aides or counselors, or it may be that these programs really do not make life easier for teachers. In any case, there is no discernible relationship between these programs and teachers' attitudes to their job during the year.

In conclusion, the strongest evidence of program effects is contained in the first part of the analysis, specifically with the analysis of inservice training focused on race relations. It can be argued that this is as it should be. ESAP was intended to assist the process of integration and teachers' attitudes are certainly an important element in that process. Programs, some of them funded by ESAP, that aim directly at changing racial attitudes are associated with the expected effect. There must be reservations about the finding, but it is a positive indication of program success and it points to areas for more complicated investigation, to be reported in Section VII.

STUDENT OUTCOMES

The analysis of student outcomes is similar to that of teacher outcomes. The issue is whether there are educational programs that can be related to differences in students' perceptions, attitudes, and self-reported behavior. The question is directly related to ESAP's goals of improving the quality of relationships in desegregating schools. The measures of student outcomes used in the analysis give some indication of the way they felt about their school, the way they saw the teachers and principal, their interaction with students of the opposite race, and their self-esteem.

Six of the program types were chosen for analysis on the basis of assumptions about the possible and likely connections between educational programs and students' attitudes and behavior. Each program variable is entered in a regression analysis and its effect on the dependent measure is represented by the standardized partial regression coefficient. This

method of analysis was chosen because it provided the most convenient way of controlling for other independent variables, in this case a measure of home background (years of education completed by student's mother) and the percent of white students in the school. Both variables were seen as possible determinants of student attitudes and were therefore entered first in the regression analyses, prior to the dummy variable identifying program type. In Table 5 the standardized coefficient for the program type variable (with associated level for t where beta is statistically significant) is the only information reported from these analyses.

The total sample of students is divided into the four race and sex groups. This is consistent with the general aim of this inquiry to supply an explanation of the ESAP effect on black male achievement. To repeat, I am not looking here at the effects of programs that were exclusively funded by ESAP, but at the effects of programs funded by many different sources in the hope of identifying effective strategies. This is the nearest approach that can be made to the preferred analysis. As a first step, student noncognitive variables were selected for analysis on the basis of their correlation with achievement scores. This, too, was justified in terms of the overall purpose of identifying intervening variables--here the student noncognitive measures--that might explain the association between program type and student achievement. Clearly, it makes more sense to investigate dependent variables that account for variance in achievement.

The analyses are presented in Table 5, which divides the results on the basis of program type. Some of the programs appear to have little effect. For example, Part A of Table 5 shows the coefficients for general inservice training programs. These, with one exception, are statistically insignificant, which suggests inservice training is not associated with either higher or lower levels of student attitudes. This may not be too surprising since a variety of objectives are subsumed under this broad category of programs and since inservice training is directed at teachers, not students.

Other programs are associated with more favorable student attitudes. One of these is inservice training that concentrates on race relations,

Table 5

MULTIPLE REGRESSION ANALYSIS OF RELATIONSHIP BETWEEN PROGRAM TYPE VARIABLES AND MEASURES OF STUDENT NONCOGNITIVE OUTCOME, BY RACE-SEX GROUPS

Dependent Variables	Black Males	Black Females	White Males	White Females
Part A: *Analysis of Inservice Training Program for Teachers*				
Student's rating of his ability compared with others in class	.06	.01	.03	-.02
Teachers seen as liking blacks and whites at same school, by student	-.03	.03	.01	.01
Student had called student of the other race on the phone	.05	.01	.03	.04
If student could choose, would go to a racially mixed school	.01	.01	-.04	-.03
Student says he hates school	-.01	.01	-.04	-.00
Student is a member of school club or sports team	.02	-.03	.03	.05* (1.99)
Part B: *Analysis of Inservice Training with Emphasis on Race or Intergroup Relations*				
Student's rating of his ability compared with others in class	.05	-.01	.03	.04
Teachers seen as liking blacks and whites at same school, by student	.10** (2.87)	.00	.05	.09*** (3.52)
Student had called student of the other race on the phone	.00	-.02	.07** (2.78)	.09*** (3.61)

Table 5--continued

Dependent Variables	Black Males	Black Females	White Males	White Females
Part B: *Analysis of Inservice Training with Emphasis on Race or Intergroup Relations* (continued)				
If student could choose, would go to a racially mixed school	.09** (2.62)	.07* (2.31)	.08** (3.17)	.08** (3.21)
Student says he hates school	.05	-.06	.04	.01
Student is a member of school club or sports team	.01	.04	-.05	-.14*** (5.53)
Part C: *Analysis of Inservice Training with Emphasis on Teaching Methods*				
Student's rating of his ability compared with others in class	-.01	-.01	.07** (2.68)	.03
Teachers seen as liking blacks and whites at same school, by student	.01	.07	-.00	-.05
Student had called student of the other race on the phone	-.06	.05	.05	.08*** (3.37)
If student could choose, would go to a racially mixed school	-.06	.04	.07** (2.87)	.05* (2.09)
Student says he hates school	-.01	-.05	.02	.04
Student is a member of school club or sports team	-.06	-.00	.02	-.01

Table 5--continued

Dependent Variables	Black Males	Black Females	White Males	White Females
Part D: *Analysis of Programs in Intergroup Relations among Students*				
Student's rating of his ability compared with others in class	.08* (2.32)	-.02	.06* (2.32)	.03
Teachers seen as liking blacks and whites at same school, by student	.02	.06* (1.98)	.03	.01
Student had called student of the other race on the phone	-.02	.04	.08** (3.10)	.09*** (3.61)
If student could choose, would go to a racially mixed school	-.01	.10** (3.13)	.06* (2.26)	.09*** (3.76)
Student says he hates school	.00	-.02	-.02	.01
Student is a member of school club or sports team	-.02	.08** (2.63)	-.03	-.07** (2.64)
Part E: *Analysis of Programs in Minority Culture or History*				
Student's rating of his ability compared with others in class	.04	.02	.05	.02
Teachers seen as liking blacks and whites at same school, by student	.04	-.02	.00	.03
Student had called student of the other race on the phone	-.02	.08** (2.61)	.07** (2.67)	.08** (3.16)

Table 5--continued

Dependent Variables	Black Males	Black Females	White Males	White Females
Part E: *Analysis of Programs in Minority Culture or History* (continued)				
If student could choose, would go to a racially mixed school	-.07* (2.15)	.00	.04	.08** (3.16)
Student says he hates school	.00	-.02	.04	.03
Student is a member of school club or sports team	-.00	.07* (2.22)	-.03	-.01
Part F: *Analysis of Program in Parent-Teacher Relations*				
Student's rating of his ability compared with others in class	.07* (2.32)	-.01	.06* (2.30)	.06* (2.45)
Teachers seen as liking blacks and whites at same school, by student	-.01	.05	.03	.01
Student had called student of the other race on the phone	.02	-.04	.06* (2.17)	.03
If student could choose, would go to a racially mixed school	.04	.04	.04	.03
Student says he hates school	.09** (2.61)	-.03	.01	.00
Student is a member of school club or sports team	.03	-.05	.08** (2.96)	.03

t- statistics in parentheses.
* = significant at the .05 percent level.
** = significant at the .01 percent level.
*** = significant at the .001 percent level.

(Part B). In several of the analyses the presence of the program is associated with more favorable student attitudes of seve ıl kinds. In particular, it is associated with an increased probability students will say they would prefer a racially mixed school over a segregated one. This is encouraging for two reasons. First, this type of student outcome might be expected to change as a result of altered teacher attitudes. Further, student attitudes that would not be predicted to respond to changes in teachers are found to have no association with the presence of this kind of program. For example, this type of inservice training does not appear to make any difference to how much students like to go to school. Second, this is precisely the kind of program that reflected ESAP's goal of achieving successful desegregation. Moreover, inservice training that has a different kind of emphasis, teaching methods (Part C), does not have the same effect on student attitudes. This too falls in line with expectations. Programs that set out to alter teachers' attitudes and behavior apparently succeed in changing the way students feel about things too.

Most of the positive and significant results in Part B are concentrated among white students of both sexes. This suggests that black and white students respond in different ways to their educational environments. But it also suggests that these program effects on noncognitive outcomes are not going to account for the relationship between ESAP and the educational performance of black males. If blacks are not affected by these kinds of interventions, then there seems little hope of sketching the connection of events between the ESAP funding on the one hand and the higher test scores of black students on the other.

More positive program effects are found in Part D, which reports the effects of programs in intergroup relations designed for students. Here too the positive and statistically significant results outweigh the number that would be expected by chance by a large factor. And again, this kind of activity is associated with the differences in attitude that might be expected. For instance, it is associated with a larger proportion of students who say they prefer integrated over segregated schools and with a larger proportion of students (white only)

who interact with students of the other race. However, the program has little effect on students' perceptions of teachers' attitudes, or on the extent to which students like going to school. This too is encouraging since once again it is the type of program that ESAP fostered and its effect is consistent with ESAP's goals. However, it is worth noting that the program effect is least obvious for black males, so it does not look as if this kind of educational activity could account for ESAP's success with black males.

Parts E and F deal with programs in minority history or culture and those designed to bring a closer relationship between parents and school. Courses oriented to minorities do not seem to make any difference to students' self-esteem, even black students' self-esteem, nor does it alter attitudes about going to school; though, for reasons that are not immediately obvious, it does have a consistent effect on the interactions among students of different races. Since this program did not alter students' outcomes to any substantial extent, and certainly not in the way that was expected, it does not look like a promising topic for further analysis. The same general conclusion holds for the analysis of programs that develop parent-teacher relations (Part F). The effect of this program is uneven, except that it is associated with higher levels of self-esteem among students. However, compared with inservice training in race relations and intergroup programs for students, the effects seem modest.

In review, the strongest evidence of special program effects were obtained in the analysis of inservice training programs that emphasized race relations and programs that develop intergroup relations among students. The analysis of these two program variables was extended, first by looking at the possible confounding effect of recency of desegregation. In the analysis of teacher outcomes it was pointed out that recently desegregated schools were more likely to have this kind of inservice training. The question was whether this difference might explain the levels of student attitudes regardless of program presence. In other words, was the program effect apparent rather than real? The regression analyses for these two program type variables (Parts B and D) were repeated with extra control variables added that grouped schools

according to the recency of desegregation. The results clearly show that recency of desegregation does not substantially alter the uncontrolled relationship between program type and student attitude.

The second step was to investigate further dependent variables. New measures of self-reported behavior and attitude were used in Table 6, following leads from Table 5. These analyses confirmed the previous finding; the positive effects of these two programs tend to be concentrated among white students. Thus inservice training that emphasizes race relations has a positive effect on white student's interaction with blacks and on their attitudes to black students, but there are no corresponding program effects on black students. This is important because previous analysis had shown that two of these racial attitude variables might explain the relative performance of black males in experimental schools. The question is whether the attitude variables might be influenced by the presence of certain kinds of education programs. It is clear that they are not so related; black students' racial stereotyping and their belief about the relationship between race and intelligence are not to be explained by this particular kind of teacher training.

As before, this kind of inservice training is associated with more favorable perceptions of teachers' attitudes and behavior, just what might be expected to result from a program like this. But again, the program does not have the same effect on blacks.

The analysis of programs in intergroup relations for students (Part B) shows comparable results, though less obvious. Most of the positive effects are concentrated among white students as before, but some of the statistically significant relationships are found for the two black groups. Thus, this program too is associated with more favorable attitudes toward students of the other race and higher levels of interaction between races. However, the effect on perception of teacher attitudes and behavior is mixed; in most analyses the relationship is negative.

In summary, the pattern of results accumulated over a substantial number of analyses indicates two programs were associated with more favorable student attitudes and behavior, though this effect is almost

Table 6

RELATIONSHIP BETWEEN PROGRAM TYPE VARIABLE AND MEASURES OF STUDENT NONCOGNITIVE OUTCOME, BY RACE AND SEX

Dependent Variables	Black Males	Black Females	White Males	White Females
Part A: *Inservice Training with Emphasis on Intergroup Relations*				
Would like more friends of the other racial group	-.02	.04	.10*** (3.83)	.08*** (3.10)
Feels uncomfortable with students of the other race (Range 1-4. 4=never)	-.03	-.01	.00	.01
Says three students talks to most are of the same race (Range 1-2. 2=no)	-.05	-.07	.08** (2.93)	.09*** (3.51)
Had helped student of other race with homework	-.03	-.05	.04	.06* (2.44)
Had asked for help with homework from student of other race	.02	-.05	.06* (2.37)	.08** (3.18)
Describes students of the other race as dumb (1=yes/ 2=no)	.04	-.04	.06* (2.07)	.06* (2.12)
Says color doesn't have anything to do with smartness (1=agree/0=other response)	-.01	-.01	.05* (2.02)	.05
Principals seen as liking blacks and whites at same school, by student	.07	.05	.00	.13*** (5.05)
Teachers seen as being unfair to whites (Range 1-2. 2=no)	.05	-.02	-.02	-.03

Table 6--continued

Dependent Variables	Black Males	Black Females	White Males	White Females
Part A: *Inservice Training with Emphasis on Intergroup Relations* (continued)				
Teachers seen as being unfair to blacks (Range 1-2. 2=no)	-.02	.02	.01	-.02
Teachers perceived as liking blacks and whites going to same school (1=yes/0=other)	.10** (2.87)	.00	.05	.09*** (3.52)
Teachers perceived as disliking blacks and whites going to same school (0=yes/1=other)	.06	.08* (2.55)	.08** (2.97)	.13*** (5.01)
Part B: *Analysis of Programs in Intergroup Relations among Students*				
Would like more friends of the other racial group	-.04	-.00	.08** (3.18)	.07** (2.63)
Feels uncomfortable with students of the other race (Range 1-4. 4=never)	-.06	.08* (2.55)	.01	.02
Says three students talks to most are of the same race (Range 1-2. 2=no)	-.03	.02	.09*** (3.67)	.06*** (2.27)
Has helped student of other race with homework	-.04	.11*** (3.93)	.05	.07** (2.76)
Has asked for help with homework from student of other race	.08	.08** (2.70)	.02	.08** (2.98)
Describes students of the other race as dumb (1=yes/2=no)	.03	.02	.00	.07** (2.70)

Table 6--continued

Dependent Variables	Black Males	Black Females	White Males	White Females
Part B: *Analysis of Programs in Intergroup Relations among Students* (continued)				
Says color doesn't have anything to do with smartness (1=agree/0=other response)	.00	.04	.05	.08*** (3.35)
Principals seen as liking blacks and whites at same school, by student	.02	.07* (2.24)	-.02	-.01
Teachers seen as being unfair to whites (Range 1-2. 2=no)	-.10** (2.98)	-.02	-.05	-.04
Teachers seen as being unfair to blacks (Range 1-2. 2=no)	-.08* (2.20)	.00	-.08** (2.99)	-.01
Teachers perceived as liking blacks and whites going to same school (1=yes/0=other)	.02	.06* (1.98)	.03	.01
Teachers perceived as disliking blacks and whites going to same school (0=yes/1=other)	-.01	.13*** (4.31)	.03	.09*** (3.68)

t- statistics in parentheses.
* = significant at the .05 percent level.
** = significant at the .01 percent level.
*** = significant at the .001 percent level.

entirely concentrated among white students. To repeat, these results are encouraging, but they should not be blown out of proportion. They do not tell us how ESAT worked; they only report on the observed relationships between program types and certain teacher and student outcomes in the experimental and control schools. Nor are these effects large ones. The pattern of results supports a positive conclusion, but it would be incautious to make bold predictions about the amount of difference that could be brought about by wider application of these programs. This is particularly true since we have only the most general idea of what the programs consist of in the first place; they would therefore be difficult to reproduce. Although some attempts have been made to test alternative explanations of the findings, such statistical devices can never be exhaustive, by definition. Therefore, there is the ever-present possibility that the observed effects are not properly attributed to the programs, but to some unmeasured or uncontrolled third influence. Despite these cautions, the results are suggestive of ways in which schools alter teacher and student attitudes, and the leads provided by these analyses direct the approach adopted in the following section, which looks again at the determinants of achievement levels.

VII. THE RELATION OF STUDENT ACHIEVEMENT TO ESAP, EDUCATIONAL PROGRAMS, AND STUDENT ATTITUDES

The first part of the Rand work with the ESAP data consisted of a reevaluation of the program's effects on student achievement by various comparisons between experimental and control schools. In this section the analysis of achievement scores is taken in two different directions. First, I look at districts where students in the ESAP school did much better than those in the control school. These instances, called outliers, are examined for possible clues about special or unusual circumstances associated with program success. The districts where experimental school students far outperform those in control schools are compared with districts where the reverse is true. The purpose is to identify dramatic differences between highly effective and ineffective implementations of the Emergency School Assistance Program. Since there are 48 school-pairs to start with, and I am looking at only a fraction of these, this part of the analysis approaches case study.

The rest of the section is not a study of ESAP per se, but of the relationship between school programs and student achievement. The objective here is to tie together various parts of the reanalysis. Certain educational programs have been shown to be associated with favorable noncognitive teacher and student outcomes, and some of these, in turn, have been found to be related to student achievement. The analysis reported here aims at relating programs, noncognitive variables, and achievement in an effort to explain performance levels with the available data on schools, teachers, and students.

ANALYSIS OF OUTLIERS

The first task is to define the outlying districts: those where students in experimental schools did either much better or much worse than those in control schools. The sample was divided into the four race-sex groups because the purpose of the analysis was to examine interactions related to these two variables. Then, a school-pair was eliminated where there were fewer than four students of a given race or

sex. This meant different school-pairs were eliminated from the four subgroups. The next series of choices concerned the measure of effectiveness. In the end, the simplest was used: the unadjusted difference between the means for experimental and control school students. An extensive examination of adjusted difference scores revealed that the same schools were located in the upper and lower ends of these distributions as were in the two tails of the distributions of unadjusted differences. The four distributions are displayed in Fig. 1.

Straightforward examination suggests that there are all kinds of departures from normality, but definitive judgments are hard or impossible because of the small number of cases. The cases at the extremes of the distributions could well be manifestations of a low density spread of considerable importance. In hope of bringing these unusual properties to light I reconstructed them using grouped rather than continuous data (Fig. 2). These two representations of the distributions suggest several things. First, there are outliers at both ends for white males, and possibly positive outliers for white females. In addition, these distributions appear to have two modes. The distribution for black females piles up at the left hand side; only the distribution for black males seems roughly normal.

The next step was to decide on a cutoff point for the definition of outliers. This necessarily involves judgment, one consideration being that a cutoff too far from the mean is going to yield a very small number of cases. Guided by this practical consideration, I set a limit such that about 10 percent of the cases would be identified as either positive or negative outliers for a given distribution. Cases over 1.5 standard deviations above the mean for each distribution were characterized as instances of highly effective ESAP treatments and cases 1.5 standard deviations below the mean as instances of highly effective control schools. In a normal distribution these limits would separate 13.4 percent of the sample. These limits are marked on Fig. 1 and districts above and below the cutoffs labeled by their serial number. Eight cases fall in the low end of the distribution, seven in the high end. Counting each appearance in the outlying group separately, there are an average of just under five cases in each distribution in the high or low

Each point represents a school-pair.
Positive differences indicate experimental school scored higher.
By race/sex groups.
At least four students per school.

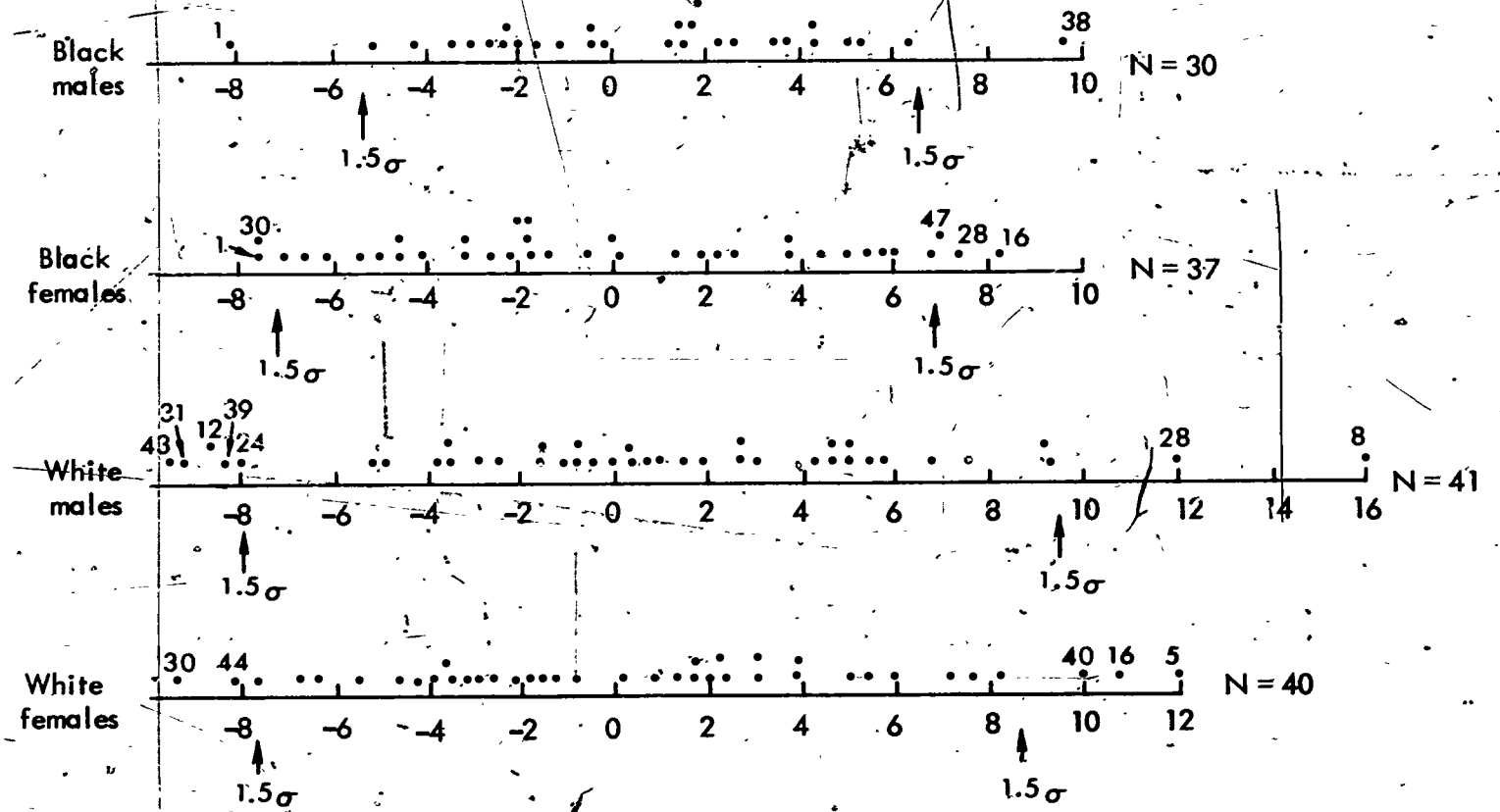

Fig. 1— Distributions of raw difference scores between experimental and control schools on achievement test score

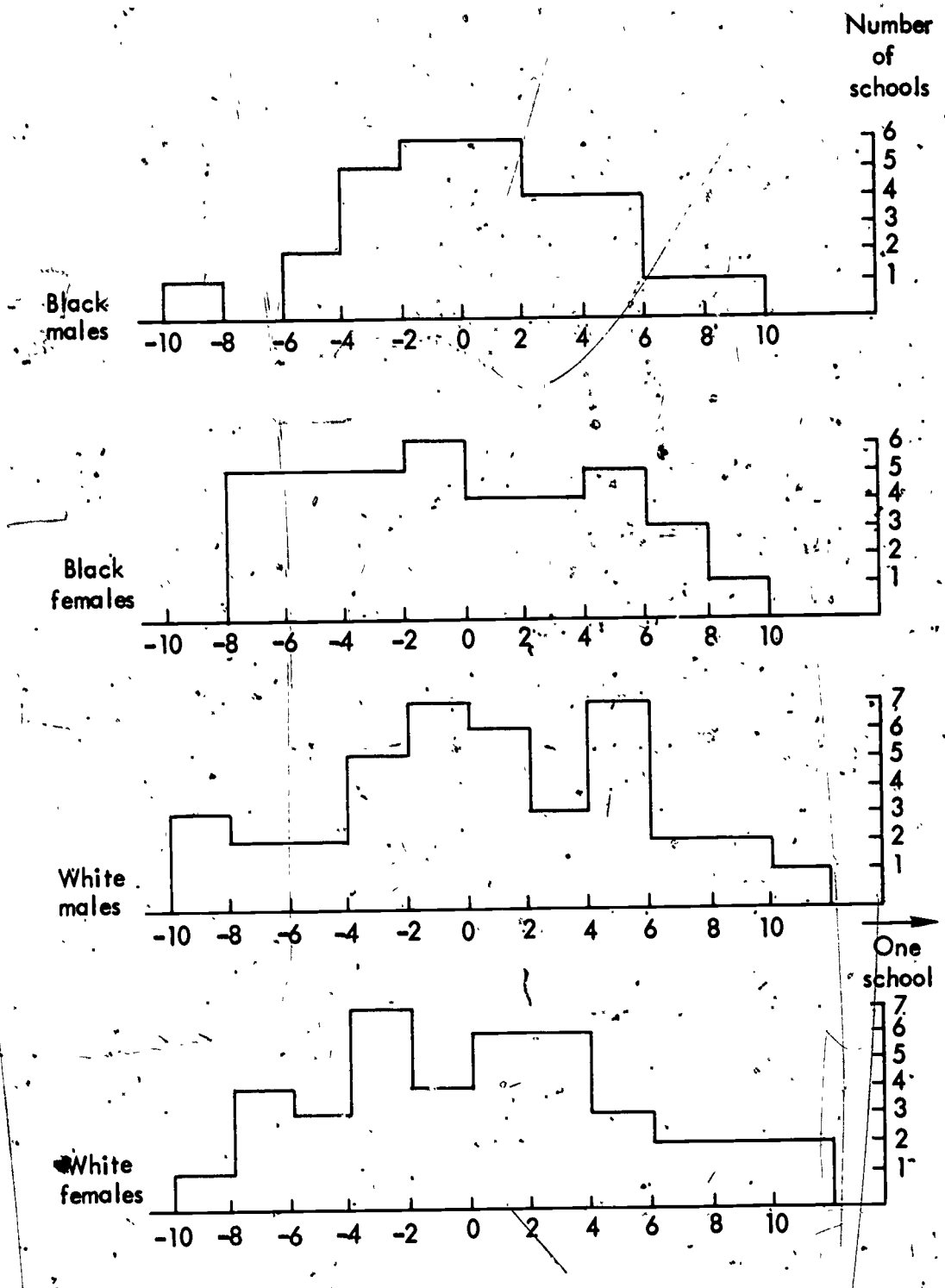

Fig. 2 — Graph of raw difference scores using grouped data

category, almost exactly 10 percent of the total number of school pairs. This is slightly under the expected fraction of the distribution. Four districts appear in more than one distribution, implying that where a district is highly effective--or ineffective--for one subgroup of students it may also be effective--or not--for another. However, if these four distributions were totally independent of one another, some of this consistency would be expected by chance. In fact, the number of schools observed in more than one distribution is below that expected for four independent distributions.

The main question is whether these outlying cases can be differentiated from one another. If they can, it might suggest reasons for the unusual success or failure of ESAP. To start with, comparisons were made between high and low outliers on a number of school characteristics such as the size, the percent of white students, the recency of desegregation, the attitude of the principal to desegregation, the per pupil expenditure, and so forth. All available indices of school-to-school differences were used. Very few showed high and low outliers were clearly different from one another. The exception is that positive outliers tend to be smaller than negative outliers (a difference of over 100 students), and they tend to be located in somewhat rural areas (established using the percent of people in the county living in communities of over 2,500). High outlying districts had fewer schools (21) than low outliers (28) and fewer students in all (about 11,000 and 15,000).

These differences are not statistically significant but they may not be fortuitous either. There are several reasons for thinking a program like ESAP would be more successful in small school districts and small schools. For example, these districts may be more manageable so that extra federal funds do not get lost. Administrators might be more careful to see the money gets used for the right purpose. Another possibility is that rural school districts are more peaceful than urban ones and therefore provide a more stable setting in which innovations can be implemented. Or, the families students come from may be more homogeneous in rural areas, and this may mean there are fewer cross-pressures to contend with when changes take place in schools.

Some of these issues can be pursued with the data. For example,

principals' reports on the level of disruption of the school can be used to check the idea that smaller, rural school districts are more peaceful places. One of these measures (frequency with which students' lockers are broken into) showed the expected difference, but the same pattern of results was not found consistently for the other variables. Another index of disruption, absenteeism rate, also showed high outliers were better off than low outliers. So, taking a selective approach to the data, the case can be made that the low outliers, where control school students did better than experimental school students, were more disorganized places. Of course, this does not prove that disorganization led to poor educational performance; the evidence does not warrant such a strong conclusion.

The other suggestions were even harder to test adequately. There was no sensible measure of how readily these schools can be controlled, either by the administration or by the principal. The closest approximations were the principals' assessment of how much effect they can have on students, and information about the superintendents' action in helping schools with the process of desegregation. Obviously, neither can be taken seriously as a measure of subtle and complicated realities.

The next question is whether the high and low outliers can be differentiated in terms of educational programs and activities. For example, are the high outliers more likely to have remedial education programs? The simplest possibility is that the low outliers were districts where the experimental design had been vitiated. That is, ESAP might not have been delivered to the experimental school, or it might have been delivered to the control school. Of the eight low outliers, there was only one district where the control school had received ESAP funding. This explanation does not seem terribly helpful, therefore, since one of these six deviant cases would be expected to show up in the lower outlier group on probabilistic grounds.

Still, the low outliers might be those districts that received a smaller ESAP grant (in the experimental school) than the high outliers. However, comparison of the two groups showed that the low outliers actually received slightly larger ESAP grants. This could be explained by the fact that these schools tended to be larger; when the ESAP grant

was expressed as a fraction of the total school budget the difference between high and low groups disappeared. Consistent with this pattern, the low outliers were found to have larger numbers of programs and activities as reported by the ESAP director or principal. But this too could be explained largely in terms of the difference in size of the two types of schools. When these tallies of the number of programs were adjusted for school size, the differences between the two groups either vanished or were reduced with one exception. A cumulative score for ESAP-funded activities indicated by the director shows high outliers had more activities even when the size of school had been taken into account.

While this makes good sense, and is encouraging, there is a danger of taking crude cumulative scores too literally. Equating programs of different kinds, as one does in making a simple cumulative score, is obviously a hazardous assumption. And no account is taken of the differences in size or funding of these activities. Thus a school that spends all its money on a remedial education program will have a "lower" score than one that decides to do five different things with the grant. Nevertheless, this finding suggested a close inspection of the ESAP-funded programs in these schools to see if there were particular strategies associated with success or failure. Accordingly, the two groups were compared for differences in the things that ESAP money had bought. In general, no pattern emerged from this analysis. The high outliers were neither more nor less likely to have used money on remedial or tutoring programs. They did not concentrate money, any more than the other group, on inservice training. Nor did they use it on special kinds of inservice training, such as training in race relations. There are no special resources or equipment found in the high and low groups. And it does not seem that extra personnel, such as teachers' aides, are more common in the successful schools. Nevertheless, even though high and low outliers do not seem to have distinctive types of programs, on one cumulative measure the high outliers turn out to be better supplied by ESAP.

To summarize, the high and low outlier groups are different from one another, but not in terms of the kinds of educational programs that

they have. They are different in terms of location and size; low outliers tend to be larger and located in urban areas. This may hint at the importance of the setting for ESAP rather than the substance of the program itself. The key to success may not be the content of the program so much as the way in which the program is implemented. This idea fits with the fact that ESAP regulations allowed school districts a good deal of latitude in devising their version of ESAP. It might be expected, therefore, that the skill with which local administrators adapt ESAP to their local needs would make a difference to the outcome of the program. But this is a guess.

EDUCATIONAL PROGRAMS, NONCOGNITIVE OUTCOMES, AND STUDENT ACHIEVEMENT

Earlier in this section I looked at unusually effective and ineffective instances of ESAP. Here I look at all the data in an attempt to explain variations in achievement by differences in educational programs and variations in students' attitudes. There are two parts to the analysis.

(1) ESAP was shown to have an effect on certain measures of students' attitudes; that is, students in experimental schools had more, or less, favorable attitudes than those in control schools. But these effects were more evident for white students than for blacks. If the program had an effect at all, it would seem that the white students were the ones affected. Most puzzling, the effect on whites was positive for males and negative for females.

Since ESAP had little effect on these noncognitive outcomes for black males, they did not look like promising intervening variables to explain the linkage between program funding and achievement scores. However, it is possible that the program affected black male achievement levels by changing white students' attitudes. The program could have thereby created a less threatening environment for black male students so that they felt better about their school work and performed more effectively on tests. This assumes that white students' attitudes are an important ingredient in determining the quality of black male students' experience in school. For example, ESAP was associated with more positive attitudes for white males in the way they felt about desegregated

schools and in the degree to which they interacted with black students. Perhaps these changes made it easier for black males to feel accepted, to feel less threatened, and consequently to improve their potential for educational performance in school. A complication enters at this point because while ESAP was associated with more favorable attitudes for white females on some of the measures examined in Table 2, most showed a reverse effect: Control students scored above experimental students. The aggregate measures of these latter variables would be expected to have a negative association with black male achievement; at least that would be consistent with the view that students' influence on one another's attitudes is not dependent on sex.

Aggregate scores were computed for white males and females separately, so that each school had an aggregate score on those noncognitive outcomes that was significantly associated with the presence of the program. These aggregate scores could then be correlated with the individual level achievement test scores for black males. Table 7 reports these correlations (first column). Even with the large number of cases involved, only one is statistically significant. However, there is some support for the idea of aggregate effects in the pattern of signs of these correlations. With a single exception, the correlations fall in line with expectations. Where ESAP is associated with a positive effect on student attitudes (all the white male attitudes and two of the white female variables), the correlations reported in Table 7 are positive. Where ESAP is associated with a negative effect on student attitudes (all but two of the white female variables) the correlations are negative. Taking the analysis one step further, a measure of student's social background was introduced. The suspicion was that this might substantially change the uncontrolled relationships reported in the first column of Table 7. The partial regression coefficients (achievement dependent variable, independent variable indicated in left hand margin, mother's education controlled) are reported in the second column. Though it is true that in three cases the signs of the correlations are changed, in general the introduction of this control did not alter the conclusion derived from the first set of results. The zero order relationship is not to be explained away in terms of student social background.

Table 7

RELATION BETWEEN AGGREGATE MEASURES OF WHITE STUDENT ATTITUDES AND BLACK MALE ACHIEVEMENT

Aggregate Measures for White Males	r	b
If student had talked to counselor during the year	.013	-.007
Student would prefer to attend a racially mixed school	.019	-.007
Student would like more friends of the other race	.060	.047
Student had helped another of the other race with homework	-.007	-.016
Student had asked another of the other race for help with own homework	.041	.040
Aggregate Measures for White Females		
If student had talked to counselor during the year	.013	-.001
Student says she is glad to go to school in the morning	-.026	-.016
Student says that when punished it is for no good reason	-.029	-.038
Student thinks principal's attitude to desegregation is favorable	-.051	-.052
Teachers seen as being unfair to white students	-.042	-.018
White students reported as complaining of favoritism on part of teachers	-.067	-.055
Black students reported as complaining of favoritism on part of teachers	-.099*	-.092*

Column 1. Zero order correlations between aggregate measures of white student attitudes and individual black male achievement scores. Aggregate measures of white students' attitudes are those for which a statistically significant difference exists between experimental and control school students (see Table 2).

Column 2. Standardized regression coefficient for measure of white student attitude with mother's education controlled.

* = significant at the .05 percent level.

It must be remembered that the correlations are not large; they account for a vanishingly small percentage of the variance in achievement. In a separate multiple regression, all five white male attitude variables were entered as independent variables and were found to account for an additional 0.5 percent of the variance in achievement over and above that explained by mother's education. The comparable analysis for the white female attitude variables had similar results; an additional 1.3 percent of the variance in achievement can be attributed to the variation in the complete set of independent variables. These results must be considered in light of the fact that only a small percentage of the variance in achievement (under 1 percent) can be accounted for by ESAP. Therefore any intervening effects are going to have to be smaller still. Seen this way, such tiny effects become more important.

(2) The analyses presented in the second part of Section VI showed that certain programs were associated with differences in noncognitive student outcomes. Specifically, inservice training for teachers that emphasized race relations and programs that helped intergroup relations among students were associated with more favorable student attitudes and more positive self-reported behavior. However, these effects were again concentrated among white students, so it did not seem likely that the key to explaining the program effect on black male achievement would be found here. Nevertheless, the inquiry was pursued, broadening the objective to include all four subgroups of students in a general investigation of the relations among program type, noncognitive outcome, and student achievement score.

Table 8 reports the analysis of the two program type variables, inservice training focused on race relations and intergroup programs for students. The four race-sex groups were analyzed separately for two reasons; the relationship between noncognitive outcome and student achievement had been found to vary across these four groups, and the effect of these two educational programs on the noncognitive outcomes had also been shown to differ depending on student race and sex. Several noncognitive variables are chosen for the analyses, the choice being based on results obtained in Sections IV and VI. Is the

Table 8

ANALYSIS OF RELATION BETWEEN PROGRAM TYPE, SELECTED STUDENT NONCOGNITIVE VARIABLES AND STUDENT ACHIEVEMENT SCORES, BY RACE-SEX GROUP

Intervening Variable		Black Males	Black Females	White Males	White Females
Part A: *Program Type: Inservice Training with Emphasis on Race Relations*					
If student had talked to a student of the other race on the phone	r	.03	.07	.09	.08
	b-1	.03	.07	.09	.08
	b-2	.02	.06	.07	.04
	R^2	.05	.06	.06	.11
If student could choose, he would prefer a racially mixed school	r	.14	.19	.19	.15
	b-1	.14	.19	.19	.15
	b-2	.14	.17	.16	.12
	R^2	.06	.09	.09	.12
If student is a member of school club or sports team	r	.07	.15	.15	.24
	b-1	.07	.14	.15	.25
	b-2	.04	.12	.12	.20
	R^2	.05	.07	.07	.15
Student would like more friends of the opposite race	r	.08	.08	.19	.08
	b-1	.08	.07	.18	.08
	b-2	.09	.08	.17	.06
	R^2	.05	.07	.09	.12
Student had asked for help with homework from student of other race	r	.08	.05	.08	-.02
	b-1	.08	.06	.08	-.02
	b-2	.07	.05	.08	-.02
	R^2	.05	.06	.07	.11
Part B: *Program Type: Programs in Intergroup Relations for Students*					
Student's rating of his ability compared with others in his class	r	.32	.39	.46	.45
	b-1	.32	.39	.47	.45
	b-2	.30	.37	.43	.39
	R^2	.14	.19	.24	.25
Student had called a student of the other race on the phone	r	.03	.07	.09	.08
	b-1	.03	.07	.10	.08
	b-2	.02	.05	.07	.05
	R^2	.05	.06	.06	.11

Table 8--continued

Intervening Variable		Black Males	Black Females	White Males	White Females
Part B: *Program Type: Programs in Intergroup Relations for Students* (continued)					
If student could choose, would go to a racially mixed school	r	.14	.19	.19	.15
	b-1	.15	.19	.19	.16
	b-2	.14	.17	.17	.12
	R^2	.07	.09	.09	.13
Student is member of school club or team	r	.07	.15	.15	.24
	b-1	.07	.14	.15	.24
	b-2	.05	.12	.12	.19
	R^2	.05	.07	.07	.15
Student would like more friends of the opposite race	r	.08	.08	.19	.08
	b-1	.08	.08	.19	.09
	b-2	.09	.08	.18	.07
	R^2	.06	.06	.09	.12

Note: Table reports zero order correlation between noncognitive variable (indicated in left-hand margin) and achievement, the standardized partial regression coefficient for this relationship when program type is introduced into the equation (b-1), the same coefficient when mother's education is also entered in the equation, and the value of R^2 when all these variables are included in the equation.

association between student attitude and performance to be accounted for by the presence of either of these two educational programs? Table 8 contains four pieces of information. First, it shows separately, by subgroup and by noncognitive variable, the direct relationship (zero order correlation) between achievement and the noncognitive variable. Below this is the standardized regression coefficient for the noncognitive variable obtained when the dummy variable identifying program type is also entered in the equation. Underneath that is another estimate of the same relationship, also a standardized regression coefficient, when mother's education is entered as an additional control. The fourth piece of information is the value of R^2 for the whole equation.

Neither of these two programs explains the direct relation between the noncognitive variable and achievement. This can be seen by comparing the first number (r) in each cell with the second (b-1). Although to varying extents students' educational performance can be attributed to their noncognitive differences, no case can be made that these relationships are due to the presence or absence of the most promising program types identified by previous analysis. The effect on the partial relationship of introducing the control for social background is a good deal more pronounced than the effect of introducing program type as the intervening variable (r compared with b-2 in each cell). The control for social background takes account of the degree to which these kinds of educational programs are located in schools with students of different levels of educational performance, and it also takes account of the relationship between background and achievement. Social background appears to be a good deal more important in accounting for the relationships between the noncognitive variables and achievement than are the measures of program type.

In summary, these two programs appear to be associated with positive and desirable outcomes, which are consistent with the general aims and intentions of ESAP: inservice training programs with emphasis on intergroup relations and programs in intergroup relations for students. However, although the changes these programs brought about in students are associated with their educational performance, there is little basis for attributing that association to the program type. Therefore, these

two programs might be judged in terms of their effects on certain non-cognitive outcomes, but they do not appear to provide the explanation of the higher achievement levels of any one of the four subgroups of students, including black male 10th graders. Of course, it should be remembered these programs are not necessarily ESAP-funded.

VIII. REVIEW

In reviewing the analysis I put aside the qualifications and caveats scattered through the text, but I do so knowing a shortened summary is vulnerable to misinterpretation. The review should be seen as the most positive and optimistic rendition of the results. A more cautious account will be found in the text, and the reader is directed there also for information about the size of the effects being discussed.

(1) Although ESAP does not alter the racial climate of the school in general, it does change the attitudes and behaviors of some groups of students, notably white males and black females, in a positive direction. These students are more likely to feel they belong in school, are happier about going there in the morning, interact more with students of the other race, and would prefer a racially mixed school. These are the kinds of dispositions and attitudes that ESAP tried to develop and can be seen either as ends in themselves or as means to the end of achieving successful desegregation.

(2) Measures of students' attitudes and behavior are associated with their achievement levels, and these relationships are stronger or weaker depending on the student's race and sex. There is no simple summary of all these differences, but achievement is somewhat related to whether the black student thinks race is linked to IQ and thinks other blacks are smart. This is suggestive of the importance of self-image in academic success. Unfortunately, ESAP does not appear to have been responsible for changing black students' attitudes in this key area.

(3) A good deal of ESAP activity was devoted to improving intergroup relationships, both among teachers and among students. Significantly, both kinds of programs are associated with more favorable attitudes and behavior in schools. For instance, teachers in schools with inservice training that concentrated on intergroup relations were more likely to favor desegregation, felt that other teachers' attitudes were also more positive, and felt more open to express themselves to

their students on the matter of race. Students in schools that had programs designed to help relations between the races were also better disposed to racially mixed schools, interacted more with students of the other race, and perceived their teachers as having more favorable attitudes to desegregation. Unfortunately, it was not possible to link these two program types to ESAP; the analysis cannot differentiate between schools where these programs are funded by ESAP and schools where they are funded from other sources.

(4) Some school districts were especially successful; there, the experimental school students did much better than the control school students on an achievement test. There were also some districts where the reverse was true. A comparison of these two extremes suggested that the best implementations of ESAP were in fairly small schools, situated in the smaller, rural school districts. This may mean that ESAP's success is determined to a degree by the local setting.

BIBLIOGRAPHY

NORC (National Opinion Research Center), *Southern Schools: An Evaluation of the Effects of the Emergency School Assistance Program and of School Desegregation*, Chicago, November 1973.

RMC (Resource Management Corporation), *Evaluation of the Emergency School Assistance Program: Final Report*, UR-163 (3 Volumes), Bethesda, Maryland, 1971.